Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.
Identification and characterisation of an exported immunogenic protein of *Mycobacterium avium* subspecies *paratuberculosis*

A thesis presented in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at Massey University, Palmerston North, New Zealand

Christine Dupont

2002
MASSEY UNIVERSITY
APPLICATION FOR APPROVAL OF REQUEST TO EMBARGO A THESIS
(Pursuant to AC98/168 (Revised 2), Approved by Academic Board 16.02.99)

Name of Candidate: Christine Dupont I.D. Number: 9911979
Degree: PhD Dept / Institute / School: IUAAS
Thesis Title: Identification and Characterisation of an Exported Immunogenic Protein of M. avium Subspecies paratuberculosis
Name of Chief Supervisor: Dr Alan Murray Telephone Ext: 7845

As author of the above named thesis, I request that my thesis be embargoed from public access until (date) 31st December 2003 for the following reasons:

☐ Thesis contains commercially sensitive information.
☐ Thesis contains information which is personal or private and/or which was given on the basis that it not be disclosed.
☐ Immediate disclosure of thesis contents would not allow the author a reasonable opportunity to publish all or part of the thesis.
☐ Other (specify): 

Please explain here why you think this request is justified:

A.J. Park patent lawyers have requested that the embargo be extended from Oct 2003 to Dec 2003 to facilitate a PCT (International Patent Application)

____________________________________________________
Signed (Candidate): ___________________________ Date: ______________________________

Endorsed (Chief Supervisor): Alan Murray Date: 11/16/2003

Note: Copies of this form, once approved by the representative of the Vice Chancellor, must be bound into every copy of the thesis.
MASSEY UNIVERSITY
APPLICATION FOR APPROVAL OF REQUEST TO EMBARGO A THESIS
(Pursuant to AC98/168 (Revised 2), Approved by Academic Board 16.02.99)

Name of Candidate: CHRISTINE DUPONT
I.D. Number: 99119179

Degree: PhD
Dept / Institute / School: IVABS

Thesis Title: Identification and characterisation of an exported immunogenic protein of Mycobacterium avium subspecies paratuberculosis

Name of Chief Supervisor: Dr Alan Murray
Telephone Ext: 7895

As author of the above named thesis, I request that my thesis be embargoed from public access until (date) October 2003 for the following reasons:

☐ Thesis contains commercially sensitive information.
☐ Thesis contains information which is personal or private and / or which was given on the basis that it not be disclosed.
☐ Immediate disclosure of thesis contents would not allow the author a reasonable opportunity to publish all or part of the thesis.
☐ Other (specify):

________________________________________
Please explain here why you think this request is justified:

part of the work is being included in a patent application which is currently being prepared.
The patent applicants will be Massey University and the Pasteur Institute in Paris.

Signed (Candidate): CHRISTINE DUPONT Date: 22.03.02

Endorsed (Chief Supervisor): Alan Murray Date: 22.03.02

Approved / Not Approved (Representative of VC): __________________________ Date: 23/03/02

Note: Copies of this form, once approved by the representative of the Vice Chancellor, must be bound into every copy of the thesis.
Abstract

Exported proteins of mycobacteria are available to interact with the immune system at an early stage of infection and are potent inducers of immune responses. Potentially exported proteins of *Mycobacterium avium* subspecies *paratuberculosis* were identified using alkaline phosphatase gene fusion technology. A library of partial gene fusions from a New Zealand clinical isolate of *M. a. paratuberculosis* was constructed in the shuttle vector pJEM11 and expressed in the surrogate hosts *E. coli* and *M. smegmatis*. The DNA inserts from a portion of the resulting clones expressing alkaline phosphatase-positive fusion proteins were partially sequenced to identify the proteins. Eleven proteins not previously described for *M. a. paratuberculosis* were identified as containing signal sequences for export. One of these, a putative lipoprotein named P22 was selected for further study. The full nucleic acid sequence of the *p22* gene was determined and the open reading frame was cloned into the mycobacterial expression vector pMIP12. This enabled P22 to be produced as a polyhistidine-tagged protein in *M. smegmatis* and facilitated purification by chromatography. N-terminal sequencing of the recombinant protein confirmed cleavage of an N-terminal signal sequence. Native P22 was detected in culture supernatants and cell sonicates of *M. a. paratuberculosis* strain 316F using rabbit antibody raised to P22. Investigation of the presence of genes similar to *p22* in other mycobacterial species, revealed *p22* was present in *Mycobacterium avium* subspecies *avium* and similar genes existed in *M. intracellulare* (88.5% identity) and *M. saofulaceum* (87.7% identity). Database searches showed P22 belonged to the LppX/LprAFG family of mycobacterial lipoproteins also found in *M. leprae* and in members of the *M. tuberculosis* complex. P22 shared less than 75% identity to these proteins. Recombinant P22 was able to elicit significantly increased interferon-gamma secretion in blood from a group of eight sheep vaccinated with a live, attenuated strain of *M. a. paratuberculosis* (strain 316F) compared to a group of five unvaccinated sheep. Antibody to P22 was detected by Western blot analysis in 10 out of 11 vaccinated sheep, in two out of two clinically affected cows and in 11 out of 13 subclinically infected cows.
Acknowledgements

Much of the work reported in this thesis was made possible by the assistance of many people and service providers. DNA sequencing was done by Lorraine Berry at the Massey DNA sequencing facility. N-terminal sequencing of P22 was done by Trevor Loo at Massey University Protein Sequencing Services. Faecal culture and ELISA tests for the cattle used, were carried out by AgResearch, Wallaceville. The serum samples from these cattle were generously donated by Dr. Cord Heuer and Solis Norton. M. a. paratuberculosis DNA samples used for PCR analysis of strain distribution were a gift from Dr. Desmond Collins. The experimental sheep were housed at Massey University Agricultural Services. I'd especially like to thank Margaret Brown and Geoff Warren for wrangling sheep and collecting blood samples and watching over the animals that were so important to this study. The mycobacterial shuttle vectors pJEM11 and pMIP12, were kindly supplied by Professor Brigitte Gicquel, Unite de Genetique Mycobacterienne, Pasteur Institute, Paris, France. Sue Copland assisted with the Southern blots used in this study and Dr. Jeremy Rae contributed in the preparation of M. a. paratuberculosis culture filtrates and isocitrate dehydrogenase assays. Support for this work was provided by Meat New Zealand, WoolPro and the IVABS post-graduate student support fund.

The people who have influenced me in the decision to embark on a career in scientific research deserve more thanks than I can mention here. My chief supervisor, Dr. Alan Murray has provided unconditional guidance and support over the years. He has a gift of always finding some positive aspect in every result and I could always count on him for encouragement through rough times. My second supervisor Dr. Keith Thompson always managed to find time for me (and my sheep). His wisdom and experience as a scientist and teacher has been greatly appreciated and admired. My third supervisor Dr. John Tweedie gave essential criticism and advice in regards to this thesis and was always willing to help and support me. Dr. Catherine Day introduced me to the wonderful world of protein expression and encouraged me in my decision to undertake PhD studies, even though it meant losing her only technician. My lab-mates Dr. Jane
Oliaro, Dr. Sandy Maclachlan, Miho Minamikawa and Dr. Becky Davies have been excellent company over the years.

Special thanks must be said to members of my family whose support has been a significant factor in my enjoyment of this work. Because of my insistence in pursuing this path, my family in Canada have suffered the (temporary) loss of their daughter and grand children. Despite this, they constantly offered their support. My husband Dr. John Lumsden who quietly encouraged me through this endeavour and always managed to help me keep things in perspective. Last but not least, my children Jamie and Genna who handled my "absence" with great maturity. I'm sure all the nights of take-aways for dinner contributed to their understanding.

I would also like to express my thanks to the sheep used in this study. Approval for the use of experimental animals was granted by the animal ethics committee of Massey University, protocols 99/155 and 00/87. Permission to carry out genetic manipulations of *E. coli* and *M. smegmatis* was granted by Massey University Genetic Technology Committee under Section 40 of the Hazardous Substances and New Organisms Act 1996, GTC application GM099/MU/179.

Construction of the *M. a. paratuberculosis* PhoA fusion library and characterisation of SodC from this study have been published (Dupont, C. & Murray, A. (2001). Identification, cloning and expression of *sodC* from an alkaline phosphatase gene fusion library of *Mycobacterium avium* subspecies *paratuberculosis*. *Microbios* **106** S1: 7-19).
# Table of Contents

Abstract.................................................................................................................................i
Acknowledgements..................................................................................................................ii
List of Figures ...........................................................................................................................ix
List of Tables .............................................................................................................................xi
List of Abbreviations ................................................................................................................xii

## Chapter 1 Literature review

1.1 History.................................................................................................................................1
1.2 Prevalence and economic impact.........................................................................................1
1.3 Host range ...........................................................................................................................3
  1.3.1 Crohn's disease .............................................................................................................3
1.4 Classification .......................................................................................................................5
1.5 Clinical signs and transmission .........................................................................................6
1.6 Control ...............................................................................................................................8
  1.6.1 Test and cull ..................................................................................................................8
  1.6.2 Vaccination ..................................................................................................................8
  1.6.3 Management ...............................................................................................................10
1.7 Detection ...........................................................................................................................11
  1.7.1 Culture .......................................................................................................................11
  1.7.2 Detection of DNA ........................................................................................................13
   1.7.2.1 IS900 insertion element .........................................................................................13
  1.7.3 Immunological tests ......................................................................................................15
   1.7.3.1 Tests for cellular responses ..................................................................................16
   1.7.3.2 Tests for humoral responses ...............................................................................16
  1.7.4 Histopathological detection .........................................................................................17
1.8 The mycobacterial envelope and its relationship to pathogenicity and immunology .........18
  1.8.1 Lipoarabinomannan ....................................................................................................19
1.9 Export of proteins in mycobacteria .....................................................................................21
  1.9.1 Export pathways and signal peptides ..........................................................................23
   1.9.1.1 Cleavage of signal peptides .................................................................................25
   1.9.1.2 Lipidation ..............................................................................................................25
   1.9.1.3 C-terminal anchoring .........................................................................................26
1.10 Searching for exported proteins of mycobacteria ..............................................................27
1.11 Components isolated from *M. a. paratuberculosis* ..........................................................29
## Chapter 2  General materials and methods

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Bacterial strains and plasmids</td>
<td>41</td>
</tr>
<tr>
<td>2.1.1 Bacterial strains</td>
<td>41</td>
</tr>
<tr>
<td>2.1.2 Plasmids</td>
<td>42</td>
</tr>
<tr>
<td>2.2 Bacterial growth and storage conditions</td>
<td>43</td>
</tr>
<tr>
<td>2.2.1 E. coli</td>
<td>43</td>
</tr>
<tr>
<td>2.2.2 M. smegmatis</td>
<td>43</td>
</tr>
<tr>
<td>2.2.3 M. a. paratuberculosis</td>
<td>44</td>
</tr>
<tr>
<td>2.3 DNA isolations</td>
<td>44</td>
</tr>
<tr>
<td>2.3.1 Isolation of plasmid DNA from E. coli</td>
<td>44</td>
</tr>
<tr>
<td>2.3.2 Isolation of genomic DNA from mycobacterial species</td>
<td>44</td>
</tr>
<tr>
<td>2.4 DNA manipulations and cloning procedures</td>
<td>45</td>
</tr>
<tr>
<td>2.4.1 Restriction endonuclease digestions</td>
<td>45</td>
</tr>
<tr>
<td>2.4.2 Electrophoresis</td>
<td>45</td>
</tr>
<tr>
<td>2.4.3 Extraction from agarose gels</td>
<td>46</td>
</tr>
<tr>
<td>2.4.4 Southern blotting and hybridisations</td>
<td>46</td>
</tr>
<tr>
<td>2.4.4.1 DNA probe preparation</td>
<td>47</td>
</tr>
<tr>
<td>2.4.4.2 Removal of probe from Southern blots</td>
<td>47</td>
</tr>
<tr>
<td>2.4.5 Polymerase chain reactions</td>
<td>47</td>
</tr>
<tr>
<td>2.4.5.1 Primer design</td>
<td>47</td>
</tr>
<tr>
<td>2.4.5.2 PCR conditions</td>
<td>48</td>
</tr>
<tr>
<td>2.4.6 DNA sequencing</td>
<td>49</td>
</tr>
<tr>
<td>2.4.7 Ligations</td>
<td>49</td>
</tr>
<tr>
<td>2.5 Bacterial transformations</td>
<td>50</td>
</tr>
<tr>
<td>2.5.1 Preparation of electrocompetent M. smegmatis</td>
<td>50</td>
</tr>
<tr>
<td>2.5.2 Transformation of E. coli</td>
<td>50</td>
</tr>
<tr>
<td>2.5.3 Transformation of M. smegmatis</td>
<td>51</td>
</tr>
<tr>
<td>2.6 Protein isolations</td>
<td>51</td>
</tr>
<tr>
<td>2.6.1 Preparation of cell lysates</td>
<td>51</td>
</tr>
</tbody>
</table>
2.6.2 Preparation of culture filtrates ................................................................. 52
  2.6.2.1 M. smegmatis .................................................................................. 52
  2.6.2.2 M. a. paratuberculosis ................................................................. 52
2.6.3 Ni²⁺-affinity chromatography ................................................................. 52
2.6.4 Size exclusion chromatography ............................................................. 53
2.7 Protein analyses ......................................................................................... 54
  2.7.1 Polyacrylamide gel electrophoresis ..................................................... 54
  2.7.2 Western blotting and immunodetections .............................................. 54
    2.7.2.1 Immunodetection of Western blots ............................................... 55
    2.7.2.2 Removal of antibody from Western blots ..................................... 55
  2.7.3 N-terminal protein sequencing ............................................................ 55
  2.7.4 Estimation of protein concentration and molecular weight .................. 56
2.8 IFN-γ assays ............................................................................................. 57
2.9 Rabbit details and immunisations ............................................................. 57
  2.9.1 Preparation of P22 for immunisation .................................................. 57
  2.9.2 Rabbit immunisation protocol ............................................................. 58
2.10 Sheep details and immunisations ............................................................ 58
  2.10.1 Experimental sheep and immunisation protocol .................................. 58
  2.10.2 Naturally infected sheep ....................................................................... 59
2.11 Naturally infected cattle .......................................................................... 60
2.12 Bioinformatics ......................................................................................... 60

Chapter 3 Identification of M. a. paratuberculosis DNA sequences encoding exported proteins

3.1 Abstract ....................................................................................................... 62
3.2 Introduction .................................................................................................. 63
3.3 Materials and methods .............................................................................. 65
  3.3.1 Construction of an M. a. paratuberculosis pJEM11 expression library ...... 65
    3.3.1.1 Extraction of DNA from M. a. paratuberculosis ......................... 65
    3.3.1.2 Preparation of pJEM11 vector DNA ......................................... 66
    3.3.1.3 Partial digestion with Sau3A of M. a. paratuberculosis DNA ...... 66
    3.3.1.4 Ligation of M. a. paratuberculosis DNA and pJEM11 and transformation into E. coli ................................................................. 66
    3.3.1.5 Plasmid isolation from the E. coli recombinant library ............... 67
    3.3.1.6 Transformation of the recombinant plasmids into M. smegmatis mc²155... 67
  3.3.2 Sequencing of DNA inserts encoding putative exported proteins .......... 68
3.4 Results.........................................................................................................................69
3.4.1 Construction of an *M. a. paratuberculosis* pJEM11 expression library .................69
3.4.1.1 Confirmation of *M. a. paratuberculosis* DNA for cloning ................................69
3.4.1.2 Subcloning of *M. a. paratuberculosis* DNA into the vector pJEM11 and expression of the library in *E. coli* ..............................................................69
3.4.1.3 Expression of the library in *M. smegmatis* ......................................................70
3.4.2 Analysis of *M. a. paratuberculosis* phoA fusions .............................................71
3.4.2.1 Sequencing of DNA inserts encoding putative exported proteins .................71
3.4.2.2 Analysis of phoA fusions ..............................................................................71
3.5 Discussion ................................................................................................................77

Chapter 4 Cloning, heterologous expression and characterisation of an immunogenic 22 kDa protein from *M. a. paratuberculosis*

4.1 Abstract.....................................................................................................................83
4.2 Introduction ...............................................................................................................84
4.3 Materials and methods ............................................................................................86
4.3.1 PCR amplification of the *p22* gene from *M. a. paratuberculosis* ....................86
4.3.2 Cloning of the *p22* open reading frame .............................................................86
4.3.3 Expression and purification of P22 recombinant protein from *M. smegmatis* ........88
4.3.3.1 Western blot analyses of P22 recombinant protein ....................................88
4.3.4 Preparation of rabbit antibody raised to P22 ......................................................89
4.3.5 PCR amplification of the p22 ORF from genomic DNA ......................................89
4.4 Results.......................................................................................................................90
4.4.1 Sequence analysis of plasmid pTB-16 and identification of the p22 open reading frame .........................................................................................................................90
4.4.1.1 Sequence analysis of pTB-16 ....................................................................90
4.4.1.2 Identification of the p22 ORF ....................................................................91
4.4.2 Sequence analysis of p22 ..................................................................................91
4.4.2.1 Sequence similarities between P22 and a family of mycobacterial lipoproteins .................................................................................................................................91
4.4.3 Cloning and expression of the p22 ORF ............................................................93
4.4.4 Purification of recombinant P22 from cell lysates ..............................................94
4.4.5 Analysis of *M. smegmatis* culture filtrates for the presence of recombinant P22 .........................................................................................................................94
4.4.6 Immune responses to P22 .................................................................................94
<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1. Relationships between <em>Mycobacterium</em> species based on 16S rRNA sequence homology</td>
<td>6a</td>
</tr>
<tr>
<td>Figure 1.2. General relationship of tests and immune responses to infection with <em>M. a. paratuberculosis</em> over time</td>
<td>16a</td>
</tr>
<tr>
<td>Figure 1.3. Common features of a signal peptide</td>
<td>24a</td>
</tr>
<tr>
<td>Figure 3.1. Schematic of the <em>E. coli/Mycobacterium</em> shuttle vector pJEM11</td>
<td>64a</td>
</tr>
<tr>
<td>Figure 3.2. Schematic representation of the location of the oligonucleotide primers designed for sequencing the <em>M. a. paratuberculosis</em> inserts in the pJEM11 constructs</td>
<td>68a</td>
</tr>
<tr>
<td>Figure 3.3. Schematic representation of the construction of the PhoA fusion library</td>
<td>69a</td>
</tr>
<tr>
<td>Figure 3.4. PCR amplification of <em>M. a. paratuberculosis</em> genetic elements</td>
<td>69b</td>
</tr>
<tr>
<td>Figure 3.5. PhoA+ recombinant <em>E. coli</em> colonies from the <em>M. a. paratuberculosis</em> pJEM11 library</td>
<td>69c</td>
</tr>
<tr>
<td>Figure 3.6. PhoA+ recombinant <em>M. smegmatis</em> colonies from the <em>M. a. paratuberculosis</em> pJEM11 library</td>
<td>70a</td>
</tr>
<tr>
<td>Figure 3.7. Restriction endonuclease analysis of selected PhoA+ clones</td>
<td>71a</td>
</tr>
<tr>
<td>Figure 4.1. Schematic of the <em>Mycobacterium</em> expression vector pMIP12 for the production of histidine-tagged recombinant proteins in <em>M. smegmatis</em></td>
<td>86a</td>
</tr>
<tr>
<td>Figure 4.2. Sequence analysis of the p22 ORF</td>
<td>91a</td>
</tr>
<tr>
<td>Figure 4.3. Amino acid sequence comparison between P22 of <em>M. a. paratuberculosis</em> and database search results</td>
<td>91b</td>
</tr>
<tr>
<td>Figure 4.4. Kyte-Doolittle plot (top) and signal sequence features (bottom) of the P22 precursor protein</td>
<td>92a</td>
</tr>
<tr>
<td>Figure 4.5. Comparison of the promoter regions of <em>M. bovis lpp-27</em> and <em>M. a. paratuberculosis</em> p22</td>
<td>92b</td>
</tr>
<tr>
<td>Figure 4.6. Restriction endonuclease digest of plasmid pMIP-p22</td>
<td>93a</td>
</tr>
<tr>
<td>Figure 4.7. Expression of recombinant P22 from <em>M. smegmatis</em></td>
<td>93b</td>
</tr>
</tbody>
</table>
Figure 4.8. Affinity chromatography of recombinant P22 ............................................. 94a
Figure 4.9. Detection of recombinant P22 from *M. smegmatis* culture filtrates......... 94b
Figure 4.10. Detection of antibody to P22 in sheep vaccinated with Neoparasec..........94c
Figure 4.11. Detection of antibody to P22 in individual sheep from a naturally infected flock. ........................................................................................................................................................................... 95a
Figure 4.12. Detection of antibody to P22 in naturally infected cattle. ...................... 95b
Figure 4.13. IFN-γ induction using Ni^{2+}-affinity-enriched P22 in Neoparasec- vaccinated sheep blood.......................................................................................... 96a
Figure 4.14. IFN-γ induction by purified recombinant P22 in Neoparasec- vaccinated sheep blood.................................................................................................. 97a
Figure 4.15. Detection of antibody to P22 from sheep vaccinated with
*M. a. paratuberculosis* strain 316F culture filtrate. .................................................. 97b
Figure 4.16. Western blot detection of rabbit antibody raised to P22.......................... 98a
Figure 4.17. Detection of native P22 in Western blots of *M. a. paratuberculosis*
strain 316F cell fractions and comparison to recombinant P22 using rabbit antibody raised to P22. ........................................................................................................ 98b
Figure 4.18. PCR amplification of the *p22* gene from 13 isolates of
*M. a. paratuberculosis*. ............................................................................................ 99a
Figure 4.19. PCR amplification from 22 mycobacterial strains using primers designed
to the *p22 ORF*. .................................................................................................... 99b
Figure 4.20. Southern blot analyses using a *p22* probe from genomic DNA of 13
mycobacterial strains.................................................................................................. 100a
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>A summary of the clinical stages of paratuberculosis</td>
</tr>
<tr>
<td>1.2</td>
<td>Components isolated from <em>M. a. paratuberculosis</em></td>
</tr>
<tr>
<td>2.1</td>
<td>Bacterial strains used in this study</td>
</tr>
<tr>
<td>2.2</td>
<td>Plasmids used in this study</td>
</tr>
<tr>
<td>2.3</td>
<td>Antibiotics and supplements used in microbiological media</td>
</tr>
<tr>
<td>2.4</td>
<td>Sheep treatment groups</td>
</tr>
<tr>
<td>3.1</td>
<td>Oligonucleotide primers designed for analysis of <em>M. a. paratuberculosis</em> DNA</td>
</tr>
<tr>
<td>3.2</td>
<td>Identification of selected <em>M. a. paratuberculosis</em> PhoA fusion proteins</td>
</tr>
<tr>
<td>4.1</td>
<td>Summary of results for detection of <em>M. a. paratuberculosis</em> by serum ELISA, faecal culture and P22 Western blot analysis</td>
</tr>
</tbody>
</table>
List of Abbreviations

\( A_{280\text{nm}} \) absorbance at 280 nm
ATCC American type culture collection
Avian PPD Purified protein derivative from \textit{M. a. avium}
BCG bacillus Calmette-Guerin
BCIP 5-bromo-4-chloro-3-indolyl phosphate
BLAST basic local alignment search tool
ConA concavalin A
dTTP deoxythymidine triphosphate
dUTP deoxyuridine triphosphate
DIG digoxigenin
DNA deoxyribonucleic acid
EDTA ethylenediamine tetraacetic acid
ELISA enzyme-linked immunosorbent assay
HPLC high pressure liquid chromatography
IFN-\( \gamma \) interferon-gamma
Johnin PPD purified protein derivative from \textit{M. a. paratuberculosis}
kan kanamycin
kb kilobase pairs
kDa kilodalton(s)
LAM lipoarabinomannan
LB Luria-Bertani
OD optical density
ORF open reading frame
PBS phosphate-buffered saline
PCR polymerase chain reaction
PhoA alkaline phosphatase
POD peroxidase
PVDF polyvinylidene difluoride
RBS ribosome binding site
SDS-PAGE sodium dodecyl sulphate-polyacrylamide gel electrophoresis
TAE Tris-acetate, EDTA
UV ultraviolet

Amino acids

A alanine \quad C cysteine
D aspartic acid \quad E glutamic acid
F phenylalanine \quad G glycine
H histidine \quad I isoleucine
K lysine \quad L leucine
M methionine \quad N asparagine
P proline \quad Q glutamine
R arginine \quad S serine
T threonine \quad W tryptophan
V valine \quad Y tyrosine
<table>
<thead>
<tr>
<th>Nucleic acids</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>adenosine</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>thymidine</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>cytidine</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>guanosine</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 1  Literature review

1.1  History

Johne’s disease was first described as early as 1826 as a chronic enteritis of cattle that could not be attributed to any known cause of diarrhoea. At that time, the disease was called pseudotuberculosis because of its remarkable similarity to intestinal tuberculosis, but without caseous necrosis. The name Johne’s disease comes from work carried out in 1895 by Johne and Frothingham who demonstrated a connection between enteritis in cattle and the presence of acid-fast micro-organisms in intestinal mucosa (Johne & Frothingham, 1895; cited by Chiodini, 1992). In 1910, Twort cultivated and characterised a mycobacterium that, in 1914, was shown to produce experimental enteritis in cattle (Twort, 1911; cited by Clarke, 1997). The organism was named, at the time, *Mycobacterium enteritidis chronicae pseudotuberculosa bovis Johne* (Twort & Ingram, 1912; cited by Chiodini, 1992). In 1932, it was designated as a distinct species named *Mycobacterium paratuberculosis*, although synonyms such as *M. johnei* were often used, and the disease was renamed paratuberculosis (Chiodini et al., 1984a; Kreeger, 1991). The organism has been reclassified as *Mycobacterium avium* subspecies *paratuberculosis* (Thorel et al., 1990b).

In New Zealand, Johne’s disease was first diagnosed in 1912 in an imported cow (O'Hara, 1983). However, it was not until 1952 that Johne's disease was reported in sheep (Williamson & Salisbury, 1952). It is currently one of the world’s most widespread bacterial diseases of domestic animals.

1.2  Prevalence and economic impact

The impact of Johne’s disease on the world economy is considered to be enormous (Chiodini et al., 1984a; Merkal et al., 1987; Pandey et al., 1987; Scott-Orr et al., 1988; de Lisle & Milestone, 1989; Behymer et al., 1991). Most of the losses have been estimated for the dairy industry, for which there have been more analyses done on the impacts of the disease. Accurate economic analyses are difficult to make because of the prolonged course of infection and the predominantly subclinical nature of the disease. There is a lack of data regarding production losses due to subclinical Johne’s disease in sheep, goats, beef cattle and deer to estimate the economic impact of subclinical disease in these animals. Estimates of the costs of subclinical infection in dairy herds take into account reduced milk
production (Buergelt & Duncan, 1978; Abbas et al., 1983a; Benedictus et al., 1987; Hutchinson, 1988; de Lisle & Milestone, 1989; Kormendy et al., 1989; Wilson et al., 1993; Nordlund et al., 1996), higher mortality rates among infected herds (Merkal et al., 1987; McNab et al., 1991; Johnson-Ifearulundu et al., 1999), and lower carcass values due to a decrease in mean body weight (Johnson-Ifearulundu et al., 1999). Studies have also shown paratuberculous cattle are culled earlier than their herdmates, resulting in increased costs of replacing these cattle (Buergelt & Duncan, 1978; Kormendy et al., 1989; Wilson et al., 1993). Higher replacement rates have also been found to occur in the Madrid region of Spain in small ruminant herds with high seroprevalences (Mainar-Jaime & Vazquez-Boland, 1998). There have been some reports of increased mastitis and impaired reproductive performance in infected animals (Merkal et al., 1975; Abbas et al., 1983a). Others have shown no association between infection and infertility or mastitis (de Lisle & Milestone, 1989; Wilson et al., 1993).

In many countries, such as the USA and New Zealand, where herds are not routinely tested for *M. a. paratuberculosis* infection, the number of herds officially known to have infection is thought to be a gross underestimate of the actual level of infection. Estimates in New Zealand, based on unofficial reports from veterinary practitioners, suggest that approximately 60% of dairy herds are infected with *M. a. paratuberculosis*. In sheep, this prevalence may be as high as 70%. The level of disease in farmed deer is thought to be around 7% (Brett, 1998). Furthermore, the reported prevalence of infected animals is a reflection of the diligence with which veterinarians and owners look for the disease. In Wisconsin, USA, approximately 33% of dairy herds show evidence of the presence of one or more infected animals (Collins et al., 1994). In Victoria, Australia, approximately 20% of dairy herds are estimated to be infected (Wraight et al., 1999) and up to 8% of Australian sheep flocks are believed to be infected, with the disease continuing to spread (Royal Society of New Zealand Daily News, 2001).

In the USA, direct and indirect losses of greater than $200/cow is suggested for infected herds. Nation-wide, the economic impact approximately ten years ago was estimated to be as much as 1.7 billion US dollars annually (Chiodini et al., 1984a; Thoen & Baum, 1988; Jones, 1989). A recent survey estimated that up to 21.6% of US dairy herds are infected (National Animal Health Monitoring System Survey, 1997). In New Zealand, total estimated losses based on known prevalence rates are around $5 million per year. Unofficial estimates of disease bring this figure to approximately $57 million per year. Most of this loss is in dairy cattle and of this, 52% is attributed to lower milk production (Brett, 1998).
1.3 Host range

*M. a. paratuberculosis* has a wide host range including many wild and domestic species of sheep, goats, deer, camelids and cattle (Williams *et al.*, 1983a; Chiodini *et al.*, 1984a; Cocito *et al.*, 1994). Wild animals that may serve as reservoirs of infection include white-tailed deer (Libke & Walton, 1975; Chiodini & Van Kruiningen, 1983), bighorn sheep and mountain goats (Williams *et al.*, 1979; 1983b), red deer (Sharp, 1997), elk (Jessup *et al.*, 1981), antelope (Dukes *et al.*, 1992), bison (Ellingson *et al.*, 1999), moose (Soltys *et al.*, 1967), rabbits (Angus, 1990; Greig *et al.*, 1997), foxes, stoats, weasels and badgers (Beard *et al.*, 2001). Infections can also be established in primates (McClure *et al.*, 1987). Asymptomatic shedders include horses (Larsen *et al.*, 1972), and pigs (Larsen *et al.*, 1971). Experimentally, a number of species can be successfully infected and be shown to shed the organism for long periods of time. These include hamsters, gerbils, guinea pigs, mice, rats, chickens, lemmings and pigeons (Greig *et al.*, 1999). The lesions and clinical signs produced in these hosts are not always characteristic of Johne’s disease. More recently, reports of the culture of *M. a. paratuberculosis* from nematodes (Whittington *et al.*, 2001) further suggest the presence of this organism is much more widely spread than previously thought.

*M. a. paratuberculosis*, like other obligate pathogenic mycobacteria, is not free living in the environment. Its presence in water or soil samples is assumed to be from deposition from infected faeces. The bacterium is very resistant to the effects of heat, cold, sunlight and dehydration compared to other pathogenic bacteria such as *Listeria* and as such it may survive for nine months or longer in soils (Kopecky, 1977; Beeman *et al.*, 1989; Grant *et al.*, 1996; Sung & Collins, 1998).

1.3.1 Crohn’s disease

Crohn’s disease is a chronic gastroenteritis of people that closely mimics Johne’s disease in terms of gross and microscopic pathology, although there are differences in extraintestinal manifestations (Van Kruiningen, 1999). The disease is thought by many to be the result of an abnormal immune response to unspecified and possibly non-specific antigens of unknown origin, resulting in the upregulation of inflammatory mediators. A survey of intestinal biopsy samples from Crohn’s patients has found significant numbers of mixed bacterial flora, including *M. a. paratuberculosis* (Tiveljung *et al.*, 1999). This supports the idea that secondary bacterial colonisers may enter primary lesions where they then elicit the chronic inflammatory syndrome characteristic of Crohn’s disease. Although the aetiology may be
multifactorial, *M. a. paratuberculosis* has received the most attention as a possible cause of the disease. Other agents that have been implicated in Crohn’s disease include *Listeria sp.*, *streptococci* and viruses such as those causing measles, rubella, mumps and herpes (Romagnani *et al.*, 1997). The incidence of the disease appears to be increasing (Munkholm *et al.*, 1992; Hermon-Taylor & Bull, 2002).

The first suggested similarity between Crohn’s disease and Johne’s disease was in 1913 (Dalziel, 1913). In 1932, the human disease was distinguished from intestinal tuberculosis by Crohn and co-workers (1932). Since then, *M. a. paratuberculosis* has been successfully cultured from a small number of Crohn’s disease patients and in these cases has been found to exist as spheroplast-like forms, which revert to bacillary forms upon extended culture (Chiodini *et al.*, 1984b; 1984c; 1986; Gitnick *et al.*, 1988; Wall *et al.*, 1993; Schwartz *et al.*, 2000). The most successful isolation of *M. a. paratuberculosis* from Crohn’s disease tissue has been by the use of Mycobacterial Growth Indicator Tube (MGIT) culture (Schwartz *et al.*, 2000). In this study, positive cultures were confirmed in 10 of 27 (37%) Crohn’s disease tissue specimens, compared to two of 36 (5.6%) control tissues (one of the control tissues was from a healthy donor and the other patient had non-specific colitis). Results from this study also suggested that positive culture and PCR results were obtained more successfully using surgically resected tissue that includes submucosa rather than using biopsy material containing only the surface mucosa. The first successful isolation of the bacillary form of *M. a. paratuberculosis* from a person was from an enlarged cervical lymph node of a young boy (Hermon-Taylor *et al.*, 1998). The fact that this patient went on to develop Crohn’s disease five years later and a resected tissue specimen tested positive for *M. a. paratuberculosis* by PCR (Hermon-Taylor, personal communication cited by Chamberlin *et al.*, 2001), has been suggested as strong evidence that *M. a. paratuberculosis* causes Crohn’s disease. In any case, it is the first proven instance of *M. a. paratuberculosis* causing disease in a human being.

Inoculation of various animal species with human isolates has failed to result in textbook Crohn’s disease as described for human beings. Furthermore, the pathological similarities between Johne’s disease and Crohn’s disease are considered to be outweighed by the differences and it is therefore argued that there is not an aetiology in common (Van Kruiningen *et al.*, 1986; 1991). Conversely, it may be argued that genetic differences in host species result in the manifestation of different pathological lesions. For example, infection with cattle strains of *M. a. paratuberculosis* in some mouse types, results in atypical
manifestations of disease, compared to those seen in cattle (Chiodini & Buergelt, 1993; Veazey et al., 1995).

Because of the difficulty in detecting spheroplast forms of the bacterium, PCR of the IS900 element (see 1.7.2.1) has been employed to identify the presence of M. a. paratuberculosis in people. There is wide variation in the reported occurrence of M. a. paratuberculosis in both Crohn’s and non-Crohn’s patients, using this technique. The numerous reports have been summarised by Quirke (2001). The reported frequencies for the tissues tested range as follows: Crohn’s 0-100%; ulcerative colitis 0-61%; non-inflammatory bowel disease 0-87.5%. Higher frequencies for IS900 PCR products from Crohn’s disease tissue compared to non-inflammatory bowel disease tissue, were claimed in eight of these studies, and equal (0%) or slightly lower frequencies were claimed by seven others. The contradictions are likely reflections of the variation in experimental design used, differences in the tissue samples collected, the techniques used for DNA isolation from these tissues, and the PCR conditions used. The high proportion of studies reporting no PCR products from any of the tissue samples is of concern as it may be argued that the samples used were not amenable to PCR and therefore represent false negatives. Until accepted, standardised methods and experimental designs are employed, the presence or absence of M. a. paratuberculosis DNA cannot be unequivocally established. Interestingly, amplification of IS900-containing sequences by PCR in cultures of human milk have been obtained from two lactating mothers with Crohn’s disease whereas none of the control samples from healthy, lactating individuals were positive for IS900 (Naser et al., 2000). Regardless of whether or not M. a. paratuberculosis causes Crohn’s disease, the accumulated studies have shown that exposure to this organism is not an uncommon occurrence among people and therefore the presence of this pathogen in animals is of concern to human beings.

1.4 Classification

Mycobacteria range from widespread innocuous inhabitants of soil and water to organisms that are pathogenic for many animal species, including people. They are non-motile rods that are defined on the basis of a distinctive staining property known as acid-fastness. This property is due to the presence of a lipid-rich cell wall that makes them relatively impermeable to various basic dyes, but once stained they resist decolourisation with organic solvents.

Mycobacteria belong to the order Actinomycetales, family Mycobacteriaceae. The G + C content of most mycobacterial DNA species ranges from 62 to 70% (Goodfellow et al.,
Figure 1.1. Relationships between *Mycobacterium* species based on 16S rRNA sequence homology. Adapted from Rogall et al. (1990).
Base composition of *M. a. paratuberculosis* DNA is 66 to 67% G+C (McFadden et al., 1987; Imaeda et al., 1988), and the genome size is $4.4 \times 10^6$ to $4.7 \times 10^6$ base pairs. *M. a. paratuberculosis* belongs to the so-called MAC complex (*Mycobacterium avium* complex), which includes the species *M. xenopi, M. lepraemurium* and the various strains of *M. avium* and *M. intracellulare*. These organisms are all slow growing, which is characteristic for pathogenic mycobacteria. MAC contains three subspecies of *M. avium*: *M. avium* subsp. *silvaticum*, *M. avium* subsp. *avium*, and *M. avium* subsp. *paratuberculosis*. Classification as subspecies for these strains is based on a large number of biochemical tests (Thorel et al., 1990b). Classification as a subspecies is further warranted by studies using DNA technology. DNA/DNA hybridisation studies and 16S rRNA sequencing studies show most *M. a. paratuberculosis* strains share nearly 100% identity to *M. a. avium* strains (Hurley et al., 1988; Imaeda et al., 1988; Saxegaard & Baess, 1988; Yoshimura & Graham, 1988; Rogall et al., 1990; Stahl & Urbance, 1990). Thus neither of these tools are able to distinguish between *M. a. avium* and *M. a. paratuberculosis*. Figure 1.1 shows the relationships of various mycobacterial species based on 16S rRNA sequence homology. Restriction fragment length polymorphism (RFLP) analyses, which is considered to be the most powerful tool for distinguishing between closely related strains, has shown that *M. a. paratuberculosis* is most closely related to *M. a. avium* serovar 2 (Chiodini, 1990). This organism, also known as *M. avium* strain 18, was previously thought to be an *M. a. paratuberculosis* strain before the application of RFLP analyses (Thorel et al., 1990a; Chiodini, 1993). RFLP analyses have attempted to differentiate between ovine and bovine strains of *M. a. paratuberculosis*. RFLP analysis using the IS1311 element (Marsh et al., 1999) and an IS900 multiplex PCR (Bull et al., 1999) have been reported to discriminate between these host strains. Conversely, analyses using the IS900 element with the restriction endonuclease *Bst* EII, has not led to a consensus in this regard (Collins et al., 1990; Pavlik et al., 1999).

### 1.5 Clinical signs and transmission

The bulk of evidence suggests that the main source of infection of young animals is faeces-contaminated pasture or bedding and/or infected colostrum or milk (Cocito et al., 1994). To date, there is no evidence that *M. a. paratuberculosis* is sexually transmitted (Larsen & Kopecky, 1970; Merkal et al., 1982) and the isolation of the bacilli from reproductive tracts of both male (Larsen & Kopecky, 1970; Larsen et al., 1981) and female (Rohde & Shulaw, 1990) cattle appears to be a result of dissemination of the bacterium from its primary site of infection in the gut (Chiodini et al., 1984a). *M. a. paratuberculosis* has infrequently been
isolated from foetuses of infected cows (McQueen & Russell, 1979; Seitz et al., 1989; Sweeney et al., 1992b). Experimental infection of adult cattle show that they may develop the disease, but young animals less than 30 days old are most susceptible (Taylor, 1953; Brotherston et al., 1961; Payne & Rankin, 1961; Nisbet et al., 1962; Larsen et al., 1975; Clarke, 1997). The Peyer's patch, which is the targeted site of uptake of M. a. paratuberculosis, is maximally developed in young ruminants and progressively involutes as the animal ages (Nisbet et al., 1962; Larsen et al., 1975; Reynolds & Morris, 1983; Clarke, 1997). This may account for the decrease in susceptibility of adult animals.

Infected animals pass through one or more of the three recognised clinical stages of the disease. These are summarised in Table 1.1. Stage I, is asymptomatic and shedding of the organism is undetectable. Stage II is also largely asymptomatic but shedding is present although usually intermittent. During this time the concentration of the organism in the intestinal mucosa and lumen progressively increases and intermittent diarrhoea may occur. Stage III is the terminal stage and is symptomatic and large numbers of bacteria are present in the faeces. Signs visible in cattle in the terminal phase (stage III) include intractable chronic diarrhoea, progressive emaciation, decreased milk production, anaemia, diffuse oedema and infertility. Animals die in a cachectic state (Cocito et al., 1994). Clinical signs may be precipitated by parturition, lactation and other stresses (Chiodini et al., 1984a). Progression of disease, based on experimental infections of young cattle, show most will advance to stage II, but only a small proportion will progress to stage III. Reports from the state of Victoria in Australia show that the number of clinical cases per year in infected dairy herds is less than or equal to one percent. This situation is thought to be similar in New Zealand (Brett, 1998). Such animals generally show clinical disease only after two to five years (Stuart, 1965; Lepper et al., 1989). For this reason, it is suggested that for every clinical case on a farm, there are approximately 20 additional infected animals. This is sometimes referred to as the "iceberg effect" for which supporting data exists (Whitlock, 1992; Whitlock et al., 1999).

In other ruminants, the clinical signs are similar with a few exceptions. Diarrhoea is not always evident in affected sheep and goats (Morin, 1982; Gezon et al., 1988; Carrigan & Seaman, 1990; Hietela, 1992; Cranwell, 1993; Cetinkaya et al., 1994; Benazzi et al., 1995; Clarke et al., 1996). In sheep, the fleece is often of poor quality and easily broken (Cranwell, 1993). Similar observations have been reported in llamas and alpacas (Belknap et al., 1994; Ridge et al., 1995). The course of disease in farmed deer is particularly acute with very young animals, sometimes under one year of age, being clinically affected (Gilmour & Nyange, 1989).
Table 1.1  A summary of the clinical stages of paratuberculosis

<table>
<thead>
<tr>
<th>Stage</th>
<th>Mycobacterial shedding</th>
<th>Clinical signs</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>undetectable</td>
<td>absent</td>
</tr>
<tr>
<td>II</td>
<td>medium and intermittent</td>
<td>largely absent, intermittent diarrhoea</td>
</tr>
<tr>
<td>III</td>
<td>high (10^{10}/g faeces)</td>
<td>present</td>
</tr>
</tbody>
</table>

Adapted from Whitlock et al. (1986) and Cocito et al. (1994).

1.6  Control

Eradication of Johne’s disease is difficult due to the inability of current diagnostic tests to detect all infected animals. The most successful disease control strategies include combinations of vaccination, test and cull and good herd management practices.

1.6.1  Test and cull

Test and cull programmes involve regular testing of all stock and the subsequent culling of infected animals. It is considered to be the most expensive option in disease control. Tests which only detect clinical cases or faecal shedders are of limited value as it has been demonstrated that removing these animals from the herd does not necessarily have an impact on the prevalence of infection (Sherman & Gezon, 1980; Thomas, 1983).

1.6.2  Vaccination

Commercially available vaccines include attenuated strains of *M. a. paratuberculosis* (eg. those used in the UK, New Zealand, Denmark and Norway) and heat-killed preparations (eg. those used in Australia, USA, Iceland, Netherlands and India). All strains used for vaccines are of bovine origin (MacDiarmid, 1989). The vaccines currently used in New Zealand for sheep, cattle and goats are Neoparasec (Merial, France) and Gudair (CZ Veterinaria, Porrino, Spain). Neoparasec is an attenuated bovine strain (316F) and Gudair is a heat-killed
Chapter 1

preparation. Vaccination with live or dead preparations appears to be equally effective in protection from disease. Vaccination does not prevent or eliminate infection but reduces the severity of clinical signs and faecal shedding (Larsen et al., 1978; Kalis et al., 2001). Acid-fast organisms within the injection site nodules can be found with both the live and heat-killed vaccines. It is interesting that sterilisation by heat does not seriously disrupt the cell wall structure and attests to the well-known tenacity of this organism. There is some evidence that the attenuated strain of the vaccine can be distributed to other parts of the body following subcutaneous inoculation, including lymph nodes and the gastrointestinal tract (Larsen et al., 1977; Collett & West, 2001; Kohler et al., 2001).

Vaccination of infected lambs with the attenuated vaccine triggers strong humoral and cellular responses and leads to reduced mycobacterial numbers in lesions, compared to unvaccinated infected controls (Gwozdz et al., 2000a). Vaccination of adult sheep and goats in known infected flocks has shown this to be an effective strategy for control of disease (Crowther et al., 1976; Corpa et al., 2000). Strong evidence of the effectiveness of vaccination against ovine Johne’s disease comes from Iceland where a vaccination programme was adopted in the 1950s. At this time, the disease was widespread and annual mortalities were high, averaging 8-9%. Vaccination has significantly reduced the economic importance of this disease (reviewed by MacDiarmid, 1989). Similarly, in Norway, vaccination of goats has reduced the prevalence of infection from 53% to 1% (Saxegaard & Fodstad, 1985). In New Zealand, experience has shown that after three years of vaccinating replacement stock with the attenuated vaccine, the prevalence of Johne’s disease is noticeably reduced and after five years the disease is almost non-existent (Brett, 1998). Even with penalties imposed on vaccinated carcasses, vaccination is still considered to be economically viable in most situations (Brett, 1998).

Vaccination has several drawbacks. Localised vaccination site reactions occur commonly in animals vaccinated with live or killed Johne’s vaccines (Collett & West, 2001). There are a few published studies on the presence of lesions resulting from vaccination (reviewed by Collett & West, 2001), and it is widely reported by veterinary practitioners and owners that nodules, especially in vaccinated cattle, can be quite large (>20 cm in diameter). These nodules can go on to rupture and discharge from the skin surface. This causes concern among farmers regarding the appearance of vaccinated animals because of a perceived effect on the sale of such animals. Macroscopically and histologically the lesions are consistent with those of mycobacterial infections (Milestone, 1989; 1992). The extent and size of the lesions varies widely, and can involve areas of the body distant from the vaccination site.
Lesions may also persist for the life of the animal and may involve the draining prescapular lymph node, even though there may be no visible injection site lesion (Milestone, 1989; Walker, 1991; Johnstone, 1992). Such granulomatous lesions in the lymph nodes of some vaccinated animals can be confused with lesions caused by bovine tuberculosis at slaughter and can lead to carcass condemnation. In human beings, where accidental vaccination has occurred, granulomas form at the injection site and persist for years causing serious discomfort and, in some cases, have required amputation of the affected digit (Whitlock et al., 1986; MacDiarmid, 1989).

Vaccination with *M. a. paratuberculosis* interferes with tuberculosis testing programmes because of immunological cross-reactions. A large percentage of vaccinated animals give positive results to bovine tuberculin testing, although the percentage of animals testing positive declines with time. These animals can usually be differentiated from tuberculosis-infected cattle by a further comparative tuberculin test (Milestone, 1989) which involves a comparison of the skin test reactions to avian and bovine tuberculosis tuberculins. However, it has been demonstrated that some cattle produce positive or doubtful skin reactions as late as two years after vaccination (Kohler et al., 2001). In some countries, the use of the vaccine is restricted to herds or areas declared free of tuberculosis. Furthermore, live vaccinated animals or those submitted for slaughter may be restricted for export (Brett, 1998).

1.6.3 Management

Recommended management practices include good hygiene, removing young stock from potential sources of infection as early as possible, segregation or removal of infected stock and obtaining infection-free replacements (Rossiter & Burhans, 1966; Merkal et al., 1975; Thomas, 1983; Collins et al., 1994). For the dairy industry, isolated calf-rearing is recommended. It is strongly felt by many investigators that control schemes involving vaccination and/or test and cull are only effective when used in combination with good management and hygiene practices (Hernandez et al., 1999; Kalis et al., 1999). Further control of the disease requires the trading of only *M. a. paratuberculosis*-free animals. This requires recognition of infected animals as early as possible, a feat that would require very sensitive and specific diagnostic tests. As yet, no test is available to accurately prove uninfected status in individual animals on a single examination. For this reason, it may be more appropriate to test for the presence of *M. a. paratuberculosis* at the herd level.
1.7 Detection

Detection of infection in the host is influenced by the stage of disease (Ridge et al., 1991; Sweeney et al., 1995). In the lengthy subclinical stage, the bacterium produces little or no detectable humoral response and the number of organisms is typically low, making direct detection of the bacterium difficult. It is important to note that the immune response between species and between individual animals can vary considerably.

1.7.1 Culture

In cattle, culture of *M. a. paratuberculosis* from faeces or tissue is currently the most accurate means of detecting infection and is considered the standard against which all other tests are compared (Stabel, 1998; Whittington et al., 1999). The success of culture is linked to the presence of sufficient numbers of *M. a. paratuberculosis* being shed from the intestine (Collins, 1996; Perez et al., 1996; Whittington et al., 1999). The sensitivity of faecal culture is very high (>95%) in clinically affected animals (Whittington & Sergeant, 2001) and overall sensitivities of 30% to 50% have been reported for cattle, based on repeated testing (Whitlock et al., 1999). In sheep, culture sensitivities of 98% have been reported for sheep with multibacillary disease and 48% for those with paucibacillary disease (Whittington et al., 1999).

Identification of *M. a. paratuberculosis* in culture is primarily based on its slow growth (up to four months) (Thoen & Baum, 1988; Sanftleben, 1990), and dependency on an external source of mycobactin (Chiodini et al., 1984a; Thorel et al., 1990a). Mycobactin is an iron-chelating compound produced by most mycobacteria *in vitro* with the apparent exception of *M. a. paratuberculosis*. There is also evidence that mycobactin is not synthesised by *M. a. paratuberculosis* *in vivo* (Lambrecht & Collins, 1993). Some workers have found variations in mycobactin-dependency using different culture media (Aduriz et al., 1995). The finding of some *M. a. avium* strains to be mycobactin-dependent further confounds identification using mycobactin-dependence (Matthews & McDiamid, 1977; Barclay & Ratledge, 1983; Thorel, 1984). In spite of this, mycobactin-dependency, together with slow growth, continue to be key features for identification of isolates of *M. a. paratuberculosis*.

Improved sensitivity of faecal culture has been obtained employing technical modifications using centrifugation, sedimentation and filtration. The disadvantage of culture is the long incubation time required for the appearance of colonies. Routinely, samples are cultured on
Herrod's egg yolk media, or less frequently on Lowenstein-Jensen media, with mycobactin supplementation. When suspected colonies appear, a Ziehl-Neelsen (acid-fast) stain is performed and the colonies are subcultured to the same media type, with and without mycobactin in order to confirm this requirement. Alternatively, *M. a. paratuberculosis* can be confirmed through PCR to identify the presence of the species-specific IS900 genetic element. Unfortunately, strains that predominantly infect sheep and goats worldwide, the so-called "ovine" strains, do not readily grow on conventional solid media such as Herrod's egg yolk or Lowenstein-Jensen media (Howarth, 1932; Dunkin & Balfour-Jones, 1935; Taylor, 1945; Carrigan & Seaman, 1990; Juste *et al.*, 1991; Dukes *et al.*, 1992; Collins *et al.*, 1993; Aduriz *et al.*, 1995; Benazzi *et al.*, 1995). Culture sensitivity of ovine strains on solid media has been reported to be only about 10% (Carrigan & Seaman, 1990). Consequently, histopathologic identification has been considered to be the standard for diagnosis of paratuberculosis in sheep (Whittington *et al.*, 1999). Recently, however, there has been success in culturing these strains in liquid media (BACTEC and MGIT systems (Becton Dickinson)), and on Middlebrook 7H11/OADC agar. Detection by culture, particularly of tissues, has been shown to be more sensitive than histopathology in some cases (Whittington *et al.*, 1999; 2000).

Liquid culture systems can detect positive samples sooner than solid media culture. This can be as early as nine days for cattle shedding large numbers of *M. a. paratuberculosis* (Damato & Collins, 1990) and four to seven weeks (vs. 12 to 16 weeks) for those shedding low numbers of organisms. BACTEC detection medium relies on metabolic utilisation of radiolabeled palmitic acid for the detection of 14C-labeled CO2, which signals the presence of growing bacteria. The medium is very low in nutrient content and is not optimal for the growth of most bacteria. For *M. a. paratuberculosis* culture, mycobactin is added. Bovine strains are readily grown using this system, and some ovine strains can be isolated this way. Non-radiometric systems are also available. It has been found that radiometric culture is more sensitive for the detection of *M. a. paratuberculosis* from infected cattle than growth on Herrod's egg yolk medium (Damato & Collins, 1990; Whittington *et al.*, 1998).

Microscopic detection of *M. a. paratuberculosis* in faeces depends on the number of bacteria present in the sample. This is typically low in early disease and high in clinically affected animals (Ris *et al.*, 1988; Kormendy, 1990). *M. a. paratuberculosis* is not homogeneously distributed in faeces, but is present in small clusters, contributing to false negatives from low faecal shedders upon microscopic examination and faecal culture (Visser, 1999).
\textit{M. a. paratuberculosis} has been cultured from milk from both clinically and subclinically infected cows. Similar to faecal culture, culture sensitivity for milk has been shown to be greater in clinically affected cows that were also shedding large numbers of bacteria (Sweeney \textit{et al.}, 1992a; Streeter \textit{et al.}, 1995). It appears that the number of organisms present in milk is generally lower than in faeces (Sweeney \textit{et al.}, 1992a; Cocito \textit{et al.}, 1994) and therefore very sensitive methods are required for detection in this medium. Culture has generally been shown to be more sensitive than PCR in detecting \textit{M. a. paratuberculosis} (Giese & Ahrens, 2000). New methods are being developed to increase the sensitivity of detection in milk. One of these methods uses immunomagnetic separation. When IS900 PCR was directly coupled to this procedure as few as \(10^3\) cfu/ml could be detected (Grant \textit{et al.}, 2000). Two samples of commercially pasteurised cows milk and one sample of bulk milk tested positive using this procedure (Grant \textit{et al.}, 2000). Several studies are currently investigating the sensitivity of various methods for use in detection of \textit{M. a. paratuberculosis} from infected individual and bulk milk samples (Giese & Ahrens, 2000; Dundee \textit{et al.}, 2001; Odumeru \textit{et al.}, 2001).

1.7.2 Detection of DNA

To date, only three subspecies-specific DNA fragments have been identified in the \textit{M. a. paratuberculosis} genome. The most widely used is the IS900 element. However, the F57 DNA fragment (Poupart \textit{et al.}, 1993), and the \textit{hspX} gene (Ellingson \textit{et al.}, 1998), have also been identified. F57 is a 620 base pair DNA fragment with 56\% identity to a region of DNA in \textit{M. a. avium}. Other than this, it shares no similarity to any known sequences, and appears to be present as a single copy in the genome (Poupart \textit{et al.}, 1993). It has not been widely used as a diagnostic marker (Poupart \textit{et al.}, 1993; Cocito \textit{et al.}, 1994). The \textit{hspX} gene of \textit{M. a. paratuberculosis} shares 60\% identity to similar genes in other mycobacteria. Its sequence suggests it encodes a protein related to the family of DnaJ heat-shock proteins (Ellingson \textit{et al.}, 1998). A 30 base pairs probe designated MP derived from this gene may be useful as a diagnostic probe.

1.7.2.1 IS900 insertion element

IS900, the first insertion element identified in mycobacteria, is unique to strains of \textit{M. a. paratuberculosis} and remains the main genetic tool for identification of \textit{M. a. paratuberculosis} (Collins \textit{et al.}, 1989; Green \textit{et al.}, 1989; Moss \textit{et al.}, 1991; Collins \textit{et al.}, 1997). It will be described in some detail here. IS900 was discovered independently by three
groups (McFadden & Hermon-Taylor, 1988; Collins et al., 1989; Murray, 1989). The first
publication of the sequence in a peer-reviewed journal was by the Hermon-Taylor group
(Green et al., 1989), who were the first to call it IS900. IS900 is 1.45 kilobase pairs in length
and lacks the terminal inverted and direct repeats which are characteristic of most insertion
sequences. IS900 is repeated 15-20 times in the genome and shows a degree of target
sequence specificity. The consensus insertion site for IS900 is 5' CATGN(4-6)*CNCCTT
3', where * denotes the site of insertion. It always inserts in the same orientation with respect
to this target site (Green et al., 1989). Analysis of the junction sites for several copies of
IS900 in the M. a. paratuberculosis genome reveal that IS900 inserts between the putative
ribosome binding site and start codon of a given gene (Doran et al., 1997; Bull et al., 1999).
IS900 contains two open reading frames. A single open reading frame (ORF 1), on the
positive strand encodes the 43 kDa protein (p43). Expression of this protein has been
demonstrated in M. a. paratuberculosis cultures and appears as 28 kDa when fully
processed. Although not fully characterised, it is postulated to be a transposase that may
direct the insertion of the element into the genome (Tizard et al., 1992). Observations
supporting this function include the stable integration of artificial transposons containing the
p43 gene into the genomes of various host strains of mycobacteria (Dellagostin et al., 1993).
ORF 2 on the negative strand encodes the hed (host expression-dependent) gene. The
protein product, Hed, is approximately 55 kDa and has been expressed in recombinant E.
coli, but not further characterised (Doran et al., 1997). The hed gene exists in IS900 devoid
of a promoter, ribosome binding site and termination codon. A promoter sequence PAN is
capable of driving expression of Hed when IS900 inserts between a RBS adjacent to PAN
and a start codon of a host-encoded gene (Murray et al., 1992). This provides the scenario of
an IS element actively 'hijacking' promoter and RBS signals for its own expression. Genes
similar to hed exist in the insertion elements IS901, IS902 and IS116. The function of Hed
is unknown but it has been postulated to be a repressor of transposition, preventing IS900
from transposing within its host (Hernandez Perez et al., 1994). Others have suggested it is
involved in alternate iron transport (Doran et al., 1994; Naser et al., 1998).

Other closely related insertion elements found in various MAC strains are IS901 (Kunze et
al., 1991), IS902 (Moss et al., 1992), IS1245 (Guerrero et al., 1995), IS1311 (Roiz et al.,
1995) and IS1110 (Dale, 1995). IS1311 is ubiquitous in MAC strains including M. a.
paratuberculosis while IS1245 is absent in M. a. paratuberculosis (Collins et al., 1997).
IS1110, which is rarely found in MAC strains (Collins et al., 1997), is related to IS900 but
differs in being highly mobile (Cousins et al., 1999). IS900 has been implicated in the
pathogenicity of M. a. paratuberculosis based on its similarity to IS901 which is found in
virulent strains of *M. a. avium* and *M. a. silvaticum* (Doran et al., 1994). Experiments with recombinant *M. smegmatis* carrying IS900 suggest the presence of IS900 may be involved in mycobactin dependency and also the slow growth rate of *M. a. paratuberculosis* (Naser et al., 1998).

PCR detection of IS900 is often used instead of mycobactin dependency for confirmation of *M. a. paratuberculosis* because of its rapidity. IS900 detection by PCR is available as a commercial kit (Idexx, USA). This test is reported to have a sensitivity of approximately 60% in infected cattle, based on cattle diagnosed by faecal culture (Whipple et al., 1992). A report of finding organisms that give positive results with IS900 PCR but are not *M. a. paratuberculosis* has placed some doubt on the routine use of this test as the ultimate confirmation of *M. a. paratuberculosis* (Cousins et al., 1999). The isolates identified demonstrated considerable similarity with IS900, particularly at the 5’ end which was previously considered to be specific for *M. a. paratuberculosis* (Vary et al., 1990). This also questions results obtained using PCR as the sole means of detection of *M. a. paratuberculosis* in Crohn’s disease tissue and suggests that further restriction endonuclease analysis on PCR products be applied to confirm IS900. As with faecal culture, the sensitivity of PCR detection of IS900 is related to the number of organisms present. So far, most studies have shown lower sensitivities with PCR on faeces, milk and other body samples such as blood. Interfering substances, commonly found in faeces also may lower sensitivities.

### 1.7.3 Immunological tests

Many different methods have been used to measure humoral and cellular responses to *M. a. paratuberculosis* infection. Sensitivity and specificity estimates of the tests vary widely, depending upon the reagents used, the competence of the personnel performing the test and the prevalence of disease in the herd. There are a variety of antigens used in the tests, making comparisons between laboratories difficult. A practical problem of any of the immunological assays used is cross-reactions with other closely related mycobacterial species. A recent study found cattle experimentally infected with *M. bovis* showed high levels of cross-reacting antibody using a commercial ELISA for the detection of *M. a. paratuberculosis* infection (Olsen et al., 2001). *M. a. avium* is the most closely related organism to *M. a. paratuberculosis*, thus any test used to diagnose Johne’s disease may also detect animals that have been sensitised to *M. a. avium*. Repeated testing may help establish the true Johne’s status of animals, as sensitivity to *M. a. avium* is considered to be transient (Billman-Jacobe et al., 1992). Comparative studies show that serological assays mainly identify animals with
Figure 1.2. General relationship of tests and immune responses to infection with *M. a. paratuberculosis* over time. Adapted from Buergelt (1999).
more advanced Johne's disease, whereas assays that detect cellular reactivity can identify animals in earlier stages (Buergelt et al., 1977; de Lisle & Duncan, 1981). There is evidence that tests for cellular responses lose their diagnostic potential with animals in the late phase of clinical disease (Riemann & Abbas, 1983), but others have reported sensitivities of 100% in clinically affected cattle (Billman-Jacobe et al., 1992). Figure 1.2 shows the general relationship between the immune responses of infected animals and the ability of various diagnostic tests to detect them.

1.7.3 Tests for cellular responses

Tests used for cellular responses include the skin test (intradermal test) and the IFN-γ EIA test. Other tests such as lymphocyte blastogenesis or lymphocyte migration inhibition test are not commonly used for diagnosis. Cellular tests use purified protein derivative (PPD) antigens. Johnin PPD is a mixture of trichloroacetic acid-precipitated and solubilized proteins from heat-treated *M. a. paratuberculosis* cultures (Office International des Epizooties, 2000). The product therefore contains a complex mixture of cytosolic, envelope and culture supernatant proteins. These tests therefore have relatively poor specificity but good overall sensitivity (Buergelt et al., 1977; de Lisle & Duncan, 1981). There is a noted tendency of the cellular responses to Johnin PPD antigen to decline with the onset of more advanced disease and with heavy faecal shedding (Whittington & Sergeant, 2001). The skin test can result in up to 50% of uninfected animals responding in an infected herd (Riemann & Abbas, 1983; Chiodini et al., 1984a). The IFN-γ EIA is considered to be the most suitable test for cellular immune responses, with a reported sensitivity of 50 to 75% as compared to faecal culture and when used as a single test in herds with low-infection status (Stabel & Whitlock, 2001). In a similar study, the IFN-γ assay detected the majority of the infected cattle on the first test and 6% more on the second test, indicating the value of consecutive testing (Billman-Jacobe et al., 1992). Studies with experimentally infected sheep indicate the IFN-γ test was 100% sensitive in detecting infection in sheep at all stages of disease, but only when they were tested consecutively (Gwozdz et al., 2000b). In these studies, both IFN-γ and antibody responses were frequently intermittent, and it was noted IFN-γ responses varied more in animals with numerous bacteria present in tissues.

1.7.3.2 Tests for humoral responses

Tests that measure circulating antibody include complement fixation, agar gel immunodiffusion (AGID), and enzyme-linked immunosorbent assay (ELISA). The
complement fixation test has been reported by numerous groups to be of poor sensitivity and specificity (Wilks et al., 1981; Colgrove et al., 1989; Reichel et al., 1999; Gwozdz et al., 2000b), but despite this, it is a frequently used test for import/export purposes worldwide (Stabel, 1998). AGID is considered to be 100% specific, but suffers from sensitivity as low as 26% (Sockett et al., 1992; Whitlock et al., 1999). The absorbed ELISA (Commonwealth Serum Laboratories, Victoria, Australia) generally has greater sensitivity and specificity than other serological tests (Riemann & Abbas, 1983; Sockett et al., 1992). Absorption of serum with environmental mycobacterial preparations, such as M. phlei, has greatly increased the specificity of this test (Yokomizo et al., 1985; Milner et al., 1987; Ridge et al., 1991; Hilbink et al., 1994). As with other serological tests, it performs best in animals with clinical disease (80% detection) and poorly with subclinically infected animals (as low as 15% detection), (Billman-Jacobe et al., 1992; Sweeney et al., 1995). In cattle, it has reported overall sensitivities of 35% (Whitlock et al., 1999), 57% (Milner et al., 1990) and 45% (Collins & Sackett, 1993), and a specificity of 99% (Milner et al., 1990; Collins et al., 1991). Similar variations are reported for sheep (Hope et al., 2000; Whittington & Sergeant, 2001). Estimated sensitivities depend to a large extent on the proportion of animals in the later stages of disease present in a herd. A comparison of the AGID and ELISA tests in sheep flocks has indicated these two tests may be detecting different populations of infected animals (Hope et al., 2000). The ELISA was more likely to detect subclinically infected animals than the AGID and the AGID test was better at detecting animals with low body condition. Although the overall sensitivities of these two tests were approximately the same (~36 to 55%), using both tests increased the sensitivity to 50 to 58%, and illustrates the usefulness of using tests in parallel.

1.7.4 Histopathological detection

Histopathological detection of Johne’s disease is largely confined to animals submitted for necropsy. Terminal ileum and ileocaecal lymph nodes are the appropriate tissues for examination (Buergelt et al., 1978; Pemberton, 1979; Benedictus & Haagsma, 1986; Kreeger et al., 1987; Gwozdz et al., 2000a). Morphological demonstration of acid-fast organisms in lesions in appropriate tissues is typical of Johne’s disease, but such organisms may not always be found in subclinical cases (Merkal, 1973; Whittington et al., 1999; Hope et al., 2000). Surgical biopsy of ileocaecal lymph node from particularly valuable animals is considered to be an aid in confirmatory diagnosis in live animals (Julian, 1975; Pemberton, 1979).
Microscopic examination of tissue lesions in animals with Johne's disease show that *M. a. paratuberculosis* organisms exist in clumps (Whitlock, 1991). Necropsy of naturally infected animals shows that typical gross lesions of chronic enteritis are similar in cattle and sheep, and are found predominantly in the distal ileum, but occasionally extend into the colon and jejunum. *M. a. paratuberculosis* can also be found in concentrated samples of circulating blood monocytes from infected cattle and sheep (Koenig et al., 1993; Gwozdz et al., 1997). This may be the mechanism for dissemination of the organism from infected sites to other organs such as liver and reproductive tissues.

1.8 The mycobacterial envelope and its relationship to pathogenicity and immunology

The envelope of mycobacteria, which consists of the plasma membrane, cell wall and in some cases a capsule, forms the interface between the mycobacterium and its host. Components of the envelope have been studied in order to explain the complicated pathology of mycobacterial diseases, since only components of the envelope, plus a limited number of secreted proteins are thought to be readily accessible to the host. The wall and capsule together provide passive resistance to host agents that would otherwise destroy the bacteria. Passive resistance may be necessary for pathogenicity, but is clearly not sufficient, and it is likely that exported components such as proteins play an important role in directing the pathogenic events that lead to disease. Components of the envelope, such as lipoarabinomannan (LAM), may play a role in ensuring that mycobacteria are phagocytosed by macrophages (Chan & Kaufmann, 1994; Tessema et al., 2001). Other antigens possibly involved in this process include the Ag85 complex, which have fibronectin binding capabilities (Abou-Zeid et al., 1988a). Many surface-linked proteins of Gram-positive pathogens are thought to be important for pathogenicity. For example, the streptococcal surface M-proteins are involved in binding components from the host that inhibit the deposition of complement on the bacterial surface and prevent phagocytosis (reviewed by Navarre & Schneewind, 1999).

Once inside the macrophage, pathogenic species such as *M. a. paratuberculosis*, *M. a. avium* and *M. tuberculosis* prevent or fail to stimulate phagosome-lysosome fusion. This effect only occurs with live mycobacteria, not killed cells (Frehel et al., 1986a; Kuehnel et al., 2001), and therefore appears to depend on some exported product(s) rather than resident components of the cell. Interference with subsequent host defense mechanisms likely
involves alterations of cytokine production. LAM is known to have such effects (Tessema et al., 2001), and other components of the envelope need to be investigated for similar abilities.

The mycobacterial cell wall is a relatively simple structure that lies outside the plasma membrane. It consists of three basic layers; cross-linked peptidoglycan linked to arabinogalactan which in turn is attached to mycolic acids. Mycolic acids are complex, long chain fatty acids and are a major component of the mycobacterial cell wall. They form an impermeable, continuous, waxy monolayer surrounding the cell. This layer is responsible for the characteristic acid-fast staining reaction of mycobacteria (Draper, 1982) and is also a component of Freund's complete adjuvant (wax D) (Stewart-Tull, 1983). Different groups of mycobacteria possess chemically different types of mycolic acids (Minnikin et al., 1975; Goodfellow & Minnikin, 1977), however there is no correlation between possession of a particular mycolic acid and pathogenicity (Daffe et al., 1983; Minnikin et al., 1984). The actual amount and arrangement of mycolic acids may be factors important for passive survival of the mycobacterium inside the phagocytic cell by protecting the cell wall and membrane infrastructure from the effects of lytic enzymes and preventing even quite small molecules from penetrating into the cell wall (Daffe & Draper, 1998).

1.8.1 Lipoarabinomannan

LAM is a complex glycolipid consisting of repeating units of arabinose and mannose linked to a phosphatidylinositol moiety (Hunter et al., 1986). It is a major component of the cell envelope and has merited special attention because of the remarkable range of activities it appears to possess. A major activity is immunomodulation of the host immune response in mycobacterial diseases. LAM stimulates macrophages to produce tumour necrosis factor alpha (TNF-α) and interleukin 1 (IL-1), IL-6 and several other interleukins (Barnes et al., 1992; Chatterjee et al., 1992; Bradbury & Moreno, 1993; Adams & Czuprynski, 1994; 1995). LAM also induces several ‘early’ genes involved in activation of macrophages that stimulate the production of nitric oxide (Roach et al., 1993; 1994; Anthony et al., 1994; Schuller-Levis et al., 1994). It has also been shown to have the ability to scavenge toxic free radicals, and block the transcription of IFN-γ-inducible genes in human macrophage-like cell lines (Chan et al., 1991). LAM has also been implicated in phagocytosis of mycobacteria by macrophages (Schlesinger et al., 1994; Venisse et al., 1995). LAM of M. a. paratuberculosis shows structural and immunological differences between bovine and ovine isolates (Sugden et al., 1987). The exact location of LAM in the mycobacterial envelope is unknown, but it is postulated to span the wall, possibly anchored to the plasma membrane.
However, there is no direct evidence for this. LAM of mycobacteria is commonly found in culture filtrates, suggesting that it may also be released into the environment of the host during its growth.

As early as 1961, descriptions of electron-transparent zones surrounding various species of pathogenic mycobacteria have been described *in vivo* (Merckx *et al.*, 1964; Draper & Rees, 1970; Armstrong & D'Arcy Hart, 1971). Capsules are thought to protect the organism against the cidal effects found in macrophages by presenting a physical barrier to lysosomal material (Brown & Draper, 1970; Armstrong & D'Arcy Hart, 1971). All species of mycobacteria so far examined have a superficial external layer, but its thickness varies considerably between species (Rastogi *et al.*, 1986). The extensive capsules of *M. tuberculosis* (Ortalo-Magne *et al.*, 1995), *M. lepraemurium* (Draper & Rees, 1973), and *M. leprae* (Hunter & Brennan, 1983; Boddinigius & Dijkman, 1989), are relatively rich in proteins as compared to the cell wall. In these cases, the capsules were found to contain only small amounts of glycolipids, some of which were species-specific. In the case of *M. tuberculosis*, the capsule was found to consist mainly (97%) of proteins and polysaccharides non-covalently linked to the cell wall (Ortalo-Magne *et al.*, 1995). This is in contrast to the previously perceived idea that mycobacteria expose a thick, waxy coat to their environment.

*M. a. paratuberculosis*, *M. a. avium* and *M. intracellulae* are also surrounded by electron transparent zones within infected cells (Rastogi *et al.*, 1989; Kuehnel *et al.*, 2001), but the nature of this region has not been characterised. The formation of the zone around *M. a. avium* has been shown to be dependent on viability (Frehel *et al.*, 1986b).

The precise location of resident capsular proteins of mycobacteria is difficult to determine because, like the rest of the capsular components, they are present in culture filtrates as a result of shedding from the surface of the cells (Ortalo-Magne *et al.*, 1995). This makes it difficult to conclude whether these proteins are present in the medium due to sloughing of capsular material or are truly secreted to the external environment. Possible resident proteins in capsules of *M. tuberculosis* are the fibronectin-binding Ag85 complex (Wiker & Harboe, 1992). Some proteins such as the 25 kDa protein of *M. tuberculosis* (Andersen *et al.*, 1991a; Ortalo-Magne *et al.*, 1995) and the 45-47 kDa antigen complex of *M. bovis* BCG and *M. tuberculosis* (Romain *et al.*, 1993; Laqueyrerie *et al.*, 1995) appear to be secreted proteins, based on their presence in culture filtrates but can be found in capsular extracts, presumably *en route* to the external environment.
1.9 Export of proteins in mycobacteria

Proteins of mycobacteria that appear in culture filtrates have been the focus of many studies investigating host immunity and development of potential vaccines and diagnostic reagents. The use of the term 'secreted proteins' to define proteins which appear in short-term culture supernatants has come under scrutiny as more recent studies indicate the vast majority of proteins found in culture supernatants of actively growing mycobacteria consist largely of proteins associated with the cell wall and membrane. It appears these are gradually released during growth as such culture supernatants are devoid of markers of autolysis such as isocitrate dehydrogenase (Andersen et al., 1991a; Wiker et al., 1991; Orme et al., 1993; Florio et al., 1997).

Different workers have identified various numbers of protein bands on SDS-PAGE gels of short-term mycobacterial culture filtrates. For *M. a. paratuberculosis*, 17-21 bands are evident (Valentin-Weigand & Moriarty, 1992; White et al., 1994). For *M. tuberculosis*, up to 100 bands have been resolved (Horwitz et al., 1995), however only approximately 11 of these predominate (Horwitz et al., 1995; Harth & Horwitz, 1999). Between 100 and 200 proteins from *M. tuberculosis* culture filtrates have been resolved using 2D electrophoresis (Elhav & Andersen, 1997; Sonnenberg & Belisle, 1997). These differences are probably a reflection of variations in culture conditions used, including the definition 'short-term', which is not consistently defined between different laboratories.

There are many descriptions of the label 'exported protein', although it has been proposed that this term should be reserved for proteins for which a signal peptide mediates their active transport across the plasma membrane (Young et al., 1990). Others have suggested that this classification should be further divided into the two categories of soluble (secreted, extracellular, non-lipidated proteins), and insoluble (cell-associated lipoproteins) (Wiker et al., 1998). For the purposes of this report the term 'exported protein' will encompass those proteins which are translocated across the cell membrane, and therefore will include both extracellular and cell envelope-associated proteins.

It has recently been proposed that culture filtrate proteins that lack signal sequences are present as a result of bacterial leakage, rather than being actively secreted by an unknown transport mechanism. Examples include Fe/Mn superoxide dismutase (SodA) (Harth & Horwitz, 1999) and glutamine synthase (Harth & Horwitz, 1997) of *M. tuberculosis*. Experimental evidence has been produced to support the hypothesis that the presence of
proteins in culture filtrates is a consequence of leakage of highly expressed and stable proteins from the cytoplasmic compartment (Tullius et al., 2001). This work is particularly enlightening in view of the past suggestion from the same laboratory that export of glutamine synthase and SodA was dependent on information contained within the protein sequence/structure. This was based on the observation of export of these *M. tuberculosis* enzymes in the heterologous host *M. smegmatis* whereas the endogenous *M. smegmatis* enzymes remained intracellular (Harth & Horwitz, 1997; 1999). The authors later proposed the *M. tuberculosis* recombinant enzymes were being expressed in greater amounts than the *M. smegmatis* enzymes (Tullius et al., 2001). Highly expressed ESAT-6 and heat-shock proteins of *M. tuberculosis* are similarly proposed to enter culture supernatants by escape from the cytoplasm (Wiker, 2001).

The assignment of a Localisation Index (LI) value to proteins has been proposed to give an indication of their subcellular localisation. LI is based on semi-quantitative comparisons of the relative amount of a protein in the culture fluid divided by its relative amount in a sonicate preparation of the corresponding washed bacilli (Wiker et al., 1991). Those proteins that are only detectable in the culture medium have high LI values (>10) and represent secreted (soluble exported) proteins. Those that are present predominantly in sonicates have very low LI values (<1), indicating that they are cytosolic components. The proteins detected in both culture filtrates and sonicates have intermediate values (>1) and represent exported lipoproteins, which are cell surface-associated. Although only a small number (~14) of proteins have been analysed by this method, there appears to be full concordance between LI values greater than 1 and presence of a signal sequence (Wiker et al., 1999). For example, Ag85A from *M. tuberculosis* possesses a signal sequence and has a LI value of 150, whereas the cytosolic protein GroES does not have a signal sequence and has a LI value of 0.04.

Although lipoproteins are predicted to be insoluble due to the attached lipid moiety, it appears that some may be soluble. Examples include the 38 kDa, 26 kDa and 27 kDa lipoproteins of *M. tuberculosis* which appear in large amounts in early culture filtrates. Detergent extraction of culture filtrates showed these proteins remained in the aqueous phase (Young & Garbe, 1991). Extractions of the insoluble sonicated pellet showed they fractionated to the detergent phase. Metabolic labeling using isotope confirmed the cell-associated proteins were lipidated, however, it was not investigated if these proteins in the culture filtrate were similarly lipidated (Young & Garbe, 1991). It may be that the proteins released to the culture supernatant, lack lipidation and therefore remain soluble. Similar comparative analyses of protein antigens from *M. bovis* BCG have concluded there is no clear correlation between the amounts of
protein found in the cell wall and culture filtrates, nor between their sensitivity to detergent extraction and release into culture supernatants (Florio et al., 1997). It appears that most of these methods are at best able to give an indication of export of a protein, but lack the ability to further localise proteins within the envelope. The presence of a signal sequence remains the key feature that distinguishes exported proteins from cytoplasmic ones (Simonen & Palva, 1993).

1.9.1 Export pathways and signal peptides

Most proteins to be exported beyond the cytoplasmic membrane of the bacterial cell, reach their destination via the general secretory pathway (Pugsley, 1993; Economou, 1999). In Gram-positive bacteria, exported proteins are either released to the environment or are held by a variety of mechanisms to the cell membrane or cell wall. In Gram-negative bacteria, exported proteins are held in the periplasmic space or are integrated into or transported across the outer membrane. The general secretory pathway involves a number of proteins, several of which have been named with the prefix Sec. For this reason, this pathway is also called the Sec-dependent pathway. This pathway is common to all bacterial and eukaryotic cells studied to date (Navarre & Schneewind, 1999). Although mycobacterial export systems have not been characterised, there is strong evidence that they possess a Sec-dependent export pathway. Database searches have revealed that homologues of the genes for the Sec-dependent pathway exist in mycobacteria (Chubb et al., 1998; Limia et al., 2001). Also, several analyses of mycobacterial exported proteins have shown that most possess typical N-terminal signal sequences (Young et al., 1992; Hewinson & Russell, 1993; Chubb et al., 1998), which are required for export via the general secretory pathway. Other export systems, such as type III secretion systems, which do not require signal sequences, have been described for the export of virulence factors in several Gram-negative species (reviewed by Cheng & Schneewind, 2000). Such systems have so far not been identified in mycobacteria through functional or genetic searches (Tullius et al., 2001).

In the general secretory pathway, protein export is dependent on the presence of a signal sequence that is recognised by a ribonucleoprotein complex called the signal recognition particle (SRP). SRP binds the h-region of the signal peptide (described below) by virtue of a hydrophobic groove and delivers it to the translocase machinery in the membrane (von Heijne, 1998). The translocase has a membrane-spanning domain made up of heterotrimers of membrane proteins. Together these create a channel through which the mature portion of the protein is transferred, leaving the signal sequence embedded in the membrane (Economou,
Signal peptidase

Type 1: $S_n^3 X S_n^1$
Type II: $L_n (A/S)(G/A) C^1$

Figure 1.3. Common features of a signal peptide. Adapted from Fekkes & Driessen (1999).

$S_n$ = small neutral side chain (e.g., A, S, G, T)
$L_n$ = large hydrophobic side chain (L is preferred)
$X$ = any amino acid residue
1999). Once the mature portion has been delivered through the membrane, it is cleaved by a signal peptidase, which scans the surface of the membrane looking for characteristic cleavage sites within the amino acid sequence of the signal sequence. The mature protein is then released (von Heijne, 1998). The information for the final cellular destination for all exported proteins resides in the N-terminal signal sequence of the protein (Pugsley, 1993; von Heijne, 1998; Economou, 1999). Exported polypeptides have been extensively studied in \textit{E. coli} and to a lesser extent in Gram-positive bacteria such as \textit{Staphylococcus aureus}, \textit{Streptococcus pyogenes}, and \textit{Bacillus subtilis}.

Signal peptides consist of 20-40 amino acids at the amino terminus, which are predominantly hydrophobic (Emr \textit{et al.}, 1980; Silhavy \textit{et al.}, 1983). As shown in Figure 1.3, signal peptides analysed from various prokaryotic species have identified common features; a short (1-5 amino acids) amino terminal n-region containing at least one positively charged residue such as arginine or lysine, followed by a core h-region which is a hydrophobic stretch of 7-20 amino acids. This is followed by a carboxyl c-region that is more polar and contains the cleavage site for the signal peptidase which removes the signal sequence from the mature polypeptide (von Heijne, 1985; 1998). Examples of known exported mycobacterial proteins containing signal peptides are MPB70 of \textit{M. bovis} BCG (Harboe & Nagai, 1984; Terasaka \textit{et al.}, 1989), MPT64 (Yamaguchi \textit{et al.}, 1989), and MPT53 (Wiker \textit{et al.}, 1999), and the 38 kDa antigen (Andersen & Hansen, 1989) of \textit{M. tuberculosis}. Hydrophobicity of the h-region and the positions of the positively charged residues are considered features essential to the function of signal sequences (von Heijne, 1981; 1982; Inouye \textit{et al.}, 1982; Emr & Silhavy, 1983; Finkelstein \textit{et al.}, 1983; Vlasuk \textit{et al.}, 1983). The net positive charge of the n-region is thought to interact with the negatively charged surface of the inner membrane (Inouye \textit{et al.}, 1982), and the core h-region, which has a high propensity for $\alpha$-helical conformation, is thought to facilitate interaction with the interior of the membrane bilayer (Briggs \textit{et al.}, 1986). Mutation studies in \textit{E. coli} demonstrate the n-region can be neutral or even negatively charged, but such peptides are processed at reduced rates (Gennity \textit{et al.}, 1990). This can be compensated for by an increase in hydrophobicity of the h-region (Phoenix \textit{et al.}, 1993). Signal peptides of Gram-positive organisms differ from those of Gram-negative by generally being longer, more hydrophobic, and more charged in their amino terminus (von Heijne & Abrahmsen, 1989; von Heijne, 1990b). Signal peptides of Gram-positive bacteria usually function in Gram-negative, but the reverse is not always true (Schneewind \textit{et al.}, 1992). The reason for this is not known but may be related to the specificity of the transport machinery.
Although some exported proteins of mycobacteria are clearly exported to the extracellular environment, many remain associated with the plasma membrane and/or the cell wall. Examples include the wall-associated 38 kDa (Florio et al., 1997) and 19 kDa (Ashbridge et al., 1989; Florio et al., 1997) lipoproteins of *M. tuberculosis*. Anchoring of polypeptides containing signal sequences to the cell envelope may occur by at least three means. Lipidation at the N-terminus of the mature protein, possession of a hydrophobic C-terminus containing the LPXTG motif, or possession of a signal sequence lacking a signal peptidase cleavage site. The first two scenarios require a signal sequence for translocation across the plasma membrane and subsequent signal peptide cleavage. In the third case, the signal peptide provides translocation but since cleavage cannot occur, the protein remains anchored in the plasma membrane (von Heijne, 1990a).

1.9.1.1 Cleavage of signal peptides

Signal peptidases are integral membrane proteins that remove signal peptides from the precursors of exported proteins. Type I signal peptidases serve to cleave ordinary polypeptides and type II signal peptidases cleave lipid-modified polypeptides (Sankaran & Wu, 1994). In both cases, the cleavage site of the signal sequence from the mature polypeptide has a common format in both eukaryotic and prokaryotic cells (Perlman & Halvorson, 1983; von Heijne, 1984). The most striking feature is the presence of small, neutral amino acids such as alanine, glycine, serine, and threonine (von Heijne, 1984; Watson, 1984). For type I signal peptidase cleavage sites, positions at -1 and -3, relative to the mature polypeptide are considered to be the most critical for cleavage (von Heijne, 1984). For type II signal peptidase cleavage sites, residues at positions -3 and +1 are the most important (Sankaran & Wu, 1994). For these, a larger, hydrophobic residue such as leucine occurs in the -3 position, and the +1 position of the mature polypeptide is always filled with cysteine to which the lipoyl moiety becomes attached prior to signal peptide cleavage (Watson, 1984).

1.9.1.2 Lipidation

A lipoprotein consensus element (lipobox), present at the signal peptidase cleavage site, signals the addition of fatty acid moieties to what will become the amino terminal cysteine in the mature polypeptide. This lipobox consists of the amino acid consensus "\( ^3L(A/S)(G/A)\) C\(^+1\)", where leucine is given preference at the -3 position, and cysteine must be present at the +1 position (Hayashi & Wu, 1990; Braun & Wu, 1994). In mycobacterial lipoproteins so far analysed, the region following the consensus lipidation sequence
characteristically consists mainly of serine and threonine residues (Nair et al., 1993). It is proposed that lipoproteins are loosely associated with the cell wall or membrane, being anchored by their hydrophobic lipid tail (Young et al., 1990; Pugsley, 1993). Lipoproteins can, at least in some cases, be released from the cell surface. This has been demonstrated in Streptococcus pneumoniae (Pearce et al., 1994), and B. subtilis (Perego et al., 1991). In M. tuberculosis, this appears to occur for several antigens, including the 19 kDa and 38 kDa lipoproteins which appear to be ubiquitous in the cell envelope layers as well as in culture filtrates (Florio et al., 1997).

Several proteins that are exported to the periplasmic space in E. coli are known to be lipidated in Gram-positive bacteria (Gilson et al., 1988). For example, the β-lactamase of Gram-negative bacteria is a soluble polypeptide located in the periplasmic space, whereas the β-lactamases of Gram-positive bacteria such as Bacillus licheniformis and Staphylococcus aureus are lipoproteins that remain tethered to the cell membrane (Navarre et al., 1996). Mutation studies of β-lactamase precursors of Gram-positive bacteria show they can be substrates for both type I and type II signal peptidases, resulting in a mixture of exported, soluble β-lactamase and membrane-anchored, lipid modified β-lactamase (Navarre et al., 1996). Since Gram-positive bacteria have no outer membrane to retain proteins, the lipooyl moiety is thought to be a targeting device to anchor them to the cell (Gilson et al., 1988; Sutcliffe et al., 1993; Sutcliffe & Russell, 1995). Since many of the periplasmic proteins of Gram-negative bacteria have been shown to have nutrient transport roles, it has been proposed that exported lipoproteins of mycobacteria may also have nutrient transport functions (Young et al., 1990).

1.9.1.3 C-terminal anchoring

Gene sequences of several surface proteins of Gram-positive bacteria has revealed a C-terminal motif consisting of LPXTG where X is any amino acid followed by a C-terminal hydrophobic domain anchorage (Fischetti et al., 1990). In this model, the precursor protein is exported from the cytoplasm via a signal sequence, but the protein is prevented from release by a hydrophobic domain. The LPXTG is thought to be a cleavage site for a peptidase that cleaves at this point and covalently links the peptide to the peptidoglycan of the cell wall (Navarre & Schneewind, 1999).
1.10 Searching for exported proteins of mycobacteria

Proteins exported by mycobacteria to the environment have been the focus of research for vaccine development and improved diagnostics (reviewed by Cooper & Flynn, 1995). A common experimental approach used to identify these proteins is analysis of the proteins present in supernatants of cultures grown in vitro. This has been done using electrophoretic or chromatographic separation followed by immunological screening of the isolated components (Andersen et al., 1991a; Raynaud et al., 1998; Gobin et al., 1999; Rowland et al., 1999; Wiker et al., 1999). For *M. a. paratuberculosis*, this has lead to the identification and isolation of ferric reductase (Homuth et al., 1998), SodA (Liu et al., 2001) and the 14 kDa protein (Olsen et al., 2000b), (see Table 1.2). These methods target those proteins exported to the extracellular environment, however those proteins that are predominantly surface-exposed may not be identified.

A more encompassing approach is to use genetic information to search existing databases for proteins similar to known exported proteins or to look for those genes encoding signal peptides. In a study by Harboe and co-workers (1998), 10 *M. leprae* proteins similar to known exported proteins from *M. tuberculosis* and *M. bovis* were identified by DNA sequence similarities. This same study also employed two different sequences based on the compiled signal sequences of known exported proteins of *M. tuberculosis* to search protein databases to identify other putative exported proteins of *M. tuberculosis*. Approximately thirty genes for putative exported proteins were identified. These proteins have not been further analysed. The difficulty in employing searches for signal sequences based on sequence identities is the lack of uniformity in the actual amino acid sequence as they are generally only similar in terms of their structural and chemical characteristics (Simonen & Palva, 1993; Izard & Dendall, 1994).

More recently, Gomez and co-workers (2000) searched the amino acid sequences of 3924 proteins, deduced from the DNA sequence of the *M. tuberculosis* genome, for the presence of signal sequences. Computer programmes were used to score potential signal peptides, and eliminate those with features such as internal transmembrane segments and lipid attachment sites. Fifty-two, or 1.3% of the proteins analysed were predicted to be secreted to the external environment.

Exported proteins can be identified using molecular approaches involving the fusion of export-competent peptides to reporter proteins such as β-lactamase and alkaline phosphatase.
These systems rely on the concomitant export and functional expression of the fused reporter protein for identification of export. β-lactamase fusions rely on the exported fusion protein to protect the cells from high levels of ampicillin. Expression of plasmid constructs in *E. coli* have been used to identify the Ag85A, Ag85C and a penicillin-binding protein from *M. tuberculosis* (Chubb et al., 1998). PhoA fusions employ recombinant *E. coli* PhoA as a marker for export. The activity of PhoA generally requires it to be exported across the plasma membrane where it can form intramolecular disulfide bonds and undergo subsequent dimerisation to form the active enzyme (Hoffman & Wright, 1985; Boquet et al., 1987; Manoil et al., 1990; Derman & Beckwith, 1995). The enzyme can be detected by the addition of a colourimetric substrate to the medium. In the study of Gomez and co-workers (2000), 10 *phoA* gene fusions were constructed in plasmids with *M. tuberculosis* genes predicted to encode signal sequences. These constructs lacked the original sequence coding for the signal peptide of the PhoA protein and instead used the predicted signal peptide sequence encoded by the cloned DNA. Expression of the fusion proteins in *E. coli* was driven by an upstream *lac* promoter. All 10 predicted exported proteins were PhoA+, indicating that this method was highly accurate in identifying exported proteins. Other studies have also shown that the appearance of PhoA phenotype is generally associated with detection of truly exported proteins, and is not a result of lytic release of PhoA fusions from the cytoplasm (Timm et al., 1994).

PhoA fusions have been used to identify proteins that are exported across the periplasmic space of *E. coli* (Michaelis et al., 1983; 1986), and for characterisation of membrane protein topology (Manoil et al., 1988). PhoA technology has been used for the identification of exported proteins of *Streptococcus pneumoniae* (Pearce et al., 1993), *E. coli* (Boquet et al., 1987), *Helicobacter pylori* (Johnson et al., 1995; Bina et al., 1997; Oliaro et al., 2000), *Vibrio cholerae* (Taylor et al., 1989), and *Bordetella pertussis* (Knapp & Mekalanos, 1988), *Staphylococcus aureus* (Williams et al., 2000) and *Actinobacillus actinomycetemcomitans* (Ward et al., 2001). In mycobacteria, PhoA fusions have identified soluble exported proteins and various lipoproteins such as ERP, DES, the 19 kDa protein, Pel and SodC (Berthet et al., 1995; Lim et al., 1995; Mdluli et al., 1995; Borich, 1997; Moran et al., 1999; Gomez et al., 2000; Dupont & Murray, 2001). It is possible that reporter fusion systems will not detect some protein products that may require special conditions provided only by the native species or require *in vivo* growth. An example is Mig of *M. a. avium* and *M. tuberculosis*. Mig is a cell wall-associated protein, first identified only in infected host-cells and not expressed *in vitro* in normal *M. tuberculosis* culture (Plum & Clark-Curtiss, 1994).
1.11 Components isolated from *M. a. paratuberculosis*

At least 44 antigenic components are recognised by rabbit immune sera produced against whole sonicated preparations of *M. a. paratuberculosis* (Gunnarsson & Fodstad, 1979b). Many of these are common to other MAC subspecies and *M. bovis*. Table 1.2 summarises components of *M. a. paratuberculosis* that have been isolated to date. Few of these have been well characterised.

Table 1.2. Components isolated from *M. a. paratuberculosis*

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Comment</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CM-1, CM-2</td>
<td>Purified from protoplasmic extracts. Used in ELISA. Component proteins not characterised but comprises 7.8% of total protoplasmic protein. No cross-reaction with <em>M. a. avium</em>, <em>M. phlei</em>, <em>M. fortuitum</em>, <em>Nocardia asteroides</em>. 34-38 kDa range.</td>
<td>Abbas <em>et al.</em>, 1983b; Behymer, 1984; Riemann, 1984; Bech-Nielsen <em>et al.</em>, 1985</td>
</tr>
<tr>
<td>A36</td>
<td>Immunodominant complex of proteins including 34 kDa protein. Induced delayed-type hypersensitivity (rabbits) and T-cell proliferation (mice).</td>
<td>Gilot <em>et al.</em>, 1992; Vannuffel <em>et al.</em>, 1994</td>
</tr>
<tr>
<td>Antigen D (bacterioferritin)</td>
<td>Originally found as one of two precipitin lines with AGID. 400 kDa protein with similarity to <em>E. coli</em> bacterioferritin, <em>M. leprae</em> and <em>M. a. silvaticum</em>. Consists of an assembly of units 17-12 kDa. Also known as the 18 kDa antigen. Isolated from sonicate. N-terminal sequenced.</td>
<td>Hunter <em>et al.</em>, 1990; Brooks <em>et al.</em>, 1991; Sugden <em>et al.</em>, 1991; Inglis <em>et al.</em>, 1994</td>
</tr>
<tr>
<td>Antigen A</td>
<td>The second precipitin line formed with AGID. 31 kDa. Purified from sonicate extracts. N-terminal sequenced. Highly similar to Ag85 complex of BCG, especially Ag85C, for the sequence determined.</td>
<td>Sugden <em>et al.</em>, 1991</td>
</tr>
<tr>
<td>LAM</td>
<td>Smear at 32 to 42 kDa. Polysaccharide. Major cell wall component. Immunodominant by Western blot analysis. Present in culture filtrates. Involved in immune modulation. Has been used in an ELISA for antibody detection. Is not specific but has good reported sensitivity.</td>
<td>Sugden <em>et al.</em>, 1989; Jark <em>et al.</em>, 1997; Reichel <em>et al.</em>, 1999. Also see section 1.8.1</td>
</tr>
<tr>
<td>Protein</td>
<td>Description</td>
<td>Source</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td>P30</td>
<td>30 kDa. Not species-specific. Exists in many <em>Mycobacterium</em> species. Isolated from whole cell lysates, using chromatography. Lymphocyte proliferation response to P30 was found only with sheep infected with live <em>M. a. paratuberculosis</em>, not lysates. Sequence not known.</td>
<td>Burrells <em>et al.</em>, 1995</td>
</tr>
<tr>
<td>70 kDa</td>
<td>Cross-reacts with 71 kDa component (Hsp70) of <em>M. tuberculosis</em>. Highly conserved in all bacterial species. Has been cloned in <em>E. coli</em>. Also similar to <em>M. bovis</em> Hsp65 and <em>E. coli</em> GroEL.</td>
<td>Hance <em>et al.</em>, 1989; Stevenson <em>et al.</em>, 1991</td>
</tr>
<tr>
<td>P43 (ORF 1)</td>
<td>Gene product of IS900 ORF1. 28 kDa when fully processed. Transposase. Has been cloned and expressed in <em>E. coli</em>.</td>
<td>Tizard <em>et al.</em>, 1992</td>
</tr>
<tr>
<td>p35</td>
<td>Identified from <em>E. coli λ</em> -library. Sequence not given. Reacts very strongly to sera of most <em>M. a. paratuberculosis</em>-infected animals, but not <em>M. bovis</em>-infected. Cross-reacts with anti-<em>M. a. avium</em> and anti-<em>M. intracellularae</em> antibodies. No recognition with Crohn’s patient sera. Southern blots show hybridisation only to MAC species.</td>
<td>El-Zaatari <em>et al.</em>, 1997</td>
</tr>
<tr>
<td>p36</td>
<td>From above study. Not recognised by sera from <em>M. a. paratuberculosis</em>-infected animals, but was recognised by Crohn’s, tuberculosis and leprosy patient sera.</td>
<td>El-Zaatari <em>et al.</em>, 1999</td>
</tr>
<tr>
<td>42 kDa</td>
<td>From culture filtrate. May be specific. Not found in <em>M. bovis</em> BCG or <em>M. a. avium</em> serotypes so far screened with Western blots, but only used two <em>M. a. paratuberculosis</em>-cattle sera for this. Not further characterised.</td>
<td>White <em>et al.</em>, 1994</td>
</tr>
<tr>
<td>Protein</td>
<td>Description</td>
<td>Source</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td>Hsp60</td>
<td>Cloned and expressed in E. coli. Highly similar to M. leprae and M. tuberculosis. Cytosolic.</td>
<td>Colston et al., 1994</td>
</tr>
<tr>
<td>Hsp70</td>
<td>Cloned and expressed in E. coli. Cytoplasmic. Recognised by M. a. paratuberculosis-infected cattle sera in ELISA.</td>
<td>Koets et al., 2001</td>
</tr>
<tr>
<td>AhpD (alkyl hydroperoxide reductase D)</td>
<td>From above study. Probably the same as antigen 17 isolated by Gunnarsson and Fodstad. Database searches show a similar protein exists in M. tuberculosis. Weak antibody recognition by M. a. paratuberculosis-infected cattle. IFN-γ stimulator with infected goats.</td>
<td>Gunnarsson &amp; Fodstad, 1979a; Olsen et al., 2000a; 2001</td>
</tr>
<tr>
<td>14 kDa (MPP14)</td>
<td>Secreted. DNA sequence not known, but has been partially sequenced for protein. Similar proteins exist in other mycobacteria. When used in combination with AhpC, in an ELISA, these antigens could distinguish cattle with paratuberculosis from cattle with tuberculosis, but sensitivity was low. Reported to produce innate and adaptive IFN-γ production in cattle.</td>
<td>Olsen et al., 2000b; 2001; Olsen &amp; Storset, 2001</td>
</tr>
<tr>
<td>Csp1</td>
<td>Identified from E. coli λ-library using sera from rabbits immunised with live but not heat-killed M. a. paratuberculosis. Fully sequenced. 25 kDa. Cytoplasmic. Also found in infected macrophage cell line and in in vitro culture. Immunoreactivity with other M. a. paratuberculosis-infected species was not investigated.</td>
<td>Bannantine &amp; Stabel, 2001</td>
</tr>
<tr>
<td>Pks7 (polyketide synthase)</td>
<td>From above study. 70 kDa. Did not demonstrate this in infected macrophage line. Immunoreactivity with other M. a. paratuberculosis-infected species was not investigated.</td>
<td>Bannantine &amp; Stabel, 2001</td>
</tr>
<tr>
<td>35 kDa protein</td>
<td>Fully sequenced. Highly similar to the 35 kDa antigen of M. leprae and M. a. avium. Not further characterised.</td>
<td>Banasure et al., 2001</td>
</tr>
</tbody>
</table>
1.12 Pathogenesis and immune responses

1.12.1 Pathogenesis

Following ingestion into the host tissue, *M. a. paratuberculosis* is preferentially captured by the M cells of the dome epithelium of the lymphoid Peyer’s patches located in the ileum (Momotani et al., 1988; Garcia-Marin et al., 1992). M cells are specialised epithelial cells whose function is to pick up foreign particles from the lining of the gut and deliver them to the lymphoid system. *M. a. paratuberculosis* takes advantage of these cells to enter the gut mucosal barrier. From there they are passed on to and phagocytosed by the subepithelial macrophages (Momotani et al., 1988; Garcia-Marin et al., 1992). This has been observed for both live and dead bacilli (Momotani et al., 1988). The mechanisms used for entry into the macrophage are not known. In *M. tuberculosis* and *M. a. avium*, adherence to macrophages has been shown to proceed via both complement and mannose receptor pathways (Schlesinger, 1993), possibly being mediated by LAM (Bermudez et al., 1991; Chan & Kaufmann, 1994).

*M. a. paratuberculosis*, as with other mycobacteria, reside and replicate inside phagosomes of the macrophage (Valentin-Weigand & Goethe, 1999). The persistence of the organism in this environment is considered to be the crucial process in the establishment and progression of disease. The classical mechanism of killing inside the macrophage involves phagosomes fusing with lysosomes to form mature phagolysosomes and the subsequent generation of oxidative metabolites including hydrogen peroxide, superoxide anion (O₂⁻), and nitric oxide, which are directly toxic to bacteria. Other bactericidal products include lysozyme, lactic acid and hydrolytic enzymes active at low pH. Killing of pathogens residing in the phagosomes of infected murine macrophages is largely through the regulated expression of nitric oxide synthase that produces nitric oxide. Defense mechanisms of many intracellular pathogens against the harmful effects of oxidative intermediates may include the production of superoxide dismutase, catalase (Raynaud et al., 1998) and reductases (Wieles et al., 1997) that neutralise or interfere with the formation of reactive oxygen intermediates inside the macrophage.
A mechanism intensely studied in mycobacteria-macrophage interactions is the ability of these organisms to prevent phagolysosomal fusion and acidification of the phagosomal compartments in which they reside. Studies using the mouse macrophage cell line J774 have shown phagolysosomes containing dead *M. a. paratuberculosis* and *M. a. avium* or the live non-pathogenic mycobacteria *M. smegmatis* and *M. gordonae* have a lower pH than those containing live *M. a. paratuberculosis* or *M. a. avium* (Sturgill-Koszycki et al., 1994; Kuehnel et al., 2001). This has similarly been shown for *M. tuberculosis* (Clemens, 1996). In the case of *M. a. avium*, failure of phagolysomes to acidify appeared to be because of the inability of the vacuolar proton pump to function properly (Sturgill-Koszycki et al., 1994).

Phagosomes containing live *M. a. paratuberculosis, M. a. avium* and *M. tuberculosis* are able to fuse with early endosomes but are capable of continuously restricting fusion to late endosomal vesicles (Frehel et al., 1986a; de Chastellier et al., 1993; Ullrich et al., 2000; Kuehnel et al., 2001). *M. tuberculosis* appears to restrict endosomal fusion by modifying the membrane of the phagosome (Russell, 1995), however it is not known if a similar mechanism is used by *M. a. paratuberculosis* and *M. a. avium*. It appears from these studies that products continuously manufactured by the organisms are responsible for these effects. An effect of the prevention of phagolysosomal fusion appears to be the sequestration of phagosomes from the MHC class II compartment resulting in the prevention of peptides binding to MHC class II molecules for subsequent presentation to CD4+ T-cells (Jensen, 1990; Pancholi et al., 1993). There is evidence that infected macrophages from sheep (Alzuherri et al., 1997) and goats (Navarro et al., 1998) with paratuberculosis have low surface MHC class II expression. More recently, *M. a. paratuberculosis* has been shown to downregulate the expression of MHC class I and class II molecules in infected macrophages *in vitro* and to prevent the upregulation of the expression of these molecules by IFN-γ or TNF-α (Weiss et al., 2001). Inhibition of the expression of MHC class molecules may correspond with the absence of T-cell activation in the very early stages of infection, and may serve to attenuate the immune response, allowing the organism to proliferate in infected cells. Phagocytosed mycobacteria are also sampled for lipids by CD1 molecules possessed by macrophages and are then presented to T-cells (Rind, 2001). Studies with *M. tuberculosis* infection of antigen-presenting cells, has shown expression of CD1 molecules is downregulated, suggesting that this is an immune-evasion mechanism for this pathogen (Stenger et al., 1998).
The lack of T-cell activation may also be due to immunological tolerance, which may exist as a result of the ubiquity of mycobacteria (Juste et al., 1994). In this situation, the immune response is thought to remain unstimulated until contact with antigen is initiated by routes other than through the gut. This may occur through recirculation of infected macrophages or subcutaneous vaccination (Juste et al., 1994). The long persistence of *M. a. paratuberculosis* infection without clinical signs may be partly explained by this hypothesis. Using these mechanisms, mycobacteria are thus able to persist within the macrophages at the site of entry and also are transported to the draining lymph nodes. T-cells are eventually activated and induce inflammatory responses, which serves to contain the pathogen at infected sites (Kaufmann, 1993). As host cell-mediated immunity emerges, intracellular killing of the pathogen occurs by activated macrophages. In this way, T-cell immune responses are essential in limiting *M. a. paratuberculosis* infection (Chiodini & Davis, 1992; Orme et al., 1992; Appelberg, 1994; Vordermeier et al., 1996).

The disease forms of paratuberculosis are commonly referred to as either paucibacillary or multibacillary (Lepper & Corner, 1983). Paucibacillary or tuberculoid, refers to a form of the disease with few bacilli found in lesions, and implies that the host has mounted an effective cell-mediated immune response against the invading organism. The term multibacillary or lepromatous refers to the widespread presence of many bacilli. The terms tuberculoid and lepromatous are descriptions adapted from the terminology used to describe human leprosy (Lepper & Corner, 1983). The occurrence of clinical signs appears to be unrelated to the histological severity of intestinal lesions (Buergelt et al., 1978). It is not known whether these pathological types are stages of the same process or if they are different pathologic processes directed by different T-cell subsets.

### 1.12.2 Immune responses

The spectrum of immune responses is to a large extent, analogous among tuberculosis, leprosy and paratuberculosis (Ellner, 1989; Cocito et al., 1994; Stabel, 2000). Early cell-mediated immune responses are considered to be the principal mechanism involved in protection (Bassey & Collins, 1997; Burrells et al., 1999). Humoral responses appear to have no involvement in protection and high serum antibody concentrations are often recorded in the presence of advanced disease (Chiodini et al., 1984a; Perez et al., 1994; Burrells et al., 1998). There have been extensive studies on the immune response to *M. tuberculosis*, particularly in mice, but many details of *M. a. paratuberculosis* infection remain
uninvestigated. Parallels between tuberculosis and paratuberculosis are difficult to make because of the differences in host range, site of infection and disease outcome.

1.12.2.1 T-cells

Animal models of tuberculosis have demonstrated the CD4+ and CD8+ T-cells (those expressing α/β T-cell receptors) and the effector molecules IFN-γ, IL-12, IL-18 and TNF-α have essential roles in the protective immune response (reviewed by Kaufmann, 2001). The phagosomal compartment of the macrophage is the starting point for MHC class II processing of antigen that leads to CD4+ T-cell activation. Activation of CD8+ T-cells occurs by processing of phagocytosed antigen via an alternative pathway that leads to MHC I presentation (Ramachandra et al., 1999).

The accepted paradigm in mycobacterial infections is that protection is mainly due to antigen-specific, MHC II class-restricted CD4+ T-cells and MHC I class-restricted CD8+ T-cells that express IFN-γ (Stenger & Modlin, 1999). The clinical observation that people with impaired T-cell function, e.g. AIDS patients, have a higher risk of developing clinical tuberculosis, supports this (Stenger & Modlin, 1999) and is confirmed by observations of disseminated mycobacterial disease in IFN-γ or IFN-γ receptor-deficient patients (Holland, 1996; cited by Orme & Cooper, 1999; Doffinger et al., 2000).

In tuberculosis, a large amount of experimental data suggests that CD4+ T-cells play a prominent role in primary immune defense (Kaufmann, 2001). Mice lacking CD4+ T-cells have been observed to have very high susceptibility to tuberculosis (North, 1973; 1974). In addition, CD4+ T-cells from immune mice can provide protection to non-immune mice against challenge with M. tuberculosis (Lefford et al., 1973; Lefford, 1975). However, a growing body of evidence suggests that CD8+ T-cells have a more dominant role in the protective immune response in tuberculosis than previously thought. In mice, depletion of CD8+ T-cells with specific monoclonal antibodies exacerbates tuberculosis (Muller et al., 1987). Sensitised CD8+ T-cells can also provide substantial protection in mice by adoptive transfer (Orme & Collins, 1984).

Numerous cytokines are involved in the Th1 response, but IFN-γ is considered to be key in the activation of macrophages, allowing them to become armed to kill the phagocytosed mycobacteria (Flynn et al., 1993; Stabel, 2000). Activated macrophages secrete IL-12, which stimulates IFN-γ production by CD4+ and CD8+ T-cells. The critical role of these
cytokines has been well established through various studies. In IL-12 knockout mice, *M. tuberculosis* infection progresses uncontrolled in the lungs with negligible inflammatory response (Cooper *et al.*, 1997), indicating IL-12 is necessary for containment of *M. tuberculosis* by granuloma formation. Indeed, a granulomatous response in uninfected livers in mice was observed with IL-12 administration (Silva *et al.*, 1998), confirming that IL-12 is capable of inducing this response in the absence of specific antigen. IL-12 therapy increases the resistance to tuberculosis in mice (Flynn *et al.*, 1995; Kobayashi *et al.*, 1996). In people, disseminated *M. a. avium* infections occur in patients defective in IL-12 secretion or IL-12 receptor signaling (Altare *et al.*, 1998a; 1998b). Disseminated *M. a. avium* infections have also been described in human patients deficient in the IFN-γ receptor (Newport *et al.*, 1996). Similarly, IFN-γ knockout mice are incapable of controlling infection after aerosol exposure to tuberculosis, and disseminated disease occurs (Cooper *et al.*, 1997). Activated macrophages also secrete TNF-α, which is involved in granuloma formation (Kindler *et al.*, 1989).

Reactive CD4+ and CD8+ T-cells also express specific cytolytic activities that lyse infected macrophages. In human beings, the CD8+ cytotoxic T lymphocytes (CTL), predominantly use a granulysin-perforin mechanism which results in apoptosis of the infected macrophage and also may directly kill the intracellular microbe (Stenger *et al.*, 1999). Granulysin is a lytic protein, belonging to a family of serine proteases found in the cytoplasmic granules of CTL and natural killer cells. In the presence of perforin, another molecule present in the cytotoxic granules of CTLs that forms pores in the membrane of infected cells, granulysin can gain access to and kill intracellular *M. tuberculosis*. Scanning electron microscopy evidence suggests that granulysin induces lesions in the bacterial surface of *M. tuberculosis*, resulting in loss of viability (Stenger *et al.*, 1999). CD8+ CTL are capable of also using the Fas-FasL lytic pathway (Ju *et al.*, 1994; White & Harty, 1998). FasL, also known as CD95L, is expressed by CTLs after exposure to antigen. CTLs expressing FasL can lyse target T-cells expressing the Fas receptor and induce apoptosis via this pathway. It is not known if the cytotoxic T-cells harbour both mechanisms of killing or if subsets of CD8+ CTL exist. Cytotoxic lysis of infected cells is also carried out by CD4-CD8- (null) T-cells, which can express either α/β or γδ T-cell receptors (Rind, 2001). These T-cells predominantly use a Fas/FasL ligand binding mechanism to lyse cells. Any viable intracellular bacteria that are released by lysis are then taken up by nearby activated macrophages and effectively killed (Orme, 1987; Kaufmann & Flesch, 1988). In acute *M. tuberculosis* infections in mice, substantial numbers of activated CD4+ and CD8+ T-cells are rapidly accumulated in the lung (Serbina & Flynn, 2001). The CD8+ T-cells have been
shown to have both cytotoxic and IFN-γ-secreting functions (Serbina & Flynn, 2001). It has been hypothesised that a substantial quantity of mycobacterial antigen is therefore presented via the class I pathway, presumably by a phagosome to cytosol transport mechanism (Kovacsovics-Bankowski & Rock, 1995).

There have been few studies done that provide information on the role of lymphocyte subsets in protection to paratuberculosis. These studies have used various animal hosts that have been experimentally or naturally infected or vaccinated. This makes comparisons difficult, however consistent observations have arisen to support the protective role of type I, cell-mediated immune responses in the host (reviewed by Stabel, 2000). Production of IFN-γ was found to be significantly greater in samples of ileum and caecal lymph node from cattle (Sweeney et al., 1998) and sheep (Gwozdz et al., 2000a) subclinically infected compared with those with clinical paratuberculosis. Animals that successfully contained infection were also found to have strong IFN-γ responses from peripheral lymphocytes. Similarly, subclinically infected sheep and cows typically exhibit strong cell-mediated immune responses to Johnin PPD (Chiodini & Davis, 1992; 1993; Burrells et al., 1999) and possess elevated levels of circulating IFN-γ (Stabel, 1996; Bassey & Collins, 1997). Immunohistochemical evidence in naturally infected goats show that CD4+ T-cells are the predominant T-cell type associated with infected macrophages in paucibacillary lesions (Navarro et al., 1998). Experimental infection in young sheep with *M. a. paratuberculosis* resulted in higher densities of CD4+ T-cells in tissues (Begara-McGorum et al., 1998). A study using lymphocytes of *M. a. paratuberculosis*-infected cows depleted of CD4+, CD8+ or γδ T-cells has shown the predominant producers of IFN-γ were CD4+ T-cells, followed by CD8+ T-cells (Bassey & Collins, 1997). γδ T-cells did not significantly contribute to IFN-γ production in the infected cows.

In animals with multibacillary disease, weak IFN-γ responses are typically found (Burrells et al., 1998; Perez et al., 1999; Gwozdz et al., 2000a). Immunohistochemical studies in goats found the numbers of CD4+ T-cells decreased as the severity of the lesion increased and the number of CD8+ T-cells correspondingly increased with the severity of the lesions in these same tissues (Navarro et al., 1998). The activity of the CD8+ T-cells was not investigated. This study also found that epithelioid macrophages containing numerous bacilli were MHC class II-negative, suggesting a lack of CD4+ T-cell activity.
T-cells possessing γδ receptors display cytotoxicity and produce IFN-γ in response to *M. tuberculosis*-infected macrophages (Tsukaguchi *et al.*, 1995). Studies in the tuberculosis mouse model show γδ T-cells have no apparent protective role in the lungs (D'Souza *et al.*, 1997). These cells may possibly be involved in recruitment of monocytes into the granuloma. Conversely, Stenger and Modlin (1999), suggest that these cells are involved in early containment of mycobacterial infections in people.

In young ruminants, 30-80% of the total peripheral blood lymphocyte population consists of γδ T-cells (Hein & Mackay, 1991; Davis *et al.*, 1996). This is in contrast with most other mammals where these cells only comprise 0.5-15% of T-cells in peripheral blood (Kaufmann, 2001). In paratuberculosis, γδ T-cells appear to have an immunoregulatory role. An increase in the proliferative responses of the γδ T-cell subset from infected cattle to killed *M. a. paratuberculosis* was observed following the late, anergic phase of the disease (Chiodini & Davis, 1992; 1993). These same researchers found that peripheral CD4+ T-cells failed to proliferate in response to antigen in vitro in the presence of γδ T-cells. Furthermore, the γδ T-cells were unable to regulate CD4+ T-cell proliferation in the presence of CD8+ T-cells, suggesting γδ T-cell proliferation was in turn regulated by CD8+ T-cells (Chiodini & Davis, 1993). The nature of these apparent inhibitions is unknown. Resistance to paratuberculosis may require the generation of an immunoregulatory population of CD8+ T-cells, which serve to down-regulate γδ T-cell cytotoxic effects on the CD4+ T-cells (Chiodini & Davis, 1993). Others have not found any variations in the γδ T-cell subset with increasing disease, but have found that a decreased number of CD4+ T-cells was associated with an increased number of CD8+ T-cells (Navarro *et al.*, 1998). Similarly, increased reactivity of CD8+ T-cells and γδ T-cells was found in the late anergic phase of the disease, perhaps supporting a suppressive role of these cells (Chiodini & Davis, 1992; 1993; Veazey *et al.*, 1994). Although these studies are not in complete agreement with each other, they point to an immunoregulatory component of the disease.

1.12.2.2 B-cells

In *M. a. paratuberculosis* infection of ruminants, antibody production to paratuberculosis infection, as with other mycobacterial infections, has no demonstrable protective effect (Chiodini *et al.*, 1984a; Dannenberg, 1989; Orme *et al.*, 1993; Perez *et al.*, 1994; Burrells *et al.*, 1998). Clinically diseased cattle often have high levels of serum antibody and severely depressed B and T-cell proliferative responses (Waters *et al.*, 1999). Emergence of humoral
response to *M. a. paratuberculosis* antigen preparations is correlated with the appearance of clinical signs and increased bacterial shedding in *M. a. paratuberculosis*-infected sheep and cattle (Chiodini *et al.*, 1984a; Perez *et al.*, 1994; Clarke & Little, 1996; Perez *et al.*, 1997; Burrells *et al.*, 1998). Elevated levels of TNF-α, IL-1 and IL-6 are associated with multibacillary paratuberculosis (Alzuherri *et al.*, 1996). IL-6, produced by the Th2 subset, is associated with B-cell and plasmacyte development (Coffman & Catty, 1986). Its elevated presence in multibacillary paratuberculosis correlates well with a strong but nonprotective antibody response (Van Snick, 1990). Similarly, elevated levels of mRNA for IL-6 have been detected in multibacillary paratuberculosis, but not in the paucibacillary form (Alzuherri *et al.*, 1996). IL-6 production has been shown to be stimulated by TNF-α and IL-1 (Elias & Lentz, 1990), which are released by activated macrophages. TNF has multiple biological effects including stimulation of the release of IL-1, induction of fever (Sundgren-Andersson *et al.*, 1998), cachexia, enteropathy and granulomatous inflammation (Beutler & Cerami, 1989; Amiri *et al.*, 1992; Garside & Mowat, 1993). IL-1 is also involved in inflammation (Dinarello, 1991). Increased IL-1 activity from peripheral blood monocytes has been recorded in chronic paratuberculous cattle (Kreeger *et al.*, 1991). Many of the clinical signs of Johne's disease can thus be explained by the activity of these cytokines. The persistence of low numbers of mycobacteria for long periods of time, may reflect a balance between host defenses and the ability of the pathogen to resist killing and actively modify the host immune response (Clarke, 1997).

1.13 Summary and aims of the thesis

Johne's disease is caused by the intracellular pathogen *Mycobacterium avium* subspecies *paratuberculosis*. The disease occurs worldwide and is responsible for large economic losses in domestic ruminant herds, predominantly of cattle and sheep. The disease is characterised by a lengthy subclinical phase in which a large majority of animals will not be detected as infected. Such animals represent a significant risk to other members in the herd as they are capable of shedding the bacterium. Testing in herds to detect subclinical infection is carried out to establish prevalence of infection so that control measures can be instituted. However, current tests for detection of infection lack sensitivity, especially in subclinically infected animals. It is therefore important to find new antigens that may allow the development of tests with greater sensitivity and specificity. A European Commission 2001 report of the Scientific Committee of Animal Health and Animal Welfare stated "there is, at present, less optimism about the feasibility of rapid control programmes for paratuberculosis without the prior development of better diagnostic tests".

39
The currently available vaccines for Johne’s disease serve to reduce the clinical signs of the disease, but do not prevent or eliminate infection. These vaccines have a number of serious short-comings, including interference with tuberculosis testing programmes and the ability to cause lesions at the injection site and other tissues in the body. These effects can cause vaccinated animals to be confused with animals with tuberculosis and can cause wastage at the processing stage, resulting in economic losses for farmers. For these reasons, vaccination is not widely practiced. There is a need to develop new, efficacious vaccines that would eliminate or reduce these side effects. The characterisation of new antigens that can stimulate strong cell-mediated immune responses would facilitate potential development of novel vaccines and diagnostic reagents. To date, there have been very few individual proteins that have been isolated and characterised from \textit{M. a. paratuberculosis}. Studies with other mycobacteria such as \textit{M. tuberculosis} and \textit{M. a. avium} have shown that early immune recognition in the infected host is largely aimed at the proteins that are exported by these organisms. This study sought to identify novel exported proteins of \textit{M. a. paratuberculosis} that have immunologic activity. The objectives of the research were as follows:

i) To construct an \textit{M. a. paratuberculosis}-alkaline phosphatase gene fusion library and express it in a heterologous host.

ii) To identify putative exported proteins from the resulting alkaline phosphatase-positive colonies through database searches using the sequences of the cloned partial genes.

iii) To determine the full sequence of selected gene(s).

iv) To express and purify selected protein(s) from a heterologous host.

v) To assess the immunogenicity of selected protein(s) using Neoparasec-vaccinated sheep and naturally infected sheep and cows.
Chapter 2  General materials and methods

2.1  Bacterial strains and plasmids

2.1.1  Bacterial strains

The bacterial stains used in this study are detailed in Table 2.1.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype/Phenotype</th>
<th>Source/Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> DH10B</td>
<td>F&lt;sup&gt;+&lt;/sup&gt; mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 endA1 recA1 deoR Δ(ara, leu)7697 araD139 galU galK nupG rspL λ</td>
<td>Life Technologies Inc., USA Sambrook et al., 1989</td>
</tr>
<tr>
<td><em>E. coli</em> XL1-Blue</td>
<td>supE44 hsdR17 recA1 endA1 gyrA46 thi rellA1 lac F&lt;sup&gt;<em>&lt;/sup&gt;[proAB&lt;sup&gt;</em>&lt;/sup&gt; lacI&lt;sup&gt;*&lt;/sup&gt; lacZΔM15 Tn10(ter&lt;sup&gt;f&lt;/sup&gt;)]</td>
<td>Stratagene, USA Sambrook et al., 1989</td>
</tr>
<tr>
<td><em>M. bovis</em></td>
<td></td>
<td>ATCC 35746</td>
</tr>
<tr>
<td><em>M. bovis</em></td>
<td></td>
<td>ATCC 35725</td>
</tr>
<tr>
<td><em>M. bovis</em></td>
<td></td>
<td>ATCC 35726</td>
</tr>
<tr>
<td><em>M. bovis</em></td>
<td>canine clinical isolate</td>
<td>Massey University, Palmerston North, NZ Gay et al., 2000</td>
</tr>
<tr>
<td><em>M. bovis</em> (KML)</td>
<td>bovine clinical isolate</td>
<td>Fielding Meat Works, Fielding, NZ</td>
</tr>
<tr>
<td><em>M. bovis</em> BCG (Pasteur)</td>
<td>attenuated vaccine strain</td>
<td>ATCC 35734</td>
</tr>
<tr>
<td><em>M. bovis</em> BCG (Glaxo)</td>
<td>attenuated vaccine strain</td>
<td>ATCC 35741</td>
</tr>
<tr>
<td><em>M. bovis</em> BCG (Japan)</td>
<td>attenuated vaccine strain</td>
<td>ATCC 35737</td>
</tr>
<tr>
<td><em>M. fortuitum</em></td>
<td></td>
<td>ATCC 6841</td>
</tr>
<tr>
<td><em>M. gordonae</em></td>
<td></td>
<td>ATCC 14470</td>
</tr>
<tr>
<td><em>M. intracellularae</em></td>
<td></td>
<td>ATCC 35848</td>
</tr>
<tr>
<td><em>M. kansasii</em></td>
<td></td>
<td>ATCC 12478</td>
</tr>
<tr>
<td><em>M. marinum</em></td>
<td></td>
<td>ATCC 927</td>
</tr>
<tr>
<td><em>M. a. paratuberculosis</em></td>
<td>attenuated vaccine strain used in Neoparasec (Merial, France)</td>
<td>Central Veterinary Laboratory Weybridge, UK</td>
</tr>
</tbody>
</table>
M. a. paratuberculosis bovine clinical isolate Massey University, Palmerston North, NZ ATCC 53950
M. phlei field isolate Massey University, Palmerston North, NZ
M. scrofulaceum ATCC 19981
M. smegmatis mc²155 electroporation-competent strain of M. smegmatis Snapper et al., 1990 ATCC 607
M. terrae ATCC 15755
M. tuberculosis H₃₇Ra ATCC 25177
M. tuberculosis New Zealand fur seal clinical isolate Massey University Palmerston North, NZ Hunter et al., 1998

2.1.2 Plasmids

The plasmids used in this study are detailed in Table 2.2.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Source/Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pJEM11</td>
<td>*E. coli/mycobacterial shuttle vector containing a truncated phoA gene, kan'</td>
<td>Lim et al., 1995</td>
</tr>
<tr>
<td>pMIP12</td>
<td>*E. coli/mycobacterial shuttle vector containing a C-terminal histidine-tag for expression in M. smegmatis. kan'</td>
<td>Professor Brigette Gicquel Pasteur Institute, Paris</td>
</tr>
<tr>
<td>pMIP-p22</td>
<td>pMIP12 containing the p22 ORF</td>
<td>this study</td>
</tr>
</tbody>
</table>
2.2 Bacterial growth and storage conditions

2.2.1 E. coli

*E. coli* strains were routinely grown in LB (Appendix 1) broth or on LB plates containing the appropriate antibiotic/supplement (see Table 2.3) at 37°C under aerobic conditions. Recombinant *E. coli* were grown overnight with vigorous aeration at 37°C in 5 ml of LB broth containing the appropriate antibiotic. For PhoA selection, kanamycin (kan) and 5-bromo-4-chloro-3-indoyl phosphate (BCIP) were added to plates. For long-term storage, overnight liquid cultures were transferred to 1.7 ml cryotubes (Nunc, Nalgene Nunc International, Denmark) and sterile glycerol was added to 20% (v/v). The mixtures were briefly vortexed then stored at -70°C.

2.2.2 M. smegmatis

*M. smegmatis* cultures were routinely grown in LB broth or on LB agar containing the appropriate antibiotic/supplement (see Table 2.3). For PhoA selection, cultures were grown on LB plates containing kan and BCIP. Cultures were generally grown from three to five days at 37°C under aerobic conditions as described for *E. coli*. For protein isolation, *M. smegmatis* containing pMIP12 constructs were grown in 600 ml of Sauton’s broth or modified Middlebrook 7H9 broth (Appendix 1) containing kan in 2 L flasks. For storage of recombinant *M. smegmatis* containing pJEM11 constructs, colonies were scraped from LB/kan/BCIP plates and transferred to 1.7 ml cryotubes containing 1 ml of 20% (v/v) glycerol in LB broth. The mixtures were vigorously vortexed to break up cell clumps, then frozen at -70°C. Alternatively, liquid cultures of *M. smegmatis* were stored as described above for *E. coli*. 
Table 2.3. Antibiotics and supplements used in microbiological media

<table>
<thead>
<tr>
<th>Antibiotic/Supplement</th>
<th>Stock concentration (mg/ml)</th>
<th>Final concentration (μg/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ampicillin (Sigma, USA)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>kanamycin sulphate (Roche Molecular Biochemicals, Germany)</td>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>BCIP (Sigma, USA)</td>
<td>40</td>
<td>25</td>
</tr>
</tbody>
</table>

2.2.3 M. a. paratuberculosis

Cultures of M. a. paratuberculosis strain 316F were routinely grown in 500 ml of modified Middlebrook 7H9 broth (Appendix 1) in 2 L flasks grown under standard aerobic conditions (see 2.2.1). The cultures were seeded with 5 ml of freshly grown mid-log phase culture (OD_{600} = 1.0) in the same media. Generally, cultures grown in this manner took three weeks to reach mid-log phase, at which time they were harvested. For growth on solid media, Middlebrook 7H10 agar containing Mycobactin J was used (Appendix 1).

2.3 DNA isolations

2.3.1 Isolation of plasmid DNA from E. coli

Plasmids were isolated from E. coli using either the BRESAspin Plasmid Mini Kit (Bresatec, Australia) or High Pure Plasmid Isolation Kit (Roche Molecular Biochemicals, Germany) or Concert Nucleic Acid Purification System (Life Technologies Inc., USA) or QIAPrep Spin Miniprep Kit (Qiagen, Germany). Briefly, 5 ml cultures were grown under standard conditions in 20 ml test tubes overnight. From this, 2-4 ml was harvested by centrifugation and treated according to the manufacturer’s instructions. Plasmid DNA was routinely eluted in 50 μl of 10 mM Tris.Cl (pH 8.0).

2.3.2 Isolation of genomic DNA from mycobacterial species

For PCR reactions, genomic DNA was extracted either by physical disruption of cells by directly adding bacterial colonies to PCR mixtures followed by heating to 95°C for 10 min or by chemical lysis followed by phenol:chloroform extraction.
Chapter 2

Chemical lysis and subsequent purification of genomic DNA was as follows. Mycobacterial cultures were grown at 37°C on 7H10 agar slopes containing OADC enrichment (Difco, MI, USA) as per the manufacturer's recommendations (see Appendix 1), in 50 ml screw-capped disposable centrifuge tubes. Cultures were grown until colonies were easily visible, which varied from three days to several weeks for rapid- and slow-growing species, respectively. The colonies were scraped from the agar with a sterile loop and added directly to 1.5 ml of 50 mM Tris.Cl (pH 8.0), 25 mM EDTA containing 25% sucrose in a 15 ml screw-capped polypropylene centrifuge tube. The mixture was then heated to 70°C for 2 h to kill the cells. To the cooled mixture, 500 μg/ml lysozyme was added and incubated for 2 - 18 h at 37°C. To this, 4 ml of 100 mM Tris.Cl (pH 8.0), 1% (w/v) SDS, 400 μg/ml proteinase K (Roche Molecular Biochemicals, Germany) was added and incubated for 2 h at 55°C. The mixture was cooled and 0.1 ml of 5 M NaCl was added. The DNA was extracted with phenol:chloroform using a 15 ml Phase Lock Gel™ Light apparatus (Eppendorf, Germany). Briefly, this consists of the addition of a gel substance to the mixture that assists in separation of the aqueous, DNA-containing phase from the organic, impurity-containing phase upon centrifugation at 12,000 x g for 5 min. The aqueous phase was washed in the Phase Lock tube with an equal volume of chloroform to remove residual phenol and the phases were separated by centrifugation as above. The resulting aqueous phase was collected and the DNA was precipitated with ethanol at -20°C for 30 min. After centrifugation at 13,000 x g for 40 min, the DNA pellet was washed with 70% ethanol and air-dried. The DNA was resuspended in 1 - 2 ml of sterile water and stored at 4°C.

2.4 DNA manipulations and cloning procedures

2.4.1 Restriction endonuclease digestions

Typically, restriction enzyme digests were performed in a final volume of 20 - 50 μl containing 0.1 - 2 μg of DNA. Enzymes were purchased from Roche Molecular Biochemicals (Germany) or Life Technologies Inc. (USA) and used with the supplied buffers at an amount of 5 - 20 units at 37°C for 2 to 18 h unless otherwise stated.

2.4.2 Electrophoresis

Agarose gels (usually of 1% (w/v)), were prepared by dissolving agarose (Life Technologies Inc., USA) in TAE buffer (Appendix 1) by heating in a microwave followed by cooling to 55°C before pouring. Electrophoresis was performed in TAE buffer at 70 - 100V using a
Bio-Rad horizontal electrophoresis unit and a Bio-Rad model Power Pac 300 power supply. Samples were loaded in 0.2 volumes of DNA loading dye (Appendix 1) along with the appropriate DNA size standard. Gels were transferred to a 10 μg/ml solution of ethidium bromide in either distilled water or TAE buffer for 15 min. The DNA was visualised and photographed under UV light using a Bio-Rad Gel Doc 2000 system. The molecular size of the DNA fragments was estimated by comparison with either 100 bp DNA Ladder (15628-019) or 1 Kb Plus DNA Ladder (10787-018) or 1 Kb DNA Ladder (15615-016) size standards (Life Technologies Inc., USA).

2.4.3 Extraction from agarose gels

Restriction endonuclease-digested DNA or PCR products were electrophoresed in agarose gels as described above. Following ethidium bromide staining, the desired DNA bands were excised under minimal UV light exposure using a scalpel blade. Purification was carried out using either a Concert Rapid Gel Extraction System kit (Life Technologies Inc., USA) or QIAquick Gel Extraction Kit (Qiagen, Germany) according to the manufacturer’s instructions. Purified products were resuspended in 30 – 50 μl of distilled water.

2.4.4 Southern blotting and hybridisations

Approximately 1 μg of restriction endonuclease-digested genomic DNA fragments were electrophoresed in 0.7% agarose gels and transferred to nylon membranes (Biodyne B, Gelman, Pall Corporation, USA) by capillary transfer using standard procedures (Sambrook et al., 1989). The DNA was fixed to the membrane by exposure to UV light for 4 min using a Bio-Rad Gel Doc 2000 transilluminator (Bio-Rad, USA). Fixed membranes were prehybridised in heat-sealed plastic bags with DIG Easy Hyb buffer (Roche Molecular Biochemicals, Germany) at 42°C for 2 h with constant shaking in a Hot Shaker water bath (Bellco Biotechnology, NJ, USA). Hybridisation was done in the same buffer at 2.5 ml/100 cm² with 25-50 ng/ml of denatured DIG-labeled probe DNA (see 2.4.3.1), at 40°C for 18 h. The hybridised membranes were washed 2 x 5 min in 2 x SSC (Appendix 1), containing 0.1% (w/v) SDS at room temperature, followed by 2 x 5 min washes in 0.7 x SSC containing 0.1% (w/v) SDS at 68°C with constant shaking. Immunodetection of hybridised probe was achieved using the DIG system (Roche Molecular Biochemicals, Germany). Briefly, washed membranes were incubated in 1 ml/cm² blocking solution (Roche Molecular Biochemicals, Germany) for 60 min at room temperature with shaking. The blocking solution volume was reduced to 20 ml/100cm² and anti-DIG antibody conjugated to alkaline phosphatase (Roche
Molecular Biochemicals, Germany) was added to a final dilution of 1:10,000 according to the manufacturer's recommendations. The blots were developed by chemiluminescence with CSPD or CPD-Star substrate (NEN, MA, USA). The developed blots were exposed to radiographic film (BioMax MR, Kodak, USA) for 5 min to 18 h, depending on signal intensity, in the presence of a single intensifying screen (Kodak Lanex Regular, Kodak, USA). Film was developed in an automated processor (Kodak RP X-OMAT Processor Model M6B).

2.4.4.1 DNA probe preparation

All probes were labeled by the incorporation of DIG-labeled dUTP (DIG-11-dUTP, Roche Molecular Biochemicals, Germany) during PCR (see 2.6.2). DIG-11-dUTP (1 573 152, Roche Molecular Biochemicals, Germany) was added to a final concentration of 20 μM in a reaction volume of 50 μl and dTTP was adjusted to a final concentration of 80 μM. All other deoxynucleoside triphosphates were added to 100 μM. To estimate the purity and yield of DIG-labeled product, approximately 2 μl of the reaction was electrophoresed in agarose gels alongside a mass ladder for quantitation (10068-013 Low Mass DNA Ladder, Life Technologies Inc., USA). Due to the presence of DIG, the PCR products routinely appeared larger than unlabeled products. For quantitation of DIG-incorporation in probes, side-by-side filter spot tests, ranging from 0.01 pg to 10 pg, were carried out as per the manufacturer's recommendations. Labeled PCR products were stored at -20°C until used for hybridisation (see 2.4.3.1).

2.4.4.2 Removal of probe from Southern blots

Probe DNA was removed from hybridised blots with 2 x 15 min washes in 0.2 M NaOH, 0.1% (w/v) SDS at 37°C. Membranes were neutralized with three, 15 min washes in 2 x SSC (Appendix 1). Membranes were stored wet in heat-sealed plastic bags at 4°C.

2.4.5 Polymerase chain reactions

2.4.5.1 Primer design

Several factors were taken into consideration when designing primers for PCR and sequencing reactions. Except for cases where a restriction enzyme site was to be incorporated for subcloning purposes, all primers were 100% specific for the template.
When a restriction enzyme site was incorporated, usually two to four nucleotides were added onto the 5' end following the restriction site to increase efficiency of restriction endonuclease cleavage. These additions were chosen based on genomic template sequence as opposed to random additions in order to facilitate primer annealing during PCR. The Amplify version 1.2 software programme was used to check for alternative hybridisation sites present in the template and the presence of primer-dimer formations. Primers were designed to be of adequate length to give good specificity, usually between 20 and 30 base pairs. Primer pairs were matched as close as possible for length and melting temperatures. Primers were synthesized by Life Technologies Inc. (USA) and were resuspended in distilled water to a final concentration of 100 μM, calculated using physical data provided by the manufacturer (see below). All resuspended primers were stored in aliquots at -20°C. To determine the volume in which to resuspend the oligonucleotides, the following formula was used:

\[
\text{Total OD (absorbance at 260 nm)} = \frac{\text{volume of water (ml)}}{(EM) \times \text{molar concentration required}}
\]

where EM is the molar extinction coefficient

### 2.4.5.2 PCR conditions

PCR reactions were routinely carried out using *Taq* DNA polymerase (Life Technologies Inc., USA). For cloning or sequencing of PCR products, Expand High Fidelity PCR System (Roche Molecular Biochemicals, Germany) or Platinum *Pfx* DNA polymerase (Life Technologies Inc., USA) were used. Reactions were carried out according to the manufacturer’s recommendations in the presence of 1.5 mM MgCl₂ for *Taq* and Expand High Fidelity or 1.5 mM MgSO₄ for Platinum *Pfx*. Deoxynucleoside triphosphates were added to a final concentration of 100 μM. Primers were added to a final concentration of 0.4 μM. Dimethylsulfoxide was included in all PCR reactions to a final concentration of 10% (v/v). The following cycling parameters were typical for a standard PCR reaction: initial denaturation at 95°C for 10 min, followed by 35 cycles of denature at 94°C for 30 s, anneal at 60°C for 30 s, and chain elongation at 72°C for 30 s, followed by a final elongation step at 72°C for 10 min. For Platinum *Pfx*, chain elongation was carried out at 68°C according to the manufacturer’s instructions. Amplification reactions were carried out in a Perkin Elmer 9600 PCR thermocycler using 0.2 ml thin-walled PCR tubes (Life Technologies Inc., USA). PCR products were analysed by agarose gel electrophoresis (see 2.4.1).
For DIG-labeling of PCR products (see 2.4.3.1), the following cycling parameters were used with *Taq* polymerase: 95°C for 10 min, followed by 35 cycles of 94°C for 45 s, 60°C for 1 min, and 72°C for 2 min, followed by a final elongation step at 72°C for 10 min.

2.4.6 DNA sequencing

Automated sequencing was carried out by Massey University DNA Analysis Service using an ABI Prism® 377 DNA Sequencer (Applied Biosystems, USA). For a single sequencing reaction 12 μl each of 0.8 pmol/μl primer and 500 ng/μl plasmid template or 30 ng/μl PCR product template were used. The reactions were performed using an ABI Prism® BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, USA). This kit uses dye-labeled dideoxynucleotides in which a specific fluorescent dye is covalently attached to each base. Products can then be visualised as fluorescent colour-coded bands that are collected (ABI Prism® Data Collection System) and interpreted by computer software (ABI Prism® DNA Sequencing Analysis Software).

2.4.7 Ligations

Where possible, inserts and vectors were digested with two restriction endonucleases to produce complementary protruding termini compatible only for ligation between vector and insert. Approximately 100 ng of vector DNA was added to insert DNA in a molar ratio of 1:3 vector to insert. Generally, ligations were done in 20 μl volumes containing 1 unit of T4 DNA ligase in supplied ligase buffer (Roche Molecular Biochemicals, Germany). Incubations were done either at room temperature for approximately 20 min, or, for convenience, overnight at room temperature. Where fragments of DNA carrying identical protruding termini were to be cloned, restriction enzyme-digested vector DNA was treated with calf intestinal alkaline phosphatase (Roche Molecular Biochemicals, Germany) using standard procedures (Sambrook *et al.*, 1989). Test ligations of the vector only were performed and used in transformations to determine the efficiency of the phosphatase treatment.
2.5 Bacterial transformations

2.5.1 Preparation of electrocompetent *M. smegmatis*

Electrocompetent *M. smegmatis* mc²155 cells were prepared by modification of a protocol originally described by (Snapper *et al.*, 1990). Briefly, a single colony from a fresh plate was inoculated into 5 ml of 7H9 media (Difco, USA), made to the manufacturer's recommendations, and grown with aeration in a 50 ml tube for three days at 37°C. Of this, 4 ml was transferred to 300 ml of the same media and the culture was grown as above to mid-log phase (OD$_{600nm}$ =1.0). The culture was then incubated on wet ice for 2 h. Cells were pelleted at 3,000 x g for 5 min at 0°C. The pellet was washed with 300 ml of ice-cold sterile 10% (v/v) glycerol, and pelleted as above. The pellet was resuspended in 800 μl of ice-cold sterile 10% (v/v) glycerol and immediately aliquotted in 100 μl volumes and frozen at -70°C.

2.5.2 Transformation of *E. coli*

Commercially purchased electrocompetent cells (ElectroMAX DH10B, Life Technologies Inc., USA) were transformed with purified plasmid or ligation mixture, or by electro-transfer of plasmid directly from *M. smegmatis*. Ligation mixtures and plasmids were dialysed against distilled water for 20 min using 0.025 μM pore size filter discs (Millipore Corporation, MA, USA) prior to electroporation to remove salts. Routinely, 20 ng of purified plasmid or 100 ng of ligation mix were electroporated with 50 μl of cells in a 0.2 cm electroporation chamber (Bio-Rad, USA) using a Bio-Rad Gene Pulser apparatus set to 1.8 kV, 25 μF capacitance and 200 ohms resistance. Transformed *E. coli* were incubated in 200 - 500 μl of LB broth for 1 h at 37°C prior to spreading on LB agar plates containing the appropriate antibiotic/supplement (Table 2.3). For transference of plasmid DNA from *M. smegmatis* to *E. coli*, a single colony, approximately 1 mm in diameter, was mixed with 25 μl of electrocompetent *E. coli* DH10B cells in an electroporation chamber. The mixture was electroporated and plated out as above.

Chemically competent DH5α (Library Efficiency DH5α competent cells, Life Technologies Inc., USA) cells were transformed with purified plasmid or ligation mixtures in the amounts stated above. Fifty microlitres of competent cells were mixed with DNA in a precooled 1.5 ml microtube and incubated on ice for 20 min. The mixture was heat-shocked at 42°C for 90
s then placed on ice for 2 min. To this was added 300 μl of LB broth and the mixture was incubated for 1 h at 37°C before plating on LB agar containing the appropriate antibiotic/supplement (Table 2.3).

2.5.3 Transformation of *M. smegmatis*

*M. smegmatis* were transformed with purified plasmid. Fifty to 100 μl of electrocompetent cells (see 2.5.1) were mixed with approximately 20 to 50 ng of dialysed plasmid DNA (see 2.5.2 above) and electroporated as for *E. coli*. To the mixture, 400 μl of LB broth was added and incubated for 2 - 3 h at 37°C before spreading onto LB agar. For selection of PhoA, cells were plated on LB agar containing BCIP and kan.

2.6 Protein isolations

2.6.1 Preparation of cell lysates

Cell lysates were prepared from cultures grown as described (see 2.2.2 and 2.2.3). Generally, cultures were harvested at mid-log phase, as determined by reading the optical density at 600 nm of a 1 ml sample of culture in a double beam spectrophotometer (Helios α, Unicam, UK). The cells were pelleted by centrifugation at 3,500 x g for 20 min. Cells were washed three times by repeated resuspension in approximately 40 ml of PBS followed by centrifugation. Washed pellets were frozen at -70°C. For each batch of culture (500 – 600 ml), 40 ml of loading buffer (Appendix 1) was added to the frozen pellet. After thawing, the suspension was transferred to a 50 ml screw-capped polypropylene centrifuge tube and sonicated on ice using a 13 mm diameter probe with amplitude set at 55% for 4 x 1 min using an ultrasonic liquid processor (Vibracell model VCX 500, Sonics, Newtown, CT, USA). The sonicated material was centrifuged at 14,000 x g for 30 min to pellet insoluble material and the supernatant was collected. The pellet was washed three times by resuspension in approximately 40 ml of PBS followed by centrifugation. The pellet was frozen at -70°C. For isolation of recombinant protein the procedure was carried out a second time. For purification of recombinant protein, the supernatant was used directly for affinity chromatography (see 2.6.3). The insoluble sonicated pellet was stored at -70°C.
2.6.2 Preparation of culture filtrates

2.6.2.1 M. smegmatis

*M. smegmatis* cultures were grown to mid-log phase (OD_{600 nm} = 1.0) in Sauton's broth or modified 7H9 media (Appendix 1) containing kan as described (see 2.2.2). Cells were pelleted by centrifugation at 5,000 x g and the culture supernatant was sterilised by filtration through a Steritop filtration device containing a 0.22 μm filter (Millipore, USA). The supernatant was concentrated by ultrafiltration in a stirred cell apparatus (Millipore, USA) using a 3,000 molecular weight cut-off Diaflo YM membrane (Amicon, USA). This was followed by further concentration by centrifugation at 6,000 x g using a 3,000 molecular weight cut-off Centriplus 3 filtration device (Amicon, USA). The resulting 480 fold concentrated preparation was stored at -20°C.

2.6.2.2 M. *a. paratuberculosis*

 Cultures of *M. a. paratuberculosis* strain 316F were grown to mid-log phase (OD_{600 nm} = 1.0) as described (see 2.2.3). Prior to harvesting, a sample of culture was plated on blood agar and incubated 18 h at 37°C to check for contamination. Cultures were harvested by centrifugation at 10,000 x g for 20 min. The culture supernatant was sterilised, concentrated and buffer exchanged into PBS to a 200 fold concentration as described in section 2.6.2.1. To check for contamination, a sample of the resulting preparation was plated on blood agar, as above, and the preparation was stored at -20°C.

2.6.3 Ni^{2+}-affinity chromatography

Histidine-tagged recombinant proteins were semi-purified from cell sonicates with Ni^{2+}-affinity chromatography using HiTrap chelating columns (Amersham Pharmacia Biotech, UK). Briefly, these columns contain cross-linked agarose beads to which imino-diacetic acid has been coupled. When charged with a suitable metal ion, such as Ni^{2+}, proteins that contain exposed histidine residues are selectively retained by the beads. Bound proteins can be competitively eluted with buffer containing imidazole. Imidazole is a derivative of histidine and serves to displace these residues on the column.

All solutions were passed through the columns with the aid of a syringe or peristaltic pump. Briefly, the columns were rinsed with five column volumes of sterile water before use, then
charged with two column volumes of 250 mM NiCl₂ followed by five column volumes of sterile water. Charged columns were equilibrated with two column volumes of loading buffer (see Appendix 1). One column volume of cell sonicate supernatant, prepared in loading buffer, was loaded onto columns at a rate of approximately 2 ml/min. Unbound proteins were removed with five column volumes of loading buffer. To determine the imidazole concentrations necessary to elute contaminating proteins, a series of washes using varying amounts of imidazole (40 mM to 1 M) in start buffer (Appendix 1), were performed. Fractions were collected manually and 15 µl samples from selected fractions were used for SDS-PAGE analysis (see 2.7.1) to locate the recombinant protein. Pooled fractions were concentrated and buffer exchanged into PBS by centrifugation at 3,500 x g in a 10,000 molecular weight cut-off Centricon 10 filtration device (Amicon, USA). For reuse, the columns were immediately stripped by the addition of two column volumes of 250 mM EDTA (pH 8.0). This removes bound Ni²⁺ ions and any remaining protein. The columns were rinsed with at least 10 column volumes of sterile water to remove EDTA. Columns were sealed and stored at 4°C.

2.6.4 Size exclusion chromatography

To further purify recombinant protein, pooled preparations of protein resulting from Ni²⁺-affinity chromatography were used in size exclusion chromatography. The samples were first concentrated by centrifugation at 3,000 x g to a volume of approximately 2 ml using a 10,000 molecular weight cut-off filter (Centricon 10, Amicon, USA), and filtered through a 15 mm diameter 0.45 µm syringe filter (Gelman, Pall Corporation, USA). A Sephacryl S-100 (16/60) gel filtration column (Amersham Pharmacia Biotech, UK) was connected to a peristaltic pump and the flow rate was fixed at 1 ml/min. The column was equilibrated with PBS, as per the manufacturer's recommendations, and the sample was applied with a syringe. Chromatography was carried out with approximately 60 ml of PBS and fractions of approximately 2 ml were collected manually throughout the procedure. Following separation, samples were analysed by SDS-PAGE to locate the recombinant protein. The identified fractions were pooled and the protein was concentrated as above.
2.7 Protein analyses

2.7.1 Polyacrylamide gel electrophoresis

For polyacrylamide gel electrophoresis (PAGE), gels were cast in a multiple casting chamber (SE 200 series, Hoefer Scientific Instruments, USA) according to the manufacturer’s instructions. Protein separations were carried out in 0.1% (w/v) sodium dodecyl sulphate PAGE gels (SDS-PAGE gels) using a vertical slab unit (Mighty Small SE 250, Hoefer Scientific Instruments, USA) according to the manufacturer’s instructions. Resolving gels, usually consisting of 15% (w/v) acrylamide/bis, were prepared with a final concentration of 0.375 M Tris.Cl (pH 8.8), 0.1% (w/v) SDS, using an acrylamide:bis ratio of 30%:2.67%. Stacking gels of 4% (w/v) acrylamide/bis in 0.125 M Tris.Cl (pH 6.8), 0.1% (w/v) SDS were similarly prepared. All samples were mixed with one volume of 2x SDS-PAGE loading buffer containing β-mercaptoethanol (Appendix 1), boiled for 5 min, cooled on ice and briefly centrifuged prior to loading the gel. The electrophoresis buffer consisted of 0.192 M glycine, 0.02 M Tris base and 0.1% (w/v) SDS. Electrophoresis was routinely carried out at a constant voltage of 130V, using a Bio-Rad model 3000Xi power supply (Bio-Rad, USA). For estimation of molecular weight, Bio-Rad prestained Precision Protein Standards (160-0372, Bio-Rad, USA) were included on each gel.

Following electrophoresis, proteins were visualised by staining with 0.125% (w/v) Coomassie Brilliant Blue R-250 (Sigma, USA) in 50% methanol, 10% acetic acid for 15 min at room temperature with shaking. Gels were destained with several changes of 30% methanol, 10% acetic acid until stained bands were easily visualised. For a permanent record, the gels were sandwiched between two wet cellophane sheets and hot air-dried in an Easy Breeze gel drier (Hoefer Scientific Instruments, USA).

2.7.2 Western blotting and immunodetections

Protein samples were routinely run on 15% SDS-PAGE gels and transferred to PVDF (Gelman, Pall Corporation, MI, USA) or nitrocellulose (Bio-Rad, USA) membranes, using a mini Trans-Blot electrophoretic transfer cell (Bio-Rad, USA), as per the manufacturer’s instructions. Transfer was done at 100V for 1 h in cold transfer buffer (Appendix 1) as follows: PVDF membrane was briefly soaked in methanol then rinsed with several changes of water. Nitrocellulose membranes were soaked in water. The membranes, filter paper, pads and gels were then equilibrated separately in transfer buffer for approximately 20 min prior to
transfer. Following blotting, the membranes were washed briefly in distilled water, then stained with 0.2% (w/v) Ponceau S (Sigma, USA) in 30% (v/v) trichloroacetic acid for 10 min with shaking. Membranes were then washed with distilled water until protein bands could be easily visualised. For a permanent record, stained, wet membranes were sandwiched between two acetate sheets and photocopied. The membranes were then destained by multiple washes in distilled water, air dried and either stored dry in heat-sealed bags or used immediately for immunodetection.

2.7.2.1 Immunodetection of Western blots

Western blots were incubated in 3.5 ml/cm² blocking solution (Appendix 1) for 1 - 2 h at room temperature with shaking. After this time, excess blocking solution was removed so that membranes were freely moving in a minimal volume and primary antibody was added at an appropriate dilution. Membranes were incubated for 2 h with shaking at room temperature or, in the case of serum, overnight at 4°C. Membranes were then washed 6 x 5 min in wash buffer (Appendix 1). Where appropriate, secondary antibody conjugated to peroxidase (POD) was added at an appropriate dilution in a minimal volume of fresh blocking solution and the membranes were incubated for 1 - 2 h at room temperature and washed as above. Blots were developed by chemiluminescence with the substrate SuperSignal West Femto (Pierce, IL, USA), as per the manufacturer’s recommendations in heat-sealed bags and exposed to radiographic film (BioMax MR, Kodak, USA) in the presence of a single intensifying screen (Kodak Lanex Regular, Kodak). Film was developed using an automated processor (Kodak RP X-OMAT Processor Model M6B).

2.7.2.2 Removal of antibody from Western blots

To reuse Western blots, antibodies from developed blots were removed immediately after final radiographic exposure by incubation in excess 62.5 mM Tris.Cl (pH 6.7), 2% (w/v) SDS, 100 mM β-mercaptoethanol at 50°C for 30 min. Membranes were washed in several changes of excess distilled water and used immediately or stored dry in heat-sealed plastic bags at 4°C.

2.7.3 N-terminal protein sequencing

Automated Edman degradative N-terminal sequencing was carried out by Massey University Protein Sequencing Services using a pulse liquid phase sequenator (Model 476A, Applied
Biosystems, USA). This instrument performs fully automated sequencing by sequential removal of the N-terminal amino acid as a phenylthiohydantoin-derivative. Derivatives were separated by HPLC and the data collected and analysed using a Model 610A Data Analysis Module (Applied Biosystems, USA). Protein to be sequenced was electrophoresed on 15% SDS-PAGE gels and transferred to PVDF membrane as described (see 2.7.2). Following transfer, the membrane was stained with Ponceau S (see 2.7.2) and the desired band(s) were excised and destained in distilled water. Approximately 1 pmol (25 ng for a protein of 25 kDa) of protein was used for sequencing.

### 2.7.4 Estimation of protein concentration and molecular weight

Protein concentrations were estimated using the Bio-Rad Protein Assay reagent according to the manufacturer's instructions in 1 ml polystyrene cuvettes using a double beam spectrophotometer (Helios α, Unicam, UK). This assay is based on the Bradford assay (Bradford, 1976) which relies on the binding of Coomassie Brilliant Blue G-250 dye to basic and aromatic amino acids. When bound, the peak absorbance of the dye shifts from 465 nm to 595 nm. Briefly, protein samples were diluted in 800 µl of distilled water to which 200 µl of the reagent was added. The samples were vortexed and the absorbance at 595 nm was read after 5 min incubation at room temperature. Standard dilutions of bovine serum albumin (Pierce, IL, USA), ranging from 1.2 µg to 10 µg were included in parallel with each protein assay to obtain a standard curve. All samples were assayed in duplicate, using the appropriately diluted reagent as a blank.

Alternatively, the purified recombinant protein concentration was estimated by absorbance at 280 nm using matched quartz cuvettes and the amount calculated using the theoretical absorbance of a 1 mg/ml solution of the protein, as determined by ProtParam tool through the ExPASy Molecular Biology Server (see 2.11).

\[
\text{Concentration (mg/ml)} = \frac{\text{measured } A_{280\text{ nm}}}{\text{theoretical } A_{280\text{ nm}}} \times \text{dilution factor}
\]

Estimation of protein molecular weight in SDS-PAGE gels was calculated by comparison to a standard curve constructed by plotting the log molecular weight versus the migration distance of protein standards.
2.8 IFN-γ assays

IFN-γ assays were performed using the whole blood Bovigam™ EIA bovine interferon test, which is suitable for the detection of ovine IFN-γ. The test was performed as per the manufacturer’s (Commonwealth Serum Laboratories, Australia) instructions, with some modifications. Briefly, blood samples were collected in lithium heparin tubes and processed within 4 hours of collection. One millilitre aliquots of blood were dispensed into 24-well tissue culture trays. Routinely, antigens were tested in duplicate. Each antigen was added in a standard volume of 67 μl to the blood aliquots and mixed for 5 min on a rotating platform shaker. The trays were then incubated 22 h at 37°C in a humidified atmosphere with 5% CO₂. From each blood/antigen aliquot, 200 μl of plasma was harvested and stored at -20°C in 96-well plates for subsequent testing. The plasma samples were assayed singly for IFN-γ using Bovigam™ EIA plates according to the manufacturer’s instructions. Absorbance readings were carried out on a MAXline, Vmax® (Molecular Devices Corp., USA) kinetic microplate reader at 450 nm.

Johnin PPD or Avian PPD (Commonwealth Serum Laboratories, Australia) was used at 12.5 μg/ml as a positive control for specific stimulation. PBS was included as a negative control. For all assays, the non-specific T-cell stimulator concavalin A (Sigma, USA) was included for all animals at 20 μg/ml to check cell viability.

Results were expressed as “corrected” absorbance at 450 nm. For duplicate wells, this was defined as the average A₄₅₀ nm of the stimulated wells (Avian or Johnin PPD or P22) minus the average A₄₅₀ nm of the PBS control wells. For single P22 stimulated wells, this was defined as the A₄₅₀ nm of the stimulated well minus the average A₄₅₀ nm of the PBS control wells. Differences between groups were calculated by the Mann-Whitney test. The software package InStat 2.01 (GraphPad Software Incorporated, USA) was used for statistical analysis.

2.9 Rabbit details and immunisations

2.9.1 Preparation of P22 for immunisation

Antigen P22 for immunisation was prepared by transferring approximately 0.05 mg of Ni⁺²-affinity enriched recombinant protein onto nitrocellulose membrane. The membrane was
stained with Ponceau S and the P22 band was excised, destained and air-dried. The blot was then fragmented inside a microfuge tube with the aid of a sterile scalpel blade. To this, approximately 300 µl of PBS was added and the material was further fragmented until it could pass through an 18 gauge, 1 1/2” needle. Five hundred microlitres of Freund’s incomplete adjuvant (F 5506, Sigma, USA) was added and the mixture was passed through the needle several times before being injected.

2.9.2 Rabbit immunisation protocol

Adult New Zealand White rabbits were individually housed under standard conditions in outdoor hutches by Agricultural Services, Massey University, Palmerston North, New Zealand. Injection of prepared antigen was done subcutaneously in the mid-scapular skin fold of one rabbit. This was repeated approximately three weeks later. To check for antibody production, 5 ml of venous blood was collected from the external ear by a scalpel nick three weeks after the second injection. Serum was harvested from the clotted blood after centrifugation at 3,000 x g for 15 min. At the same time, serum was similarly collected from a naïve rabbit for use as control serum. The sera were aliquotted and stored at -20°C. Following confirmation of antibody production, the rabbit was exsanguinated and serum was harvested as above.

2.10 Sheep details and immunisations

2.10.1 Experimental sheep and immunisation protocol

Two mobs of Romney-cross wethers were used in this study. Mob 1 of 11 four month-old sheep was purchased in December 1999 and was housed on pasture with water ad libitum for the duration of the study. At five months of age, six sheep (see Table 2.4) were chosen at random and vaccinated subcutaneously in the right side of the upper neck with Neoparasec (Merial, France) as per the manufacturer’s instructions. The remaining five sheep were kept as unvaccinated controls. An initial blood sample from all animals in the mob, taken prior to vaccination, was tested using the commercial kit Bovigam™ (Commonwealth Serum Laboratories, Australia) to establish their immune status to Johnin PPD antigen. Blood samples were then routinely taken at four weekly intervals.
Mob 2 of 14 four month-old sheep was purchased in December 2000. They were similarly housed and blood samples were routinely taken at four-week intervals. At five months of age, the sheep were randomly divided into various treatment groups as shown in Table 2.5. All animals were tested prior to vaccination with Bovigam™ as above. Vaccination was done subcutaneously in the right side of the upper neck at five months of age with 1 ml of selected vaccine preparation (see Table 2.4). Culture filtrate for immunisations was prepared as described (see 2.6.2.2). Forty milligrams of concentrated culture filtrate was mixed with 1 ml of Neoparasec adjuvant and injected subcutaneously as above.

**Table 2.4. Sheep treatment groups**

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>Sheep #</td>
</tr>
<tr>
<td>Neoparasec</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>131</td>
</tr>
<tr>
<td></td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>136</td>
</tr>
<tr>
<td>unvaccinated</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>130</td>
</tr>
<tr>
<td></td>
<td>133</td>
</tr>
<tr>
<td></td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>137</td>
</tr>
<tr>
<td>unvaccinated</td>
<td></td>
</tr>
</tbody>
</table>

* died April 7, 2001 ryegrass staggers

• died March 7, 2001 ryegrass staggers

2.10.2 Naturally infected sheep

"Limestone Downs" is a large commercial sheep and beef farm located in the south Auckland region of Port Waikato. The farm had not previously been examined for the presence of Johne's disease. Seventy aged ewes (>2 yr old) were culled from the flock due to their general poor condition over the previous six months and serum samples were taken. Upon necropsy, it was found that several factors contributed to the ill-thrift, including facial
eczema, pneumonia, dental problems and parasitic infections. Johne’s disease, although present in the flock, was not considered to be the major cause of the poor condition of these animals. Relevant tissues were examined for Johne’s disease by Dr. Keith Thompson, pathologist, Institute of Veterinary, Animal and Biomedical Sciences, Massey University, Palmerston North, New Zealand. Based on gross pathology, four sheep were diagnosed as having Johne’s disease. Sera from these animals were used in this study along with sera from ten other randomly selected animals. Following histopathology, one of these ten animals was diagnosed as being in the early stages of disease, and another was deemed a Johne’s disease suspect (unconfirmed). The results of these examinations and details of the diagnostic criteria are presented in Appendix 2.

2.11 Naturally infected cattle

Serum samples from nineteen cows from four naturally infected dairy herds from the Manawatu region were used in this study. The infection status of the individual animals was confirmed by cultivation of faeces and subsequent confirmation of *M. a. paratuberculosis* by IS900 PCR. ELISA tests (Commonwealth Serum Laboratories, Victoria, Australia) for serum antibody were also done. Both these tests and interpretation of the results were done by Wallaceville Animal Research Center, Upper Hutt, New Zealand. In addition, sera from two clinically affected cows from two other herds were used. Both of these cows had typical clinical signs of Johne’s disease. Animal 27 had acid-fast organisms in the faeces and animal 25 was confirmed at slaughter with gross lesions typical of Johne’s disease.

2.12 Bioinformatics

DNA and protein databases were searched to identify similarities between cloned *M. a. paratuberculosis* sequences and those existing on the databases. Expected values less than 0.001 were generally considered to be significant. Databases accessed through the National Centre for Biotechnology Information (NCBI) BLAST server (http://www.ncbi.nlm.nih.gov/BLAST/) and European Bioinformatics Institute (EBI) Fasta3 server (http://www2.ebi.ac.uk/fasta3/) were used for completed genomes. Preliminary sequence data for *M. a. avium* and *M. a. paratuberculosis* was obtained from The Institute for Genomic Research (TIGR) website (http://www.tigr.org). Analyses of *M. a. paratuberculosis* signal sequences were done with SignalP World Wide Web Prediction Server (http://www.cbs.dtu.dk/services/SignalP), iPSORT World Wide Web Server...
The programme Neural Network Promoter Prediction (Reese et al., 1996) was used to search for promoter regions in the cloned DNA sequences. DAS – Transmembrane Prediction Server was used to search for the presence of predicted transmembrane regions in translated sequences. Protein and nucleic acid alignments were done using MacVector™ version 6.0.1 software, Oxford Molecular Group.
Chapter 3 Identification of *M. a. paratuberculosis* DNA sequences encoding exported proteins

3.1 Abstract

To identify potential exported proteins of *Mycobacterium avium* subspecies *paratuberculosis*, a library of alkaline phosphatase (*phoA*) gene fusions was constructed in the vector pJEM11 and expressed in *Escherichia coli* and *Mycobacterium smegmatis*. Clones that expressed *M. a. paratuberculosis* gene sequences containing signals for transport fused to *phoA* in the correct frame for translation were identified by culturing on solid media containing the PhoA chromogenic substrate 5-bromo-4-chloro-3-indoyl phosphate. A portion of the *M. a. paratuberculosis* DNA inserts were sequenced from the resulting PhoA⁺, blue colonies. The partial gene sequences adjacent to *phoA* were used to search DNA and protein databases. Twenty-five of the 33 sequences obtained had similarities to other mycobacteria or *Streptomyces* genes and included identities with a copper/zinc superoxide dismutase, a cutinase, a penicillin-binding protein, a serine/threonine protein kinase and three lipoproteins. One sequence was identical to the previously characterised gene encoding the *M. a. paratuberculosis* 34 kDa protein. Seven sequences had no significant similarities to any of the sequences on the databases.

N-terminal regions were determined for 21 of the 33 translated open reading frames fused to *phoA*. These sequences were further analysed for the presence of conserved signal peptide elements and transmembrane-spanning regions. Twelve had evidence of signal sequences and four had predicted transmembrane segments. The remaining five had no evidence of characteristic hydrophobic stretches of amino acids that might function as transmembrane segments. Of the twelve sequences with indefinable N-termini, eight had predicted transmembrane-spanning regions.
3.2 Introduction

Proteins exported by pathogenic organisms including mycobacteria, have been the focus of research for vaccine development and improved diagnostics. In addition, these proteins represent potential targets for drug intervention since they are usually more accessible to drugs than are intracytoplasmic proteins. The rationale for targeting exported proteins has been based on the hypothesis that these are the proteins to which the cells of the host organism are first exposed. Consequently, these proteins may influence the course of the immune response, giving rise to sensitised T lymphocytes that mediate host resistance, leading to initial control and containment of infection. Indeed, a number of studies (reviewed by Anderson, 1997) indicate that the emerging host response to *M. tuberculosis* infection is restricted to relatively few antigenic targets found largely in the exported fraction of the protein repertoire. Furthermore, in the case of tuberculosis, the requirement for the viable bacterium for the generation of specific and long-lasting immunity (Mackaness, 1967; Collins, 1984; Orme, 1988) also supports the idea that actively exported proteins are the key protective elements.

The strong immunogenicity of culture filtrate proteins of *M. tuberculosis* has been demonstrated in numerous laboratories (Rook *et al.*, 1986; Abou-Zeid *et al.*, 1988b; Collins *et al.*, 1988; Andersen *et al.*, 1991a; 1991b; Havlir *et al.*, 1991; Huygen *et al.*, 1992; Orme *et al.*, 1992). In particular, low-mass proteins (4 kDa – 35 kDa) constitute the major targets of T lymphocytes (Andersen *et al.*, 1992). This has also been recorded for culture filtrate proteins of *M. a. avium* (Pais *et al.*, 2000). Consequently culture filtrate proteins have been actively sought after as candidates for new subunit and DNA vaccines. A common experimental approach used to identify exported proteins of mycobacteria has been direct analysis of the proteins present in culture filtrates derived from cultures grown *in vitro*. Several antigens from *M. tuberculosis*, *M. bovis* and *M. a. avium* have been isolated from culture filtrates using chromatographic separation or electrophoresis often combined with immunological screening using immune sera or by lymphocyte stimulation assays such as blastogenesis or IFN-γ production (Fukui & Yoneda, 1961; Nagai *et al.*, 1981; Harboe *et al.*, 1986; Abou-Zeid *et al.*, 1988b; Andersen *et al.*, 1991b; Fifis *et al.*, 1991; Harboe *et al.*, 1992; Boesen *et al.*, 1995; Gulle *et al.*, 1995; Laqueyrerie *et al.*, 1995; Florio *et al.*, 1997; Sinha *et al.*, 1997; Coler *et al.*, 1998; Gobin *et al.*, 1999; Rowland *et al.*, 1999; Weldingh & Andersen, 1999; Olsen *et al.*, 2000b; Pais *et al.*, 2000).
Figure 3.1. Schematic of the *E. coli*/Mycobacterium shuttle vector pJEM11. Important features include a kanamycin resistance marker (kan'); origins of replication for mycobacteria (ori Myco) and *E. coli* (ori *E. coli*); transcriptional terminator (Terminator); truncated gene of *E. coli* alkaline phosphatase (*phoA*) devoid of a promoter, ribosome binding site, start codon and signal sequence region; multiple cloning site for insertion of DNA fragments. Adapted from Lim et al. (1995).
Another approach has been to use genetic information to search existing databases for proteins similar to known exported proteins or segments containing export signal peptides. Similar genomic approaches have been used to identify vaccine candidates for *Staphylococcus aureus* (Williams *et al.*, 2000), *Streptococcus pneumonia* (Wisemann *et al.*, 2001) and *N. meningitidis* (Pizza *et al.*, 2001). Gene fusion methodology based on PhoA, has been successfully used to identify exported proteins in a number of microorganisms. Mycobacterial exported proteins identified using this technology include: ERP (Berthet *et al.*, 1995; Lim *et al.*, 1995), the 19 kDa protein (Lim *et al.*, 1995), DES (Lim *et al.*, 1995; Jackson *et al.*, 1997), Pel (Borich, 1997), SodC (Dupont & Murray, 2001), and several other proteins (Mdluli *et al.*, 1995; Moran *et al.*, 1999; Gomez *et al.*, 2000). PhoA fusions have also been used to identify exported proteins of *Streptococcus pneumoniae* (Pearce *et al.*, 1993), *E. coli* (Boquet *et al.*, 1987), *Helicobacter pylori* (Johnson *et al.*, 1995; Bina *et al.*, 1997; Oliaro *et al.*, 2000), *Vibrio cholerae* (Taylor *et al.*, 1989), *Bordetella pertussis* (Knapp & Mekalanos, 1988), *Staphylococcus aureus* (Williams *et al.*, 2000) and *Actinobacillus actinomycetemcomitans* (Ward *et al.*, 2001).

The activity of *E. coli* PhoA generally requires it to be exported across the plasma membrane, where it can form intramolecular disulfide bonds (Derman & Beckwith, 1995) and undergo subsequent dimerisation to form the active enzyme (Kim & Wyckoff, 1989). Successful export of the enzyme and thus activity, depends on the presence of a recognised signal sequence, which directs PhoA transport across the plasma membrane. The vector pJEM11, shown in Figure 3.1, allows fusions between phoA and genes encoding exported proteins. In this vector, phoA exists devoid of a promoter region, ribosome binding site, start codon and signal sequence coding region (Lim *et al.*, 1995). Thus, PhoA activity becomes dependent upon translation of a fused DNA segment, in the correct frame, which encodes the necessary elements for export. The enzymatic activity of the PhoA can then be readily detected with the substrate BCIP. This results in the appearance of blue colonies when plated on solid medium containing the substrate.

This chapter describes the construction of a library of *M. a. paratuberculosis* phoA fusions in the vector pJEM11 and identification of putative exported proteins of *M. a. paratuberculosis*. 
3.3 Materials and methods

3.3.1 Construction of an *M. a. paratuberculosis* pJEM11 expression library

3.3.1.1 Extraction of DNA from *M. a. paratuberculosis*

Approximately 3 mg of lyophilised *M. a. paratuberculosis* (New Zealand field isolate ATCC 53950) was resuspended in 0.6 ml of extraction buffer (100 mM NaCl, 25 mM EDTA (pH 8.0), 10 mM Tris.Cl pH 8.0, 0.5% (w/v) SDS) before adding 200 µg of proteinase K (Roche Molecular Biochemicals, Germany). The mixture was incubated at 50°C for 18 h and then 100 µl of 5 M NaCl and 120 µl of 6.7% (w/v) cetyltrimethylammonium bromide (Aldrich Chemical Company, USA) in 0.5% (w/v) NaCl was added. The digest was mixed with an equal volume of 25:24:1 phenol:chloroform:isoamyl alcohol and centrifuged at 15,800 × g for 5 min. The aqueous phase was collected and the phenol:chloroform:isoamyl alcohol extraction was repeated, as above. The aqueous phase was mixed with an equal volume of chloroform and centrifuged again. The DNA was precipitated with the addition of 1 volume of 100% isopropanol to the collected aqueous phase. After 18 h at -20°C, the DNA was pelleted by centrifugation at 15,800 × g for 30 min at 4°C. The pellet was washed with 1 ml of 70% ethanol, air dried at room temperature and resuspended in 100 µl of TE buffer (Appendix 1) containing 100 µg/ml RNAsase A (Life Technologies Inc., USA) and incubated for 18 h at 37°C. The DNA concentration was calculated based on absorbance at 260 nm.

To check the condition of the DNA, an 8 µl sample (~11 µg) was electrophoresed on a 0.7% agarose gel, stained with ethidium bromide and visualised under UV light (see 2.4.2).

To confirm the identity of the isolated DNA, PCR was done on 1 µl samples of a 1:10 dilution of the DNA in a final volume of 50 µl using Expand High Fidelity PCR system (Roche Molecular Biochemicals, Germany) as described in section 2.4.5.2. For confirmation of the genus *Mycobacterium*, the primers 246 and 264 were used to amplify the 16S rRNA gene (see Table 3.1). To confirm the subspecies *M. a. paratuberculosis*, the genetic elements F57 (Poupart et al., 1993) and IS900 (Green et al., 1989) were amplified using the primers F57-f, F57-r (Poupart et al., 1993) and 90, 91 (Sanderson et al., 1992), respectively (see Table 3.1). PCR products were electrophoresed on agarose gels and photographed as described in section 2.4.2.
3.3.1.2 Preparation of pJEM11 vector DNA

pJEM11 plasmid DNA was purified from transformed *E. coli* DH10B cells. The plasmid DNA was quantitated based on its absorbance at 260 nm. Eleven micrograms of plasmid was digested to completion with 2.5 units of *Bam* HI at 37°C for 2 h and was extracted from agarose gels as described (see 2.4.3). The resulting digested plasmid DNA was dephosphorylated using 2 units of alkaline phosphatase (Boehringer Mannheim, Germany) for 1 h at 37°C to prevent recircularisation of the plasmid. The DNA was then purified by agarose gel extraction. To estimate the success of the alkaline phosphatase treatment, 0.5 units of T4 DNA ligase (Roche Molecular Biochemicals, Germany) was added to 0.95 µg of prepared vector in a total volume of 10 µl using the supplied buffer. After incubation for 30 min at room temperature, the mixture was dialysed and electroporated into 100 µl of *E. coli* XL-1 cells. Fifty microlitres of the transformation mix was spread on an LB/kan plate and incubated at 37°C for 18 h.

3.3.1.3 Partial digestion with *Sau* 3AI of *M. a. paratuberculosis* DNA

To establish the conditions necessary to obtain a suitable partial digest of the genomic DNA, 2.8 µg was digested using 0.05 units *Sau* 3AI in a volume of 25 µl at 22°C. From this, 6 x 4 µl samples were removed at 30 s intervals and immediately mixed with 5 µl of DNA loading dye (Appendix 1) containing 2 µl of 0.5 M EDTA (pH 8.0) to stop the reaction. These samples were electrophoresed on a 1% agarose gel along with 5 µl (700 ng) of undigested genomic DNA. The gel was ethidium bromide stained and photographed. A partial digest was indicated by the disappearance of the predominating high molecular weight band and the appearance of a gradual smear ranging in size from 10,000 - 100 base pairs. From the 1.5 and 2 min digestion lanes, DNA ranging in size from approximately 200 - 3,000 base pairs was extracted from the agarose gel using a QIAquick Gel Extraction Kit (Qiagen, Germany). The resulting DNA fragments were eluted in a final volume of 50 µl of TE buffer. A 10 µl sample was run on a 1% agarose gel to check recovery.

3.3.1.4 Ligation of *M. a. paratuberculosis* DNA and pJEM11 and transformation into *E. coli*

Approximately 1 µg of *Bam* HI, alkaline phosphatase-treated pJEM11 vector DNA and 0.7 µg of *Sau* 3AI partially digested *M. a. paratuberculosis* DNA were ligated using 1.0 unit of
T4 DNA ligase in a total volume of 35 µl at room temperature for 30 min. The mixture was dialysed and electroporated into 100 µl of E. coli DH10B cells. To this, 500 µl of LB broth was added and incubated for 1 h at 37°C. Aliquots of approximately 70 µl each were plated onto nine LB/kan/BCIP plates and incubated at 37°C for approximately 18 h. A representative plate was selected and colonies were counted on a quarter of the plate to estimate the total number of resulting transformants. The total number of blue transformants was also counted. Blue E. coli colonies were restreaked onto LB/kan/BCIP plates for confirmation of PhoA+ phenotype. Each of these colonies was cultured in LB/kan broth for glycerol storage at -70°C.

3.3.1.5 Plasmid isolation from the E. coli recombinant library

To each of the nine resulting plates, 1.6 ml of LB/kan broth was added and the colonies were resuspended with the aid of a rubber spatula. The mixtures from each plate were transferred to 2 ml microtubes and the plasmid DNA was extracted using a BRESAspin Plasmid Mini Kit (Bresatec, Australia) with the modification that all reagent volumes were doubled due to the high concentration of cells. Each of the nine plasmid preparations was eluted into 50 µl of TE buffer, for a total volume of 450 µl. The plasmids were stored at -20°C until used. To confirm that the plasmid collection contained M. a. paratuberculosis genomic DNA, PCR was carried out on a 1 µl sample using the primers 90, 91 (Table 3.1) for amplification of IS900 in a total volume of 50 µl using Expand High Fidelity PCR system (Roche Molecular Biochemicals, Germany) as described in section 2.4.5.2. The resultant product was agarose gel extracted and sequenced for confirmation using the 90 primer (Table 3.1).

3.3.1.6 Transformation of the recombinant plasmids into M. smegmatis mc2155

From the plasmid mixture extracted from E. coli, 2 x 10 µl aliquots were dialysed and used to transform two, 100 µl aliquots of electrocompetent M. smegmatis cells, prepared as described (see 2.5.3). To each of the resulting transformations, 500 µl of LB broth was added and the cells were incubated at 37°C for 2 h. From this, 80 µl samples were spread onto 15 LB/kan/BCIP plates, which were then incubated at 37°C. After five days, the plates were transferred to 4°C for a further 54 days. To estimate the total number of colonies (blue and white), a quadrant of a representative plate was counted. Over the 59 days, blue colonies were picked daily and designated numerically, as they appeared. These were restreaked onto fresh LB/kan/BCIP plates and grown at 37°C until they turned blue and then were transferred to
Figure 3.2. Schematic representation of the location of the oligonucleotide primers designed for sequencing the *M. a. paratuberculosis* inserts in the pJEM1 constructs. JEM1 binds to the complementary strand of the transcriptional terminator sequence and JEM2 binds to the coding strand of the *phoA* gene. Sequences of the primers are listed in Table 3.1.
glycerol for storage at -70°C.

3.3.2 Sequencing of DNA inserts encoding putative exported proteins

The DNA inserts from the pJEM11 constructs were partially or completely sequenced as described in section 2.4.6. Extracted plasmid DNA was generally used as template. In some cases, PCR products obtained with the primers JEM1 and JEM2, were directly sequenced from the *M. smegmatis* colonies. Primers JEM1 and JEM2 (Table 3.1) were used for sequencing. As shown in Figure 3.2, JEM1 binds to the complementary sequence of the transcriptional terminator, approximately 100 base pairs from the *Bam* HI restriction site of pJEM11. JEM2 binds to the coding sequence of the *phoA* gene, 141 base pairs from the *Bam* HI site of pJEM11.

To obtain plasmid DNA for sequencing of inserts, the pJEM11 constructs were first transferred from individual *M. smegmatis* colonies to *E. coli* DH10B cells. This was done to avoid the lengthy and more difficult extraction of sufficient amounts of plasmid DNA from *M. smegmatis*. Transferred pJEM11 constructs were purified from *E. coli*. The primer JEM2 was used for sequencing of both plasmid and PCR products to obtain the DNA sequence of the cloned insert immediately adjacent to the *phoA* coding region.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Product size (base pairs)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>JEM1*</td>
<td>5’CGAGCTGCAGTGGGATGACC</td>
<td>~240</td>
<td>Pasteur Institute</td>
</tr>
<tr>
<td>JEM2*</td>
<td>5’TCCGCCTGAGCAGCCCCGTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>90*</td>
<td>5’GTCGAGGGCCGTCGATTAG</td>
<td>400</td>
<td>Sanderson <em>et al.</em>, 1992</td>
</tr>
<tr>
<td>91*</td>
<td>5’GAAGTGGATCGCCACGTTGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F57-T*</td>
<td>5’GGATCTGCAGCCCAGATAG</td>
<td>618</td>
<td>Poupart <em>et al.</em>, 1993</td>
</tr>
<tr>
<td>F57-5*</td>
<td>5’ATCTAGACAGCCAGCAGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>246*</td>
<td>5’AGAGTTTATCTAGGTTCGCTAG</td>
<td>1,030</td>
<td>Boddinghaus <em>et al.</em>, 1990</td>
</tr>
<tr>
<td>264*</td>
<td>5’TGCACACAGCCACACAGGGA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*sequencing primer designed for inserts in pJEM11 constructs (see Figure 3.2)
*primer designed to amplify the 5’ region of IS900. Primer 90 binds at position 22-41 and primer 91 binds position 402-421 within the IS900 element
*primer designed to amplify the F57 genetic element of *M. a. paratuberculosis*
*primer designed to amplify the mycobacterial 16S rRNA gene
Figure 3.3. Schematic representation of the construction of the PhoA fusion library. *M. a. paratuberculosis* Sau 3AI fragments were cloned into the *Bam* HI site of the pJEM11 vector. Cloned sequences containing a promoter, ribosome binding site (RBS), start codon (ATG) and signal sequence in-frame with the *phoA* gene are expressed as PhoA fusion proteins in *E. coli* and/or *M. smegmatis*. Export of the PhoA fusion product across the cell membrane results in blue colouration of colonies when grown on media containing the substrate BCIP.
Figure 3.4. PCR amplification of *M. a. paratuberculosis* genetic elements. The species-specific elements IS900 (lane 1, 400 base pairs) and F57 (lane 2, 618 base pairs) and the mycobacterial genus-specific 16S rRNA gene (lane 3, 1,030 base pairs) from *M. a. paratuberculosis* extracted genomic DNA are shown. Primer pairs 90, 91 and F57-f, F57-r and 246, 264 (detailed in Table 3.1), respectively, were used. PCR products were electrophoresed on a 1% agarose gel, stained with ethidium bromide and photographed under UV light. Lane M contains a DNA size standard (base pairs) for comparison.
Figure 3.5. PhoA$^+$ recombinant *E. coli* colonies from the *M. a. paratuberculosis* pJEM11 library. Colonies producing exported PhoA fusion proteins turn blue in the presence of the substrate BCIP. (a) PhoA$^+$ recombinant *E. coli* colonies on a plate containing BCIP from the original library. (b) Selected *E. coli* colonies from the library: 1, Eco-3; 2, Eco-4; 3, Eco-20; 4, Eco-22; 5, Eco-24; 6, Eco-16; 7, pJEM11 no insert; 8, 9, 10, pJEM11 “white” (PhoA$^-$) recombinant colonies.
3.4 Results

3.4.1 Construction of an *M. a. paratuberculosis* pJEM11 expression library

In order to identify *M. a. paratuberculosis* gene sequences encoding exported proteins, a library of *M. a. paratuberculosis* phoA fusions was created in the vector pJEM11 (see Figure 3.1) and expressed in *E. coli* and *M. smegmatis*. The construction of the *M. a. paratuberculosis* library and subsequent detection of exported fusion proteins is represented in Figure 3.3.

3.4.1.1 Confirmation of *M. a. paratuberculosis* DNA for cloning

To confirm the identity of the DNA extracted from *M. a. paratuberculosis*, PCR of the mycobacterial 16S rRNA gene (Boddinghaus *et al.*, 1990) and the species-specific genetic elements F57 (Poupart *et al.*, 1993) and IS900 (Green *et al.*, 1989) was carried out as described (see 3.3.1.1). The results are presented in Figure 3.4. Identities of both the IS900 and F57 PCR products were confirmed by sequencing using primers 90 and F57-f, respectively, as described in section 2.4.6.

3.4.1.2 Subcloning of *M. a. paratuberculosis* DNA into the vector pJEM11 and expression of the library in *E. coli*

*M. a. paratuberculosis* DNA was prepared and subcloned into pJEM11 as described in sections 3.3.1.3 and 3.3.1.4. Since *M. smegmatis* mc^2^155 has a 10^4^ lower rate of transformation efficiency as compared to *E. coli* (Snapper *et al.*, 1990), the plasmid library was first transformed into *E. coli* to ensure the highest possible proportion of constructs could be recovered.

After 18 h at 37°C a representative *E. coli* plate had 1,200 colonies. This resulted in an estimated total of 10,800 colonies from nine plates. Of these, 17 colonies were blue. These were designated Eco-1 to Eco-17. Upon a further 18 h storage of the library at 4°C, seven more blue *E. coli* colonies resulted. These were designated Eco-18 to Eco-24. Thus, a total of 24 blue *E. coli* colonies, representing 0.2% of the library were obtained. Figure 3.5 shows colonies from the original library and some subcultured clones. After this time point, all the *E. coli* colonies on the plates appeared pale blue, making selection of PhoA^+^ phenotype...
Figure 3.6. PhoA+ recombinant *M. smegmatis* colonies from the *M. a. paratuberculosis* pJEM11 library. Colonies producing exported PhoA fusion proteins turn blue in the presence of the substrate BCIP. (a) PhoA+ recombinant *M. smegmatis* colonies on a plate containing BCIP from the original library. (b) Selected *M. smegmatis* colonies from the library: 1, pTB-17; 2, pTB-16; 3, pTB-310; 4, pTB-38; 5, pTB-184; 6, pTB-61; 7, pJEM11 no insert; 8, pJEM11 “white” (PhoA+) recombinant colony.
beyond this time point difficult. For this reason, no more colonies were collected.

Using the formula of Clarke and Carbon (1976), (see below) and based on an average insert size of 1,700 base pairs (see 3.4.1.3), it was calculated that 12,729 clones would represent a complete library. Since 10,800 clones were present in the original *E. coli* library, this represents 84.8% of the *M. a. paratuberculosis* genome.

\[
N = \frac{\ln (1 - P)}{\ln (1 - (I/G))}
\]

where  
\[ P = 99\% \text{ probability (0.99)} \]  
\[ N = \text{number of clones necessary} \]  
\[ G = \text{size of target genome (base pairs) (4.7 \times 10^6 \text{ base pairs, (Cocito et al., 1994)})} \]  
\[ I = \text{average size of cloned fragments (1,700 base pairs)} \]

To confirm that the library contained *M. a. paratuberculosis* DNA, plasmid was extracted from the pooled *E. coli* colonies and PCR for IS900 was carried out. The resulting 400 base pair product was sequenced and was confirmed to be IS900.

### 3.4.1.3 Expression of the library in *M. smegmatis*

A sample of the plasmid collection was used to transform *M. smegmatis*. After three days incubation at 37°C on LB/kan/BCIP plates, the first blue *M. smegmatis* colony appeared and was designated pTB-1. Over the next two days, a further 45 blue colonies appeared, designated pTB-2 to pTB-46. The plates were transferred to 4°C and 473 more blue colonies appeared over 54 days. A total of 519 blue *M. smegmatis* clones were collected. From day two to day 20, the average number of blue colonies that appeared per day was 19, whereas, from day 21 to day 56 the average was four. During this time, the remaining *M. smegmatis* colonies were a uniform yellow-white in colour. PhoA+ colonies with a varying intensity of blue, were therefore easily distinguished (see Figure 3.6). A quadrant count of a representative plate from day five gave approximately 1,000 colonies. The total number of *M. smegmatis* colonies was estimated to be 60,000, of which 0.9% were blue.

A six-fold increase in the number of colonies was obtained in *M. smegmatis* (60,000) compared to *E. coli* (10,800). This was a result of amplification of the plasmids in *E. coli* prior to transfer to *M. smegmatis*. Multiple copies of the plasmids were therefore expected to be present in the *M. smegmatis* library. This expectation was confirmed by analysis of the
Figure 3.7. Restriction endonuclease analysis of selected PhoA+ clones. Digested plasmids were electrophoresed on 2% agarose gels along with DNA size standards (lane M). The gels were stained with ethidium bromide and photographed under UV light. (a) Digestion of plasmids with Kpn I and Apa I which flank the Bam HI site into which the inserts were cloned. pJEM11 vector DNA can be seen at the 12,000 base pair marker. Insert sizes were deduced from comparison with standard size markers. Lanes 1 to 9 contain plasmid from M. smegmatis clones pTB-30, pTB-20, pTB-1, pTB-56, pTB-59, pTB-107, pTB-112, pTB-114, pTB-164, respectively. Lanes 10 to 13 contain plasmid from E. coli clones Eco-4, Eco-20, Eco-22, Eco-24, respectively. Plasmids possessing unique banding patterns, for example pTB-30 and pTB-20 (lanes 1 and 2) were sequenced. Plasmids possessing identical digestion patterns, such as pTB-1 and pTB-56 (lanes 3 and 4) were further analysed by digestion with Hinf I. (b) Digestion of plasmids with Hinf I. Hinf I was used to discriminate between plasmids which appeared similar upon digestion with Kpn I and Apa I. Lanes 1 to 5 and 7 to 10 contain plasmid from M. smegmatis clones pTB-141, pTB-107, pTB-144, pTB-59, pTB-164, pTB-1, pTB-56, pTB-74, pTB-38, respectively. Lane 6 contains plasmid from the E. coli clone Eco-24. As an example, pTB-1 and pTB-56 (lanes 7 and 8) appear to contain identical DNA inserts.
pJEM11 plasmid constructs (see Table 3.2). The average insert size was calculated from PhoA+ constructs analysed by agarose gel electrophoresis. The plasmids were digested with Apa I and Kpn I, whose recognition sites flank the Bam HI site in the vector used for cloning (see Figure 3.1). Only unique inserts were included in calculations. From the 128 plasmids analysed, the average insert size was 1,700 base pairs. The inserts generally ranged in size from 300 - 3,000 base pairs with a median of 1,400 base pairs.

3.4.2 Analysis of M. a. paratuberculosis phoA fusions

3.4.2.1 Sequencing of DNA inserts encoding putative exported proteins

In an attempt to identify the genes encoding potential exported proteins of M. a. paratuberculosis, sequencing of insert DNA adjacent to phoA was carried out on selected PhoA+ clones. Plasmid DNA was used for sequencing reactions because direct PCR amplification of the inserts proved to be difficult. From initial sequencing of some of the blue E. coli and M. smegmatis clones, it was apparent that many contained the same inserts. Thus, to reduce sequencing redundancies, the clones were first screened for presence of unique inserts. This was done by restriction endonuclease digestion of the plasmids. Digestion with Kpn I and Apa I, which flank the DNA inserts in pJEM11 (see Figure 3.1), produced DNA fragments of various sizes when electrophoresed on 2% agarose gels (see Figure 3.7 a). Plasmids that gave unique banding patterns were sequenced. Those that produced similar patterns with Kpn I and Apa I were further digested with Hinf I (see Figure 3.7 b). Hinf I is a relatively frequent cutter of DNA (G^↓ ANT) and gave further information as to the similarity of these inserts. Of the 131 clones analysed by digestion of plasmid DNA, three did not contain insert. In summary, insert DNA was fully or partially sequenced from 63 individual clones. Of these, 33 were unique in sequence and the rest were redundant.

3.4.2.2 Analysis of phoA fusions

To identify the open reading frames (ORFs) responsible for expression of the PhoA fusion products, DNA and protein databases were searched using DNA sequences and/or deduced amino acid sequences of the inserts. In the first instance, all DNA sequences were submitted to the TIGR M. a. paratuberculosis database. In all cases, alignments with extensive identity (>90%) were obtained (data not shown), confirming the sequences used for database searches were contiguous and not the result of Sau 3AI fragment ligations. All but one of the sequences (pTB-209) were also present in the TIGR M. a. avium database.
Where possible, the N-terminal amino acids methionine or valine, were predicted using a variety of information. Alignment with the most similar protein from database searches was the main method used for assignment. The presence of in-frame stop codons, such as TAG, in the DNA sequence, the presence of adjacent ORFs, predicted ribosome binding sites and potential promoter sequences were also taken into consideration. A summary of the results of the protein database searches is presented in Table 3.2. The most significant alignments and predicted signal peptide cleavage sites, where present, are included in Appendix 3. The N-terminal regions predicted for 21 of the 33 translated PhoA fusions analysed are described in Table 3.2 a and b. Analyses of the in-frame predicted N-terminal amino acid sequence using SignalP or iPSORT, revealed that 11 of these had recognisable signal peptide cleavage sites (see Table 3.2 a). Preliminary investigation of the recombinant 34 kDa protein, partially encoded by pTB-41, showed two bands in Western blot analyses using a C-terminal monoclonal antibody (not shown), and thus this protein appeared to have a cleavable N-terminal signal sequence when expressed in *M. smegmatis*. Therefore this clone was also included in this group. In all, 12 of the 21 inserts for which translated N-termini were located had evidence of cleavable signal peptides. Of the remaining nine sequences in this group of 21, four had transmembrane segments predicted with the programme "DAS" (see 2.12). N-termini were not determined for 12 sequences (see Table 3.2 c), either because sequencing results did not extend far enough to reach these regions or there was insufficient database information available to provide clues as to where these might be located. Of these 12 translated sequences, eight possessed predicted transmembrane segments. Seven sequences (pTB-13, pTB-26, pTB-73, Eco-4, pTB-20, pTB-209 and pTB-281) showed no significant identity to any of the annotated databases available. Signal sequences were predicted by SignalP in three of these (pTB-13, pTB-26 and pTB-73) using methionine residues as potential N-terminal amino acids for the entered peptide sequences.
### Table 3.2. Identification of selected *M. a. paratuberculosis* PhoA fusion proteins

<table>
<thead>
<tr>
<th>Clone</th>
<th>Insert size (base pairs)</th>
<th>Most significant database search result</th>
<th>%id</th>
<th>%sim</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTB-1</td>
<td>2,800</td>
<td>Putative membrane protein ML0081 <em>M. leprae</em></td>
<td>61</td>
<td>72</td>
<td>Similar to Rv3853</td>
</tr>
<tr>
<td>pTB-4</td>
<td>1,200</td>
<td>34 kDa <em>M. avium</em> subsp. <em>paratuberculosis</em></td>
<td>100</td>
<td></td>
<td>(De Kesel <em>et al.</em>, 1993; Gilot <em>et al.</em>, 1993). Surface exposed, 72% nucleic acid identity to <em>M. leprae</em> (<em>Silbaq et al.</em>, 1998).</td>
</tr>
<tr>
<td>pTB-13</td>
<td>800</td>
<td>N/A</td>
<td></td>
<td></td>
<td>Similarity to cytochrome P450 (Rv2266) exists in downstream sequence.</td>
</tr>
<tr>
<td>pTB-14</td>
<td>520</td>
<td>Serine esterase cutinase <em>M. a. avium</em></td>
<td>100</td>
<td></td>
<td>Similar to Rv1984c (68% amino acid identity), (<em>Weldingh and Andersen</em>, 1999).</td>
</tr>
<tr>
<td>pTB-26</td>
<td>2,580</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eco-24</td>
<td>1,200</td>
<td>Hypothetical protein 39.8 kDa</td>
<td>68</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>-------------------------------</td>
<td>----</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td>pTB-88</td>
<td></td>
<td>Rv2190c</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pTB-144</td>
<td></td>
<td><em>M. tuberculosis</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pTB-164</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pTB-106</td>
<td>750</td>
<td>putative D-alanyl-D-alanine carboxypeptidase</td>
<td>75</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>pTB-118</td>
<td></td>
<td><em>M. leprae</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pTB-119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pTB-143</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pTB-177</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTB-162</td>
<td>1,450</td>
<td>Probable membrane protein (Rv2345)</td>
<td>47</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>pTB-73</td>
<td>1,350</td>
<td>N/A</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Weakly similar to *Listeria* invasion-associated protein (predicted secreted). Signal sequence predicted, but could not locate exact M or V start.**

**Contains conserved peptidase S11 domain. Has hydrophobic stretches at both N and C termini. Also similar to *M. tuberculosis* hypothetical penicillin-binding protein Rv3330.**

**Hydrophobic stretch at N-terminus and around position 180.**

Contains similarity to probable membrane protein *Streptomyces* running in opposite direction.

### Signal sequence not predicted for N-terminus

<table>
<thead>
<tr>
<th>pTB-166</th>
<th>300</th>
<th>Hypothetical membrane-associated 18.9 kDa lipoprotein (LppP)</th>
<th>62</th>
<th>75</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><em>M. tuberculosis</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**18.9 kDa lipoprotein has predicted signal peptide. Lipobox is present in N-terminus of cloned sequence. Predicted transmembrane segment present amino acids 5-13.**

<table>
<thead>
<tr>
<th>pTB-187</th>
<th>1,050</th>
<th>Probable lytB</th>
<th>83</th>
<th>89</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><em>M. leprae</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Similar to *E. coli* lytB involved in penicillin tolerance and control of stringent response. Upstream sequence has similarity to Rv1109c running in opposite direction.**

<table>
<thead>
<tr>
<th>pTB-191</th>
<th>800</th>
<th>Probable 50S ribosomal protein L11 (gene <em>rp1K</em>)</th>
<th>94</th>
<th>95</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><em>M. leprae</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>pTB-201</th>
<th>1,950</th>
<th>Probable phosphoenolpyruvate carboxykinase (PckA)</th>
<th>83</th>
<th>89</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTB-249</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pTB-310</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>M. tuberculosis</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Predicted transmembrane segment present amino acids 126-134.**

<table>
<thead>
<tr>
<th>pTB-235</th>
<th>1,200</th>
<th>Hypothetical protein (Rv2298)</th>
<th>38</th>
<th>47</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><em>M. tuberculosis</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Of the aldoketo reductase 2 family.**

<table>
<thead>
<tr>
<th>pTB-413</th>
<th>3,300</th>
<th>Probable 6-phosphogluconate dehydrogenase <em>M. leprae</em></th>
<th>87</th>
<th>88</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Pentose phosphate pathway. Predicted transmembrane segment present amino acids 18-25.**

<table>
<thead>
<tr>
<th>pTB-494</th>
<th>900</th>
<th>Hypothetical transcriptional regulator (Rv0452)</th>
<th>76</th>
<th>82</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><em>M. tuberculosis</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Contains tetR family of bacterial regulatory proteins conserved domain.**

<table>
<thead>
<tr>
<th>pTB-433</th>
<th>1,800</th>
<th>Hypothetical protein Rv1473</th>
<th>87</th>
<th>92</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><em>M. tuberculosis</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Conserved ABC transporter signature, ATP-binding protein.**
<table>
<thead>
<tr>
<th>Clone</th>
<th>Percentage</th>
<th>Predicted Membrane Segment Present</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTB-519</td>
<td>650</td>
<td>Contains 4Fe-4S iron-sulfur cluster domain. Predicted transmembrane segment present amino acids 36-45.</td>
<td></td>
</tr>
<tr>
<td>Eco-4</td>
<td>1,410</td>
<td>N/A</td>
<td>Upstream sequence has similarity to Rv3242c running in opposite direction.</td>
</tr>
<tr>
<td>Eco-20</td>
<td>3,400</td>
<td>N/A</td>
<td>Upstream sequence has similarity to Rv0191 running in opposite direction.</td>
</tr>
<tr>
<td>Eco-22</td>
<td>1,410</td>
<td>N/A</td>
<td>Predicted transmembrane segment present.</td>
</tr>
<tr>
<td>pTB-20</td>
<td>1,900</td>
<td>N/A</td>
<td>TCA cycle. Predicted transmembrane segment present.</td>
</tr>
<tr>
<td>pTB-57</td>
<td>1,550</td>
<td>Citrate synthase (Rv0896) M. tuberculosis</td>
<td>Two predicted transmembrane segments present.</td>
</tr>
<tr>
<td>pTB-74</td>
<td>2,300</td>
<td>Hypothetical methyltransferase (Rv0560c) M. tuberculosis</td>
<td>No sequence identity found with M. a. avium database</td>
</tr>
<tr>
<td>pTB-107</td>
<td>2,000</td>
<td>3-methyl-2-oxobutonate hydroxymethyl transferase (PanB) M. tuberculosis</td>
<td></td>
</tr>
<tr>
<td>pTB-141</td>
<td>2,000</td>
<td>N/A</td>
<td>Upstream sequence has similarity to katE M. avium in different reading frame.</td>
</tr>
<tr>
<td>pTB-179</td>
<td>3,100</td>
<td>Probable 1-deoxyxyllose-5-phosphate synthase M. tuberculosis</td>
<td>In the biosynthetic pathway to isoprenoids, thiamin and pyridoxol. Predicted transmembrane segment present.</td>
</tr>
<tr>
<td>pTB-184</td>
<td>1,200</td>
<td>Glucose-6-phosphate dehydrogenase M. avium</td>
<td>Pentose phosphate pathway. Predicted transmembrane segment present.</td>
</tr>
<tr>
<td>pTB-190</td>
<td>1,400</td>
<td>Hypothetical transcriptional regulator Streptomyces coelicolor</td>
<td>tetR family of bacterial regulatory proteins. Predicted transmembrane segment present.</td>
</tr>
<tr>
<td>pTB-281</td>
<td>3,100</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>pTB-371</td>
<td>650</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>pTB-437</td>
<td>1,200</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>pTB-458</td>
<td>1,400</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>pTB-470</td>
<td>1,700</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

% id refers to the percentage of identical amino acids possessed between the M. a. paratuberculosis sequence and the best protein match from the database. % sim refers to % id plus similar amino acids. Transmembrane segments were predicted with the programme "DAS". (a) lists clones for which an N-terminal region has been determined for the PhoA fusion product and a signal sequence has been predicted by SignalP or iPSORT and/or for
which experimental evidence suggests a signal sequence exists. (b) lists clones for which an N-terminus has been determined for the PhoA fusion product, but a signal sequence was not predicted. (c) lists clones for which the N-terminus of the PhoA fusion has not been determined. Names of proteins are designated as by NCBI database annotations. Proteins with the prefix Rv, refer to hypothetical proteins derived from ORFs of the *M. tuberculosis* H₃₇Rv complete genome (Cole et al., 1998). N/A = not applicable (no significant similarity found). In most cases, the expected value $E<0.001$ was used to determine significance of the match. The database alignments are shown in Appendix 3.
3.5 Discussion

The construction and partial characterisation of an *M. a. paratuberculosis* phoA fusion library was carried out for the purpose of identifying novel exported proteins from this pathogen. The library was expressed in both *E. coli* and *M. smegmatis*. Approximately 100 unique PhoA+ clones, representing 0.9% of the plasmid library, were obtained. Thirty-three clones containing gene sequences fused to phoA were selected for further analyses.

Seven of the 33 sequences analysed by the BLAST algorithm were not similar to any of the annotated nucleic acid or protein databases. Six of these sequences matched sequences in the incomplete *M. avium* database with greater than 90% identity. It is not known if these sequences are specific to the *M. avium* complex. A single sequence that did not match any sequence in the *M. a. avium* database may be specific to *M. a. paratuberculosis*. Further genetic analyses are required to investigate these possibilities.

Signal sequences were predicted in 12 of the 21 sequences for which N-terminal regions could be identified (Table 3.2a). Potential transmembrane segments were identified in four others (Table 3.2c) and in a further 10 for which the N-terminal regions were not located (Table 3.2c). Transmembrane segments can allow the insertion of proteins into the cell membrane. In doing so, they function like signal sequences, however they are not removed and serve to anchor the proteins to the membrane. For fusion proteins, this can lead to sequestration of PhoA outside the cell membrane resulting in phosphatase activity. This is the principle upon which PhoA fusions have been used to study membrane protein topology (Manoil & Beckwith, 1986). Although pTB-166 was included in the group with those proteins for which an N-terminal signal sequence was not predicted, a lipidation consensus was recognised by both SignalP and iPSORT, and the resulting predictions for a signal sequence were reported as inconclusive by these programmes. The partial protein sequence had identity to a predicted lipoprotein of *M. tuberculosis* for which a signal sequence was recognised. In view of this, it is possible that the protein encoded by pTB-166 was exported via recognition of a signal peptide.

There were nine sequences for which signal sequences or transmembrane segments were not identified. In four of these (Eco-4, pTB-20, pTB-209 and pTB-281), it was not possible to locate the N-terminal regions with any certainty as these sequences showed very little similarity to any of the sequences in the annotated databases. Various combinations of ORF
sequences beginning with methionine or valine were analysed but no transmembrane segments or signal sequences were reliably predicted for these. N-terminal regions were located for the remaining five sequences (pTB-187, pTB-191, pTB-235, pTB-494 and pTB433). The mechanism by which they may have been transported is not known. Export systems, such as type III secretion systems, which do not require signal sequences, have been described for the export of virulence factors in several Gram-negative species (reviewed by Cheng & Schneewind, 2000). To date, there is no evidence to suggest such systems exist in mycobacteria (Tullius et al., 2001). It is therefore unlikely that these fusion proteins were exported by some unidentified mechanism. Thus, it appears that these proteins may be of cytoplasmic origin. There are several explanations for the apparent export of these fusion proteins. Cytoplasmic PhoA has been shown to fold into its native, active conformation in cells that are no longer growing (Derman & Beckwith, 1995). This has been observed in E. coli and although the phenomenon has not been investigated in M. smegmatis, it is possible that the fusion to phoA of gene sequences lacking export signals, in the correct frame for expression, can result in PhoA activity in cells whose growth has ceased. An observation supporting this possibility is that all of the proteins for which N-terminal signal sequences were identified, were isolated from the M. smegmatis library within the first 20 days of the 54 day incubation. Furthermore, several predicted “housekeeping” genes coding for proteins such as glucose-6-phosphate dehydrogenase and citrate synthase, which would be predicted to be cytoplasmic, were isolated late in the incubation of the M. smegmatis library. Another explanation relates to the reported observation of some cytoplasmic proteins to appear in the external media as a result of leakage in apparently viable, healthy bacteria (Tullius et al., 2001). It is postulated that as cells divide and remodel, more highly expressed proteins accumulate extracellularly. Levels of expression of the fusion proteins was not investigated in this study and so it is not known if this accounts for their appearance.

The finding of “housekeeping” genes may also be explained by the possible presence of internal sequences, such as transmembrane segments, able to mimic export signals or promoters in E. coli and/or M. smegmatis. Fortuitous sequences that are recognised by E. coli RNA polymerase σ factor have been reported in SV40 DNA (Lavialle et al., 1982). In an attempt to identify promoters using computer searches (Neural Network Promoter Prediction), an E. coli-like promoter was located in the middle of the gene coding for a probable carboxy kinase (from pTB-249) and was upstream from an in-frame ATG codon. This may be evidence that M. smegmatis can utilise promoter-like sequences for expression when these are correctly placed in relation to an ATG codon. The presence of the transmembrane segment identified in this protein may have allowed the PhoA reporter to
cross the cell membrane and become active.

Of the clones analysed, only two E. coli-like promoters were identified in correct positions for the expression of the putative gene products (pTB-1 and Eco-24). Other researchers have reported the lack of resemblance to E. coli promoter consensus sequences in regions upstream from the putative ATG start codon for many mycobacterial genes (Dale & Patki, 1990). Mycobacterial promoters have also been shown to generally function poorly in E. coli (Sirakova et al., 1989; Das Gupta et al., 1993). This is a likely explanation for the appearance of proportionately more PhoA⁺ colonies isolated in M. smegmatis than in E. coli in this study. Interestingly, the construct Eco-3 was expressed in E. coli, but not when transferred to M. smegmatis (data not shown). It was not recognised as having an E. coli-like promoter as determined by the Neural Network algorithm. The promoter sequences used by E. coli in this case may have been different from those used for predictions, or this gene may have been part of an operon, in which case the promoter may have been present in the sequences further upstream that were not sequenced.

It is not known how many exported proteins are actually produced by M. a. paratuberculosis. ³⁵S-methionine labeling of proteins present in culture filtrates has shown approximately 20 bands present in SDS-PAGE gels (Valentin-Weigand & Moriarty, 1992; White et al., 1994). More than 30 proteins have been found in M. tuberculosis and M. bovis culture filtrates (Andersen et al., 1991a; Nagai et al., 1991; Young et al., 1992). However, many exported proteins, such as lipoproteins, are targeted to the envelope matrix where they are surface-exposed, rather than secreted to the extracellular environment. The advantage of a genetic method such as the PhoA reporter system, is its capability to detect exported proteins regardless of their targeted compartment. Furthermore, the use of this system directly provides the partial DNA sequence of the gene, which facilitates rapid identification of the protein by database searches.

Most exported mycobacterial proteins isolated to date possess signal sequences (Young et al., 1992; Hewinson & Russell, 1993; Chubb et al., 1998). Some of the more interesting M. a. paratuberculosis proteins identified as having signal sequences in this study, included a copper/zinc superoxide dismutase, a cutinase, a penicillin-binding protein and two lipoproteins. These genes or their proteins have not been previously described for M. a. paratuberculosis. However, similar proteins from other species have been investigated and their significance will be summarised.
Superoxide dismutases are metalloproteins that prevent damage by oxygen-mediated free radicals by catalyzing the dismutation of superoxide into molecular oxygen and hydrogen peroxide (Schinina et al., 1989). Superoxide is a by-product of aerobic respiration and is produced by a number of reactions, including the generation of microcidal reactive oxygen species in the phagolysosome of infected macrophages. Copper/zinc superoxide dismutases (SodC) were originally thought to exist only in eukaryotes, but have since been discovered in several bacterial species (reviewed by Battistone et al., 1998), including *M. tuberculosis* (Wu et al., 1998). In *M. tuberculosis*, SodC has been located to the cell surface (Wu et al., 1998) and has been shown to be necessary for resistance to killing during the oxidative burst in infected macrophages (Piddington et al., 2001). The SodC enzymes of *Salmonella typhimurium* (De Groote et al., 1997), *Neisseria meningitidis* (Farrant et al., 1997; Wilks et al., 1998) and *Haemophilus dureyi* (San Mateo et al., 1998) have similarly been implicated in virulence. The *M. a. paratuberculosis* SodC has been cloned and characterised in the course of this study (Dupont & Murray, 2001). Its role in the pathogenesis of *M. a. paratuberculosis* has yet to be investigated.

*M. tuberculosis* encodes a family of six cutinases, all of which are predicted to be exported and for which the function is unknown. In this study, the cutinase identified from the library, was similar to a cutinase isolated from *M. tuberculosis* culture filtrate which has been shown to elicit delayed-type hypersensitivity responses in *M. tuberculosis*-infected guinea pigs, as well as IFN-γ release from lymphocytes in tuberculosis patients and in infected mice (Weldingh & Andersen, 1999). Investigation of the immunological activity of the *M. a. paratuberculosis* cutinase identified in this study would be of interest.

Penicillin-binding proteins are enzymes responsible for the final steps in peptidoglycan formation in cell wall synthesis and are involved in β-lactam antibiotic resistance (Waxman & Strominger, 1983). In *E. coli*, disruption of both of its penicillin-binding proteins is lethal (Suzuki et al., 1978). *E. coli*-like penicillin-binding proteins have been identified in the genomes of *M. tuberculosis* and *M. leprae* (Keer et al., 2000; Mahapatra et al., 2000), but the proteins have not yet been characterised. In *M. smegmatis* three genes encoding penicillin-binding proteins have been investigated and have been shown to be involved in stationary-phase survival (Keer et al., 2000). It would be of interest to investigate the role of penicillin-binding proteins in *M. a. paratuberculosis*.

One of the proteins identified for which the N-terminal sequence was not determined, but
which possessed a predicted transmembrane segment, was a probable serine/threonine protein kinase. Serine/threonine protein kinases are a family of enzymes found in many bacterial species that are responsible for the phosphorylation of proteins. They are principally involved in pathogenicity and differentiation into new developmental states, such as spore formation (Av-Gay et al., 1999). Serine/threonine protein kinases of *Yersinia pseudotuberculosis* (YopO) and *Pseudomonas aeruginosa* (PSTK) have been shown to be associated with virulence in mouse models (Hakansson et al., 1996; Wang et al., 1998). The *M. tuberculosis* genome encodes 11 serine/threonine protein kinases (Cole et al., 1998), of which two (PknB and PknD) have been isolated and are predicted to be transmembrane proteins (Peirs et al., 1997; Av-Gay et al., 1999).

Another protein identified in the library was the previously characterised *M. a. paratuberculosis* 34 kDa protein (De Kesel et al., 1992; 1993; Gilot et al., 1993). Although no signal sequence was predicted using computer programmes, Western blot analyses in this study and by De Kesel and co-workers (1993) indicated that the protein was cleaved at the N-terminus, as evidenced by the presence of two closely-spaced bands. These workers also showed the protein was located at the cell surface and was released extracellularly in stationary phase *M. a. paratuberculosis*, which is consistent with an exported envelope-associated protein. N-terminal sequencing of these two products would confirm if a cleavage site existed. The 34 kDa protein was strongly recognised by sera of infected cattle (De Kesel et al., 1992). A C-terminal recombinant polypeptide fragment (a362) of this protein has been used as a reagent for ELISA tests and was reported to have high specificity and sensitivity (De Kesel et al., 1993; Vannuffel et al., 1994). Specific antibodies to the fragment have been used for immunohistological detection of *M. a. paratuberculosis* in intestinal tissues and lymph nodes of infected cattle and correctly diagnosed the 16 *M. a. paratuberculosis*-infected cattle tested (Coetsier et al., 1998). Although only a small number of cattle were used in these studies (<30) these data support the notion that individual antigens may be employed for improved detection of *M. a. paratuberculosis* infection.

Three putative *M. a. paratuberculosis* lipoproteins were identified in this study. Lipoproteins are generally highly immunogenic in *vivo* and have been implicated as protective antigens in diseases caused by *Haemophilus influenzae* (Green et al., 1987), *Pseudomonas aeruginosa* (Finke et al., 1990), and *Borrelia burgdorferi* (Fikrig et al., 1992). In mycobacteria, lipoproteins such as the 38 kDa protein of *Mycobacterium tuberculosis* (Young et al., 1986; Harboe & Wiker, 1992) and the 19 kDa proteins of *M. tuberculosis* (Jackett et al., 1988; Rees et al., 1993) and *M. intracellularae* (Nair et al., 1992) have been shown to induce both
humoral and cellular immune responses. A DNA vaccine containing the gene for the 38 kDa protein has shown protection against tuberculosis in mice (Zhu et al., 1997). One of the putative lipoproteins identified in this study was similar to the MK35 lipoprotein of *M. kansasii*. MK35 was found to induce specific delayed-type hypersensitivity responses in infected guinea pigs and did not sensitize animals to tuberculosis skin testing (Armoa et al., 1995). In the present study, investigations using the full ORF of the *M. a. paratuberculosis* MK35-like protein showed 69% identity with MK35, 63% identity with the putative *lpqN* gene in *M. tuberculosis*, and 100% identity to a sequence in *M. a. avium*. Preliminary PCR and Southern blot analyses suggested a similar gene was present in *M. intracellulare*, but it was not detected in any of the other limited number of mycobacteria species tested in this study, including members of the *M. tuberculosis* complex. The entire protein has been cloned and expressed in *M. smegmatis* for future analyses.

Another putative lipoprotein identified in this study was similar to the 27 kDa protein of *M. bovis* which was reported to induce humoral and cellular immune responses in cattle with tuberculosis (Bigi *et al.*, 1997). This protein was selected for further investigations, the details of which are presented in Chapter 4.
Chapter 4  Cloning, heterologous expression and characterisation of an immunogenic 22 kDa protein from M. a. paratuberculosis

4.1 Abstract

An exported 22 kDa protein, originally identified from the M. a. paratuberculosis phoA fusion library, was characterised and assessed for its immunoreactivity. The open reading frame was cloned and the protein was expressed in M. smegmatis as a C-terminal polyhistidine-tagged recombinant protein. P22 possessed a characteristic signal sequence and a consensus sequence for lipidation. N-terminal sequencing of the recombinant protein confirmed cleavage of a signal peptide and native P22 was identified in M. a. paratuberculosis strain 316F culture supernatants and cell sonicate fractions. Database searches revealed P22 was present in M. a. avium (100% identity) and was similar to members of the mycobacterial LppX/LprAFG family of lipoproteins, identified in M. tuberculosis, M. bovis and M. leprae. PCR analysis and sequencing showed genes similar to p22 were also present in M. intracellulare and M. scrofulaceum. Humoral immune responses of sheep and cattle to P22 were evaluated by Western blot analyses. Recombinant P22 was recognised by 10 out of 11 sera from M. a. paratuberculosis-vaccinated sheep and by five out of 14 sera from sheep belonging to a naturally infected flock. The protein was recognised by sera from 11 out of 13 subclinically infected cows, diagnosed by faecal culture, and by sera from two clinically affected cows. Furthermore, P22 was recognised by four of the subclinically infected cows that were negative on a commercial ELISA but positive on faecal culture. Cellular immune responses of vaccinated and unvaccinated sheep to recombinant P22 were evaluated using a whole blood IFN-γ EIA assay. P22 induced a significant level of IFN-γ in the vaccinated sheep groups compared to unvaccinated control group.
4.2 Introduction

Protective immunity against Johne’s disease, as with other mycobacterial diseases, is characterised by strong T-cell type 1 responses, dominated by the production of IFN-γ. Stimulation of IFN-γ by proteins exported by mycobacteria has led to the general consensus that these proteins are major immune targets in the early phase of infection. Research to improve the efficacy of tuberculosis vaccines has shown some promising results in animal models with subunit vaccines based on exported proteins such as Ag85A (Huygen et al., 1996; Lozes et al., 1997; Tanghe et al., 2000), the 38 kDa protein (Zhu et al., 1997) and MPT64 (Li et al., 1999). It is possible that such proteins could similarly be identified for use in new vaccines against Johne’s disease.

The available immunodiagnostic tests for Johne’s disease have limitations in sensitivity and specificity. The use of crude antigen preparations, such as PPD can result in immunologic cross-reactions in animals sensitised to other mycobacteria, including those infected with M. bovis (Olsen et al., 2001). Studies with bovine tuberculosis have shown that the specificities of tests for cellular reactivity can be increased using single antigens, such as ESAT-6 and MPB70, although reduced sensitivities are reported (Buddle et al., 1999; Pollock et al., 2000). However, test sensitivities can be increased using a ‘cocktail’ of specific antigens or interpreting tests with individual antigens in parallel (Lyashchenko et al., 1998; Rhodes et al., 2000). This approach has proved useful in serological studies with M. a. paratuberculosis antigens. The low sensitivities of antigens AhpC and the 14 kDa antigen of M. a. paratuberculosis were improved when they were used in parallel (Olsen et al., 2001). Furthermore, these antigen tests had significantly higher specificity compared to that of a commercial ELISA and could differentiate between cows infected with M. bovis and those infected with M. a. paratuberculosis. In bovine tuberculosis, similar results have been shown using tests for cellular reactivity with purified proteins used singly or in combination (Lyashchenko et al., 1998; Buddle et al., 1999; Pollock et al., 2000; Rhodes et al., 2000). The use of well characterised components of M. a. paratuberculosis may provide the means to improve the specificity and sensitivity of immunodiagnostic assays for M. a. paratuberculosis infection.

The aim of this study was to partially characterise a 22 kDa exported protein (P22) from M. a. paratuberculosis, which was identified from an alkaline phosphatase partial gene fusion library, described in Chapter 3. The open reading frame encoding the entire protein was
identified, cloned and expressed as a histidine-tagged recombinant protein in *M. smegmatis*. Recombinant P22 was investigated for its ability to elicit IFN-γ secretion from peripheral blood of sheep vaccinated with a live, attenuated strain of *M. a. paratuberculosis*. Antibody production to P22 was evaluated using sera from vaccinated sheep and naturally infected sheep and cattle.
Figure 4.1. Schematic of the *Mycobacterium* expression vector pMIP12 for the production of histidine-tagged recombinant proteins in *M. smegmatis*. Important features include an *E. coli* origin of DNA replication (ori *E. coli*) for propagation of plasmid in *E. coli*; a *Mycobacterium* origin of DNA replication (ori Myco); a kanamycin resistance marker (kan'); the mycobacterial β-lactamase promoter (*pblaF*; Timm *et al.*, 1994); a ribosome binding site (RBS); start of translation codon (ATG); multiple cloning site; a C-terminal histidine x 6 tag; double translational stop sequence (TAG TGA); a transcriptional terminator. The plasmid was a gift from Professor Brigette Gicquel, Pasteur Institute, Paris.
4.3 Materials and methods

4.3.1 PCR amplification of the p22 gene from *M. a. paratuberculosis*

Genomic DNA extracted from *M. a. paratuberculosis* ATCC 53950 (see 3.3.1.1) was used as a template for PCR amplification of p22. Oligonucleotide primers were designed to the 5' and 3' ends of the entire ORF and are shown below. The forward primer lpp27-fBam, was designed to the 5' end of the predicted ORF using the sequence obtained from the pJEM11 construct pTB-16 (see 4.4.1.2). The predicted start codon of the gene is underlined. A *Bam* HI site, shown shaded, was incorporated for cloning. The reverse primer lpp27-rKpn, was designed to the 3' end of the ORF using sequence data from the *M. a. paratuberculosis* TIGR database (see 4.4.1.2). A *Kpn* I site, shown shaded, was incorporated for directional cloning and the TGA stop codon was omitted to allow read-through to produce the histidine x 6 tag coded by the vector.

lpp27-fBam  5' GATGCGATCCATGAGCCCGCCGCCT
lpp27-rKpn  5' TGAGGATTACCGAGCTCAACCCGGCTTG

PCR was done using 1 µl of a 1:10 dilution of genomic DNA template in a volume of 50 µl. The conditions used were an initial melting temperature of 95°C for 10 min followed by 35 cycles of 94°C for 30s, 55°C for 30s, 68°C for 1 min, and a final extension at 68°C for 10 min using platinum Pfx polymerase, in the presence of 10% (v/v) dimethylsulfoxide.

4.3.2 Cloning of the p22 open reading frame

For expression of p22 in *M. smegmatis*, the ORF was cloned into the vector pMIP12 (see Figure 4.1). The 725 base pair PCR product was extracted from a 1% TAE agarose gel. The purified product was eluted in 30 µl of distilled water and a 6 µl sample was run on a 1% agarose gel to check its recovery. The remaining DNA was digested with 40 units of *Kpn* I in 40 µl at 37°C overnight. This was followed by digestion with *Bam* HI in a total volume of 50 µl at 37°C overnight. The digested product was purified and resuspended in 30 µl of sterile water using a QIAgen gel extraction kit, following the recommendations for restriction endonuclease removal. From this, 6 µl was electrophoresed on a 1% TAE agarose gel to estimate its concentration. The remainder was used for ligation with pMIP12, as described below.
The vector pMIP12 was isolated from previously transformed *E. coli* DH10B cells. Plasmid DNA from two 4 ml cultures grown in LB/kan broth was resuspended in 100 µl of distilled water. To quantitate DNA concentration, 20 µl was mixed with 230 µl of TE buffer and the absorbance at 260 nm was read in a 0.5 cm quartz cuvette using a double beam spectrophotometer (Helios α, Unicam, UK). Approximately 2.6 µg of plasmid DNA was digested with 20 units of Kpn I in 20 µl for 4 h at 37°C. To confirm completion of digestion, 1 µl was removed and electrophoresed on a 1% TAE agarose gel and photographed. To the remaining 19 µl, 30 units of Bam HI were added and the volume was brought up to 30 µl with water. The mixture was incubated for 18 h at 37°C. The digested vector was electrophoresed on a 1% low melting point agarose gel (SeaPlaque GTG agarose, FMC BioProducts, Rockland, ME, USA) in TAE buffer and ethidium bromide stained for visualisation. The gel slices were cut out in a minimal amount of agarose, transferred to a 1.5 ml microfuge tube and melted at 68°C for 10 min. From this mixture, 6 µl was removed and immediately added to a microfuge tube containing 24 µl of insert DNA, prepared above. To this was added 1 unit of T4 DNA ligase in a total volume of 40 µl and the mixture was incubated at 14°C overnight. This was dialysed and electroporated into 35 µl of *E. coli* ElectroMax DH10B cells as described. To this, 200 µl of LB broth was added and the mixture was incubated at 37°C for 1 h. Of this, 50 µl was spread onto a LB/kan plate and incubated overnight at 37°C.

To confirm the presence of the insert, PCR was carried out on five of the resulting colonies using the pMIP12 primers BlaF3 and R2 (see below), in the presence of *taq* polymerase in 20 µl. Eight microlitres from each PCR was electrophoresed on a 1% TAE agarose gel and ethidium bromide stained. The forward primer BlaF3 binds at approximately 150 base pairs within the *blaF* promoter and is designed to the coding strand. The reverse primer R2 binds 44 base pairs from the Kpn I site and is designed to the complementary strand of the transcriptional terminator. The expected size of the PCR product for P22 insert was approximately 900 base pairs.

**BlaF3** 5' TCGCGGACTACGGTGCC

**R2** 5' TCGAACTCGCCCGATCCC

From a resulting PCR-positive colony, plasmid DNA was extracted from 4 ml of broth culture and resuspended in 50 µl of 10 mM Tris (pH 8.0). Eight microlitres of plasmid was
digested with 10 units of \textit{Kpn I} in 10 µl at 37°C overnight. This was followed by digestion with 10 units of \textit{Bam HI} in a final volume of 25 µl at 37°C overnight. One microlitre was electrophoresed on a 1% TAE gel and photographed after ethidium bromide staining. To confirm the insert identity and the correct insertion for expression, plasmid was used as template for sequencing using the primers Bla3 and R2.

4.3.3 Expression and purification of P22 recombinant protein from \textit{M. smegmatis}

For transfer of pMIP-p22 into \textit{M. smegmatis}, 1 µl of the plasmid DNA was electroporated into 20 µl of \textit{M. smegmatis} cells. A 250 µl aliquot of cells was spread on a LB/kan plate and incubated at 37°C for three days. To confirm the presence of pMIP-p22, one colony was picked for PCR analysis using the primers R2 and Bla3 as described above, in a volume of 100 µl. The resulting 900 base pair product was gel-extracted and sequenced using the same primers.

For preparation of recombinant protein, a single colony of pMIP-p22-transformed \textit{M. smegmatis} was picked and inoculated into 15 ml of Sauton’s broth or modified Middlebrook 7H9 broth (Appendix 1) containing kan and grown at 37°C with vigorous shaking for approximately three days. This was used to inoculate 600 ml of the same media, and the culture was grown and harvested and lysates were prepared as described.

The resulting sonicate supernatant was used in Ni$^{2+}$-affinity chromatography, employing four 5 ml columns connected in series and attached to a peristaltic pump. Imidazole concentrations of 40 mM, 250 mM and 1M were used, in the first instance, to determine the elution profile of recombinant P22. Samples from each collected elution were used in Western blot analyses to determine where the majority of the recombinant P22 eluted. Having determined this, the procedure was repeated using two washes of 100 mM of imidazole prior to the 250 mM elution, to further purify the recombinant protein. Pooled fractions were concentrated as described, in preparation for Western blotting, IFN-γ assays or size exclusion chromatography.

4.3.3.1 Western blot analyses of P22 recombinant protein

Western blots were carried out as described in section 2.7.2. For detection of histidine-tagged P22 from \textit{M. smegmatis}, anti-histidine x 6 POD conjugated antibody (Roche
Molecular Biochemicals, Germany) was added to blots at 1:500 (v/v) dilution and incubated for 1 - 2 h with shaking at room temperature. Blots were subsequently washed and developed.

For detection of antibody to P22 from animal sera, identical amounts of recombinant protein (approximately 0.5 µg) were electrophoresed in individual lanes, alongside a molecular weight standard, on SDS-PAGE gels. The protein was transferred to PVDF or nitrocellulose membranes and stained with Ponceau S to visualise the bands. The lanes were numbered at the top of the membrane and then cut into individual strips.

For immunodetection, single strips were incubated with serum in individual small glass screw-capped bottles in 2 ml of blocking buffer as described (see 2.7.2.1). The strips remained in the bottles for washing and were then pooled in plastic trays for incubation with appropriate secondary antibody and washing. To avoid cross-contamination, strips from different experimental groups, e.g. unvaccinated and vaccinated, were pooled separately during this stage. The following secondary POD conjugated anti-IgG heavy and light chain antibodies were used: goat anti-rabbit (A 6154), donkey anti-sheep (A 3415), rabbit anti-bovine (A 7414), (Sigma, USA).

4.3.4 Preparation of rabbit antibody raised to P22

Approximately 0.05 mg of recombinant P22 was used to immunise a rabbit on two occasions (see 2.8.2). To confirm the production of antibody to P22, immunodetection of Western blots of P22 was done using a 1:1,000 dilution of naïve and immunised rabbit sera, followed by anti-rabbit IgG POD conjugated antibody at 1:20,000 dilution.

4.3.5 PCR amplification of the p22 ORF from genomic DNA

PCR amplifications were carried out on mycobacterial genomic DNA using the primer pairs lpp27-fBam and lpp27-rKpn (see 4.3.1). The reactions were carried out in 20 µl using Taq DNA polymerase. Template DNA consisted of 2 µl of a 1:50 dilution of purified DNA or, alternatively, single bacterial colonies were added directly to the PCR mixture.
4.4 Results

4.4.1 Sequence analysis of plasmid pTB-16 and identification of the p22 open reading frame

To identify the *M. a. paratuberculosis* gene encoding the putative exported protein expressed as a partial PhoA fusion product from plasmid pTB-16, the cloned insert was partially sequenced.

4.4.1.1 Sequence analysis of pTB-16

A 625 base pair segment of sequence was obtained adjacent to *phoA* from pTB-16. This segment contained two potential ATG start codons in-frame with *phoA*, located 60 and 69 base pairs upstream of the *phoA* junction. At 120 base pairs upstream of *phoA*, was an in-frame TGA stop codon, suggesting the start of the ORF encoding the fusion protein was one of these two preceding ATG sequences. Database searches using the 69 base pair nucleic acid sequence containing both ATGs, resulted in no matches. However, when sequences reaching to the stop codon were included, two matches were obtained. The highest scoring match was to the *M. tuberculosis*/*M. bovis* *lprG/lpp-27* gene (identities 28/30 93%; EMBL accession AJ000500.1). The second was to the related *M. leprae* *lprG* gene (identities 24/25 96%; EMBL accession AL583918.1). Accompanying database annotations described these genes as encoding putative lipoproteins containing probable N-terminal signal sequences. These preliminary alignments encompassed only 30 base pairs in the region of the potential start codons. Given such a short segment of alignment, further evidence of the similarity of this sequence to *lpp-27/lprG* was achieved by investigation of the sequences further upstream of the stop codon. The arrangement of genes in pTB-16 was compared to the arrangement upstream of *lprG* in *M. tuberculosis*. Database searches resulted in significant identity (82% over 364 base pairs) to *ribC* (riboflavin synthase alpha chain; EMBL accession MT072921.1) of *M. tuberculosis*. This similarity began approximately 50 base pairs upstream from the stop codon, included the *ribC* ATG start of translation and continued to the end of the 625 base pair segment of obtained sequence. The arrangement of *ribC* and *lpp-27/lprG* in *M. tuberculosis*, *M. bovis* and *M. leprae* databases and in pTB-16, revealed that the two genes were adjacent to one another in all cases, with *ribC* encoded in the reverse direction on the opposite strand from *lpp-27/lprG*. Thus, it appeared that pTB-16 was expressing the 5’ end
Figure 4.2. Sequence analysis of the p22 ORF. The predicted start codon is in red. The cloned portion of the ORF in the original pJEM11 PhoA+ construct (pTB-16) is indicated in blue. The entire ORF is 705 base pairs in length. The corresponding translated 235 amino acid sequence is shown as derived from the DNA sequence using the program Translate from the ExPASy Molecular Biology Server.
Figure 4.3. Amino acid sequence comparison between P22 of *M. a. paratuberculosis* and database search results. Identity to *M. a. avium* (TIGR database) was 100% and is shown in number 1. Descriptions of the remaining Blast results, numbers 2 to 7, are included with the accession numbers in red. Identical amino acids are highlighted. Pink represents the protein segments sharing the least identity and blue represents segments sharing the greatest identity. To the right of the alignments are the BLAST results for % identity and % similarity of the proteins compared to the query sequence of *M. a. paratuberculosis* P22. The lipoprotein consensus sequence, characteristic of PS00013 of ExPASy Prosite database (http://www.expasy.ch/prosite/), is boxed.
of a gene related to *lpp-27/lprG*. As the protein encoded by *lpp-27* from *M. bovis* (the 27 kDa protein/P27) has been reported to be immunogenic, further investigation of this gene from *M. a. paratuberculosis* was initiated.

### 4.4.1.2 Identification of the p22 ORF

To obtain the full sequence of the P22 ORF for cloning, the closely related *M. a. avium* incomplete database (accessed through TIGR) was searched using the 69 base pair 5’ end of the ORF, which included both potential start codons. This database was used because, unlike the *M. a. paratuberculosis* database, contiguous sequences flanking the resulting alignment were provided. A 100% identity was obtained and the *M. a. avium* sequence was examined for the first in-frame stop codon, thus defining the 3’ end of the ORF. This *M. a. avium* ORF sequence was in turn used to search the incomplete *M. a. paratuberculosis* database (accessed through TIGR). The match was 99%, with a single base change approximately two-thirds of the way from the 5’ end of the ORF. This did not alter the translated sequence.

To define the start codon of the ORF, the *M. a. paratuberculosis* sequence was translated using the programme Translate, ExPASy Molecular Biology Server, and used in protein database searches. Comparison with the N-termini of the resulting aligned sequences (see Figure 4.3) allowed prediction of the N-terminal methionine, encoded at 60 base pairs in the pTB-16 sequence. The DNA and translated sequence of the defined ORF is presented in Figure 4.2. It encodes a precursor protein of 235 amino acids of calculated molecular weight 24.4 kDa. Further experimental analysis of the recombinant protein showed it was processed into a mature protein of 216 amino acids with a calculated weight 22.3 kDa (see 4.4.3) and was therefore designated P22. The G+C content of the *p22* ORF was 66.4%, which is consistent with the published G+C content for *M. a. paratuberculosis* (Imaeda *et al.*, 1988).

### 4.4.2 Sequence analysis of p22

#### 4.4.2.1 Sequence similarities between P22 and a family of mycobacterial lipoproteins

The translated DNA sequence was used in database searches to find the extent of similarity of the full-length protein with others in the database. The identity to *M. a. avium* (TIGR database) was 100%. The remaining alignments are presented in Figure 4.3. The first 70 amino acids of P22 share the least identity with the sequences presented. For *M. leprae* LprG, *M. tuberculosis/bovis* LprG and *M. tuberculosis* LprA, the identities for this N-
Figure 4.4. Kyte-Doolittle plot (top) and signal sequence features (bottom) of the P22 precursor protein. Hydrophobic regions are shown below the line at 0.00 and hydrophilic regions are above. Predicted signal sequence domains and the cleavage site, derived from N-terminal sequencing of the mature recombinant protein from M. smegmatis, are shown below. ↓ = positively charged    ↓ = hydrophobic    ↓ = polar
### Figure 4.5. Comparison of the promoter regions of *M. bovis* *ipp*-27 and *M. a. paratuberculosis* *p*22.

The predicted -35 and -10 regions of *M. bovis* are shaded blue. The essential promoting region in *M. bovis*, determined by deletion analysis (Bigi et al., 2000), is underlined. For *M. bovis*, the predicted ATG start codon is shaded black and the predicted ribosome binding site (RBS) is shaded red. Nucleotides shown shaded grey in *M. a. paratuberculosis* are shared by *M. bovis*.

<table>
<thead>
<tr>
<th></th>
<th>M. <em>ptb</em></th>
<th>M. <em>bovis</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GCTAATTGCCCTCTCAGGCTGGTGACGAGGCACAACCGCCTGTGAGAC</td>
<td>CTTCCCAACAATTCCGGTGAAACTCCAGGTTAAGCAGCAATAGCCCAATCCCGGGAC</td>
</tr>
<tr>
<td></td>
<td><strong>-35</strong></td>
<td><strong>-10</strong></td>
</tr>
<tr>
<td></td>
<td>ACCGGGCCCCCCTGGTCACCGGGGAGTCACCAGGACCCCTCTACGATGCTGGGTATAG</td>
<td>CCGCACCCCACTGCGACCGGCCAACTCAGAGACCCCTCTACGATGCAGGTATG</td>
</tr>
</tbody>
</table>

The RBS (ribosome binding site) is shaded red.
terminal region were 65.8%, 50% and 32.9%, respectively.

These similar proteins belong to the mycobacterial LppX/LprAFG family of lipoproteins, the function of which is unknown. All have putative signal sequences and potential lipid attachment sites, based on the motif around the cleavage sites in their mature N-terminus (von Heijne, 1989). The P22 sequence was therefore examined for similar features associated with an exported protein. The N-terminal region contains a characteristic signal sequence, shown in Figure 4.4. It is composed of a positively-charged n-region of six amino acids, followed by a core hydrophobic h-region of 10 amino acids and a c-region of three amino acids containing the signal peptide cleavage site. The strong positive n-region charge, dominated by three R residues, is a common feature of signal peptides of Gram-positive bacteria (von Heijne & Abrahmsen, 1989; von Heijne, 1990b). The cleavage site (obtained by N-terminal sequencing of the recombinant protein from *M. smegmatis*; see 4.4.3) appeared to be typical of cleavage sites used by signal peptidase I enzymes, which cleave non-lipoprotein precursors. It contained the conserved sequence AXA, which immediately precedes the cleavage site (von Heijne, 1985). In the precursor protein, this sequence is ATA at positions 17-19. At positions 21-25 of the precursor protein is the sequence IAGCS which is similar to the consensus sequence for lipidation (lipobox) and cleavage: L(A/S)(G/A)↓C(S/G), based on predicted and known lipoproteins of Gram-negative and Gram-positive bacteria (Braun & Wu, 1994; Sutcliffe & Russell, 1995). The N-terminal sequencing result of the recombinant protein (see 4.4.3) showed this cleavage site was not used in *M. smegmatis*.

The DNA sequences upstream of the start codon at 69 base pairs contained no clear ribosome binding site, as based on the *E. coli* consensus 5' AGGAGG. However, a G-rich motif was found just before the ATG at 60 base pairs. This is in agreement with the predicted ribosome binding site for *M. bovis lpp-27* (Bigi *et al.*, 1997), (see Figure 4.5) and similarly favors this ATG as the true start codon. Determinations of the transcription initiation point and essential promoting region of *lpp*-27 have been carried out (Bigi *et al.*, 2000), and these sequences were compared to those upstream of the *p22* ORF. As shown in Figure 4.5, there is little similarity between the promoter region of *M. bovis lpp*-27 and that for *M. a. paratuberculosis* *p22*. The greatest region of similarity is 27 base pairs immediately upstream of the predicted ATG start codon. This segment of sequence was not essential for transcription of *lpp*-27 (Bigi *et al.*, 2000).
Figure 4.6. Restriction endonuclease digest of plasmid pMIP-p22. The p22 ORF was directionally cloned into the Kpn I and Bam HI sites of the vector pMIP12. The plasmid was digested with Kpn I and Bam HI and electrophoresed in a 1% agarose gel, ethidium bromide stained and photographed under UV light. pMIP12 vector, approximately 7,000 base pairs in size, and the 705 base pair p22 insert are indicated. A DNA size standard is shown in the left lane for comparison.
Figure 4.7. Expression of recombinant P22 from \textit{M. smegmatis}. Protein was prepared from sonicated cells harbouring plasmid pMIP-p22 or pMIP12. Resulting soluble and insoluble fractions were electrophoresed in 15\% SDS-PAGE gels. (a) Coomassie Blue stained gel. (b) Western blot analysis using 1:500 anti-histidine x 6 POD conjugated antibody. Lane M, molecular weight standard; lane 1, pMIP12 insoluble; lane 2, pMIP12 soluble; lane 3, pMIP-p22 insoluble; lane 4, pMIP-p22 soluble. Lanes 5 to 8 contain protein fractions collected from Ni$^{2+}$-affinity chromatography of the soluble fraction of cells harbouring pMIP-p22. Lane 5, unbound protein; lane 6, 40 mM imidazole elution; lane 7, 250 mM imidazole elution; lane 8, 1 M imidazole elution. Recombinant histidine-tagged P22 protein is indicated by the arrows.
4.4.3 Cloning and expression of the p22 ORF

For the expression of p22 in a mycobacterial host, the ORF was cloned into the vector pMIP12 and used to transform *M. smegmatis*. A restriction endonuclease digest of the resulting pMIP-p22 plasmid is shown in Figure 4.6. DNA sequencing of the insert confirmed the identity of the cloned gene and that the correct reading frame was in place for expression.

In order to increase the likelihood of producing P22 in a form resembling that from the original host (*M. a. paratuberculosis*), it was expressed in the fast-growing species *M. smegmatis*. To aid purification and detection of the recombinant protein, the gene was expressed with a C-terminal histidine x 6 tag from the vector pMIP12. As shown in Figure 4.7, the recombinant protein was detected from cell sonicates in both the soluble and insoluble fractions by Western blot analysis using a monoclonal anti-histidine x 6 POD conjugated antibody. The recombinant protein was clearly recognised by the antibody (Figure 4.7 b) with no other signal obtained, confirming specificity for the histidine x 6 tag. The protein was further isolated from the soluble fraction by Ni\(^{2+}\)-affinity chromatography followed by elution with imidazole (see 4.3.3). Recombinant P22 could be seen in the 250 mM imidazole elution in SDS-PAGE gels with Coomassie Blue staining (see Figure 4.7 a). Western blot analysis showed the 250 mM imidazole elution contained most of the P22, with a slightly smaller amount present in the 40 mM elution and none was detected in the 1 M elution. There was a small amount of the protein still present in the flow-through, indicating that not all of the protein was bound to the column. This may be because the column was saturated with bound protein.

The apparent molecular weight of the recombinant protein from SDS-PAGE was 23 kDa, which was smaller than the calculated weight of 25.9 kDa for the full-length recombinant protein, expressed from pMIP12. This suggested that the full-length protein may have been cleaved at the N-terminus, as was predicted from its amino acid sequence, and so N-terminal sequencing of the 23 kDa band was carried out, as described (see 2.7.3). The resulting amino acid sequence obtained, **LL'GCS**, was consistent with an N-terminal cleavage predicted at amino acid position 19/20 in the native protein by SignalP (see Figure 4.4, amino acids 20 – 25). The theoretical size of the mature recombinant product based on this cleavage was 23.6 kDa and was similar to the apparent molecular mass from SDS-PAGE.
Figure 4.8. Affinity chromatography of recombinant P22. Cell lysates from *M. smegmatis* containing PMIP-p22 were applied to Ni\(^{2+}\)-affinity columns and bound proteins were eluted with imidazole as described. Fraction samples of 15 µl were electrophoresed in 15% SDS-PAGE gels and stained with Coomassie Blue. Lane M, molecular weight standard; lanes 1 to 14, samples from sequential elution fractions. The arrow indicates P22.
Figure 4.9. Detection of recombinant P22 from *M. smegmatis* culture filtrates. Concentrated culture filtrates from mid-log phase *M. smegmatis* containing the plasmid pMIP-p22 were prepared. Equivalent amounts (~3 μg) were electrophoresed in duplicate in 15% SDS-PAGE gels. (a) Coomassie Blue stained gel. (b) Western blot analysis using 1:500 anti-histidine x 6 POD conjugated antibody. Lane M, molecular weight standard; lane 1, *M. smegmatis* pMIP12 culture filtrate; lane 2, *M. smegmatis* pMIP-p22 culture filtrate. Mature recombinant histidine-tagged P22 is indicated by the arrow.
Figure 4.10. Detection of antibody to P22 in sheep vaccinated with Neoparasec. Western blots of recombinant P22 were individually incubated with 1:1,000 dilution of sera. Anti-sheep IgG POD conjugated antibody was used at 1:20,000. (a) Mob 1 sheep. Lane M, molecular weight standard (kDa); lanes 1 to 6, pooled three and seven months post-vaccination sera from animals 124, 127, 129, 131, 132 and 136, respectively. Lanes 7 to 12, pooled one and two month pre-vaccination sera from the same animals. Lane 13, anti-histidine x 6 POD conjugated antibody control. (b) Mob 2 sheep. Lane M, molecular weight standard (kDa); lanes 1 to 5, post-vaccination sera from animals 507, 578, 587 (three months post-vaccination), 598* (two months post-vaccination), and 560* (one month post-vaccination), respectively. Lanes 6 to 10, one month pre-vaccination sera from the same animals, respectively. Lanes 11 to 14, sera from unvaccinated animals 599, 569, 527 and 538, respectively, taken at the equivalent of three months post-vaccination.

* died after this time
4.4.4 Purification of recombinant P22 from cell lysates

It was apparent from Figure 4.7 that the recombinant protein was fully eluted from the column using 250 mM imidazole, however contaminating proteins were also present in the eluate when collected in batches of one column volume. When washed with 100 mM imidazole and eluted in a series of 1 ml fractions, the protein was recovered in a more enriched form, as shown in Figure 4.8.

4.4.5 Analysis of M. smegmatis culture filtrates for the presence of recombinant P22

To search for the presence of P22 in culture supernatants of M. smegmatis containing plasmid pMIP-p22, concentrated preparations of culture filtrates were prepared. Immunodetection of Western blots using anti-histidine x 6 POD conjugated antibody was carried out using standard procedures. Figure 4.9 shows the protein was present in culture filtrates of M. smegmatis containing pMIP-p22, and not in control M. smegmatis containing vector only. The apparent molecular weight was 23 kDa and was consistent with the size determined from SDS-PAGE from M. smegmatis sonicates.

4.4.6 Immune responses to P22

To investigate immune responses to P22, Neoparasec-vaccinated animals were used in the first instance. It was hypothesised that vaccinated animals would make an immune response to antigens possessed by the bacteria in the vaccine. In this respect, they may be similar to infected animals and therefore could be used to evaluate the immunogenicity of P22. M. a. paratuberculosis vaccinated cattle and sheep have similarly been used to evaluate various mycobacterial proteins (Valentin-Weigand & Moriarty, 1992; Koets et al., 1999; 2001).

4.4.6.1 Humoral immune responses to P22

Serum samples from Neoparasec-vaccinated and naturally infected sheep (see section 2.10 for details) were used in individual immunoblot assays to determine their humoral response to P22. A high level of antibody to P22 was produced by 10 of the 11 Neoparasec-vaccinated sheep in Western blots analyses, demonstrating the protein was immunogenic (see Figure 4.10). There was a very faint band produced by sera from three of the sheep (507, 578 and 560) prior to vaccination and also in two naïve animals (599 and 569). Control sera
Figure 4.11. Detection of antibody to P22 in individual sheep from a naturally infected flock. Western blots of recombinant P22 were individually incubated with 1:500 dilution of serum. Anti-sheep IgG POD conjugated antibody was used at 1:40,000 dilution. Blots were developed by chemiluminescent detection. Lane M, molecular weight standard (kDa); lanes 1 to 14, sheep number 48, 44, 40, 36, 6, 13, 2, 23, 51, 32, 25, 27, 26 and 43, respectively. * animals diagnosed with Johne’s disease (see Appendix 2).
Figure 4.12. Detection of antibody to P22 in naturally infected cattle. Western blots of recombinant P22 were individually incubated with 1:500 dilution of serum. Anti-bovine IgG POD conjugated antibody was used at 1:20,000 dilution. Blots were developed by chemiluminescent detection. Lanes M, molecular weight standard (kDa); lane 1 antihistidine x 6 POD conjugated antibody control; lanes 2 to 14, subclinical cattle which were positive on at least one faecal culture (animal 24, 2, 275, 144, 327, 181, 115, 34, 49, 517, 168, 58, 68, respectively); lanes 15 and 16, clinically affected cattle (animal 27 and 25); lanes 17 to 22, cattle that were negative on all faecal culture and serum ELISA tests (animal 211, 132, 193, 97, 174, 53). The position of P22 is indicated on both sides by arrows. See Table 4.1 for ELISA and faecal culture results.
from the remaining animals prior to vaccination and from unvaccinated control animals did not react with the protein (not all are shown).

To investigate if antibody to P22 was present in naturally infected animals, Western blots were probed with sera from sheep (see 2.10.2) and cattle belonging to flocks and herds known to have Johne's disease present. Results with sheep sera are shown in Figure 4.11. A total of five of 14 sheep had antibody to P22. Antibody to P22 was detected in two of the five sheep with confirmed Johne's disease (see Appendix 2) and in two sheep that showed no evidence of Johne's disease. A weak band was produced with one animal (26) that had a single suspicious lesion without acid-fast organisms in an examined lymph node but was not confirmed as having Johne's disease. Immunoblots were repeated several times with consistent results.

Figure 4.12 shows the results of P22 Western blot analysis using cattle sera. A summary of the results for serum ELISA, faecal culture and P22 Western blot analysis is shown in Table 4.1. Antibody to P22 was present in sera from 11 of 13 subclinically infected cattle that were positive on at least one previous faecal culture. A variety of strong and weak bands were produced. Three of the four cows that were positive on all ELISA and faecal culture tests (24, 144 and 49 but not 2) also produced major bands with P22. Two subclinically infected cows (181 and 68) did not have detectable antibody to P22 and were also negative on the last ELISA and previous faecal culture. However, other animals with similar test results, such as 58 and 34, produced strong bands with P22. Antibody to P22 was present in four subclinically infected cows (327, 517, 58, 115) that were negative on all three serum ELISA tests. Antibody to P22 was also present in both clinically affected cows (27 ands 25). Cow 27 produced a very strong band with P22. Of the six cows that were negative on all ELISA tests and faecal cultures, one (53) produced a weak band with P22.
Figure 4.13. IFN-γ induction using Ni⁺²-affinity-enriched P22 in Neoparasec-vaccinated sheep blood. Whole blood was incubated with 12.5 μg/ml Avian PPD in duplicate wells and 2.6 μg/ml, 0.64 μg/ml and 0.32 μg/ml Ni⁺²-affinity enriched P22, in single wells. IFN-γ assays were performed as described. Results were expressed as “corrected” absorbance at 450 nm. For Avian PPD, this was defined as the average A₄₅₀ nm of the Avian PPD-stimulated wells minus the average A₄₅₀ nm of the PBS control wells for that animal. For each P22 concentration, this was defined as the A₄₅₀ nm of the P22-stimulated well minus the average A₄₅₀ nm of the PBS control wells for that animal. See Appendix 5 for raw data.
Table 4.1. Summary of results for detection of *M. a. paratuberculosis* by serum ELISA, faecal culture and P22 Western blot analysis

<table>
<thead>
<tr>
<th>Animal #</th>
<th>ELISA 1</th>
<th>ELISA 2</th>
<th>ELISA 3</th>
<th>Faecal culture 1</th>
<th>Faecal culture 2</th>
<th>P22</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>ND</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>275</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>144</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>327</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>181</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>115</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>34</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>49</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>517</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>168</td>
<td>ND</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>58</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>68</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>27#</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>25*</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>211</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>132</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>193</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>97</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>174</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>53</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Serum ELISA and faecal culture were carried out and interpreted as positive or negative by AgResearch, Wallaceville, Lower Hutt, New Zealand. Sera and faecal culture tests were done in six-monthly intervals. Sera used for Western blot analysis was from the last collection date, corresponding to ELISA 3.

ND = not done

* Johne's disease diagnosed on clinical signs and gross pathology
# Johne's disease diagnosed on clinical signs and acid-fast organisms in faeces

4.4.6.2. Cell-mediated immune responses to P22

To investigate if recombinant P22 could stimulate CMI responses in Neoparasac-vaccinated animals, whole blood IFN-γ assays were carried out. In the first instance, Ni²⁺-affinity enriched recombinant P22 was tested in single wells in three different concentrations (2.6 μg, 0.64 μg and 0.32 μg). Results are shown in Figure 4.13. There was a significant difference (p<0.01) between the Neoparasac-vaccinated and unvaccinated group in the IFN-γ responses to P22 at all three concentrations. Eight of the nine vaccinated animals showed IFN-γ production to all three concentrations of P22, often in a concentration-dependent manner.
Figure 4.14. IFN-γ induction by purified recombinant P22 in Neoparasec-vaccinated sheep blood. Whole blood was incubated in duplicate wells with 12.5 μg/ml Johnin PPD and 1 μg/ml size-exclusion purified P22. To demonstrate a concentration-dependent response, Neoparasec vaccinated animals 124, 127 and 129 and unvaccinated animals 128, 133 and 137 were similarly tested with 5 μg/ml purified P22. PBS was included as a negative control in duplicate wells. IFN-γ assays were performed as described. Results were expressed as “corrected” absorbance at 450 nm, defined as the average A_{450 nm} of the stimulated wells minus the average A_{450 nm} of the PBS control wells for that animal. See Appendix 6 for raw data.
Figure 4.15. Detection of antibody to P22 from sheep vaccinated with *M. a. paratuberculosis* strain 316F culture filtrate. Western blots of recombinant P22 were individually incubated with 1:500 dilution of serum as described. Anti-sheep IgG POD conjugated antibody was used at 1:40,000 dilution. Blots were developed by chemiluminescent detection. Lane M, molecular weight standard (kDa); lanes 1 to 5, one month post-vaccination animals 571, 513, 514, 512 and 551, respectively; lanes 6 to 10, pre-vaccination, same animals.
Animal 136 had very low IFN-γ production to P22 and to Avian PPD. The low response to Avian PPD was consistent in this animal over the previous five months of testing (data not shown). None of the unvaccinated animals had notable IFN-γ responses to P22 however, three of these animals (128, 133, 569) had comparatively large responses to Avian PPD, especially animal 569. Reactions of control animals to Avian PPD were not uncommon and appeared occasionally in various animals during testing over the previous five months (data not shown).

To demonstrate that the immunologically active component in the Ni²⁺-affinity preparation was P22, the protein was purified by size-exclusion chromatography and tested for its ability to stimulate IFN-γ production in whole blood. Results are shown in Figure 4.14. There was a significant difference (p<0.05) in the IFN-γ responses to 1 μg of P22 between the vaccinated and unvaccinated group, despite the low responses of animals 578 and 587 in the vaccinated group and the notable response in animal 133 in the unvaccinated group. Animal 136 was not included in the assay because of its consistently poor response to Avian and Johnin PPD. To see if the IFN-γ production to P22 was concentration-dependent, three vaccinated animals were chosen from Mob 1, along with three unvaccinated animals for testing using 5 μg of purified P22. Only three animals from each group were tested due to limited amounts of purified P22. Two animals in the vaccinated group had evident concentration-dependent responses to the antigen and the third had a slight concentration-dependent response.

4.4.7 Localisation of P22 in M. a. paratuberculosis

To investigate whether antigen P22 is exported in its native host, M. a. paratuberculosis, two approaches were taken. An indirect approach was to determine if animals vaccinated with M. a. paratuberculosis strain 316F culture filtrate produced antibody to P22. A direct approach was to determine its location in Western blots of culture filtrate and cell sonicate fractions using rabbit antibody raised to P22.

4.4.7.1 Detection of P22 with serum from sheep vaccinated with M. a. paratuberculosis strain 316F culture filtrate

Serum samples from culture filtrate-vaccinated sheep (see 2.10.1) were used to probe individual Western blots of recombinant P22. Results are shown in Figure 4.15. All five vaccinated animals had antibodies to the antigen at one month post-vaccination. Antibody
Figure 4.16. Western blot detection of rabbit antibody raised to P22. Ni\textsuperscript{2+}-affinity enriched recombinant P22 was electrophoresed in SDS-PAGE gels and transferred to PVDF membranes for subsequent immunodetections with the following sera: lane 1, P22-immunised rabbit serum (1:1,000); lane 2, naïve rabbit serum (1:1,000); lane 3, control anti-histidine x 6 POD conjugated antibody (1:500). Lane M, molecular weight standard (kDa). Secondary anti-rabbit IgG POD conjugated antibody was used at 1:20,000. Blots were developed using chemiluminescent detection.
Figure 4.17. Detection of native P22 in Western blots of *M. a. paratuberculosis* strain 316F cell fractions and comparison to recombinant P22 using rabbit antibody raised to P22. (a) *M. a. paratuberculosis* cell fractions. Lane M, molecular weight standard (kDa); lanes 1 and 4, 10 μl (0.5 mg) of 200 fold concentrated culture filtrate; lanes 2 and 5, 10 μl of equivalent 200 fold concentrated soluble cell lysate fraction; lanes 3 and 6, 10 μl of equivalent 200 fold concentrated insoluble cell fraction. Lanes 1 to 3 were probed with anti-P22 rabbit sera. Lanes 4 to 6 were probed with naïve rabbit sera. (b) Native P22 in *M. a. paratuberculosis* culture filtrate and recombinant P22 from *M. smegmatis*. Lane M, molecular weight standard (kDa); lane 1, 5 μg culture filtrate; lane 2, 5 μg culture media only; lane 3, 1 μg Ni<sup>2+</sup>-affinity purified recombinant P22 from *M. smegmatis* lysate; lane 4, 0.1 μg Ni<sup>2+</sup>-affinity purified recombinant P22 from *M. smegmatis* lysate, incubated with 1:500 dilution anti-histidine x 6 POD conjugated antibody (control). Rabbit sera was used at 1:1,000 dilution. Secondary anti-rabbit IgG POD conjugated antibody was used at 1:20,000 dilution. Blots were developed using chemiluminescent detection as described. Arrows indicate the location of P22.
was not present prior to vaccination.

4.4.7.2 Production of rabbit antibody raised to P22 and detection of P22 in cellular fractions

To determine the cellular localisation of native P22 in *M. a. paratuberculosis* strain 316F, rabbit antibody was raised to recombinant P22. As shown in Figure 4.16, the anti-serum recognised recombinant P22 on Western blots. Naïve rabbit serum did not recognise the protein. The serum was used to probe Western blots of equivalent amounts of *M. a. paratuberculosis* strain 316F culture filtrate and soluble and insoluble cell fractions produced by sonication. Results are shown in Figure 4.17 (a). A single band of apparent molecular weight 24.2 kDa was detected in the culture filtrate, which was slightly larger than the molecular weight of 22.3 kDa calculated for the mature native protein based on its amino acid composition. This band was also present in greater intensity in both the soluble and insoluble cell preparations. Several other weak bands of higher molecular weight were produced in the soluble and insoluble cell fractions with the naïve and P22-immunised rabbit sera. To confirm native P22 was migrating slower than expected, it was directly compared to Ni⁺-affinity chromatography prepared P22 produced from *M. smegmatis*. The recombinant protein was previously shown to migrate with an apparent molecular weight of 23 kDa. As shown in Figure 4.17 (b) native P22 migrated slightly slower than recombinant P22.

4.4.8 Species distribution of the p22 gene

To determine if sequences related to the p22 gene existed in other mycobacterial species, nucleic acid database searches, PCR and Southern blot analyses were carried out. Databases searches revealed the gene was present in *M. a. avium* (99% identity) and similar genes existed in *M. tuberculosis/M. bovis* (*lprG/lpp-27*; 75% identity) and *M. leprae* (*lprG*; 74.6%), (see Figure 4.3 for corresponding protein alignments). The similarity between the *lpp-27* and *p22* ORFs was variable depending on the region of the gene. Within the first 210 base pairs of the 5’ end of the ORF, identity was only 67% (see Appendix 4). This was reflected in the lower similarity in the N-terminal region of the protein sequence, noted previously (see Figure 4.3). Much lower identity was found with *M. tuberculosis lprA* (53%), *lprF* (52%) and *lppX* (*p22* of *M. bovis*)(49%). A low level of similarity to *M. leprae lppX* (48%) was also found.
Figure 4.18. PCR amplification of the *p22* gene from 13 isolates of *M. a. paratuberculosis*. PCR reactions were carried out as described (see 4.3.4) using purified DNA as template and primer annealing at 62°C. Samples of 5 μl were electrophoresed in 1% agarose gels, stained with ethidium bromide and photographed under UV light. Lane M, DNA size standard; lanes 1 to 3, New Zealand farmed deer isolates type C1; lanes 4 to 6, New Zealand sheep isolates type S1; lane 7, New Zealand sheep isolate type S5; lanes 8 and 9, South African sheep isolates type I; lane 10, New Zealand cattle isolate type C1; lane 11, Faeroe Islands sheep isolate type S2. The DNA from the five IS900-typed strains were gifted by Dr. Desmond Collins, AgResearch, Wallaceville Animal Research Centre, New Zealand. Lane 12, ATCC 53950; lane 13, strain 316F; lane 14, negative control (water).
Figure 4.19. PCR amplification from 22 mycobacterial strains using primers designed to the \( p22 \) ORF. PCR reactions were carried out using primers lpp-27fBam and lpp27-rKpn at two annealing temperatures. Samples of 5 \( \mu \)l were electrophoresed on 1\% agarose gels, stained with ethidium bromide and photographed under UV light. Simultaneous amplification of the 16S rRNA gene, using primers 246 and 264 was carried out for each sample as a positive control (a single representative is shown). (a) 55\(^\circ\)C annealing temperature. (b) 62\(^\circ\)C annealing temperature. Lane M, DNA size standard; lane 1, \( M. \) \( a. \) \( paratuberculosis \) ATCC 53950; lane 2, \( M. \) \( a. \) \( paratuberculosis \) 316F; lane 3, \( M. \) \( intracellularae \); lane 4, \( M. \) \( scrofulaceum \); lane 5, \( M. \) \( fortuitum \); lane 6, \( M. \) \( terrae \); lane 7, \( M. \) \( phlei \); lane 8, \( M. \) \( smegmatis \); lane 9, \( M. \) \( kansasii \); lane 10, \( M. \) \( gordonae \); lane 11, \( M. \) \( marinum \); lane 12, \( M. \) \( bovis \) 35746; lane 13, \( M. \) \( bovis \) 19210; lane 14, \( M. \) \( bovis \) KML; lane 15, \( M. \) \( bovis \) 35725; lane 16, \( M. \) \( bovis \) 35726; lane 17, \( M. \) \( bovis \) canine isolate; lane 18, BCG (Pasteur); lane 19, BCG (Glaxo); lane 20, BCG (Japan); lane 21, \( M. \) \( tuberculosis \) \( H_37\)Ra; lane 22, \( M. \) \( tuberculosis \) fur seal isolate; lane 23, negative control (water); lane 24, representative positive control 16S rRNA from \( M. \) \( a. \) \( paratuberculosis \) ATCC 53950. The 725 base pair \( p22 \) PCR product is indicated by the white arrow. The red arrow shows a weak 725 base pair product produced in the \( M. \) \( tuberculosis \) complex at 55\(^\circ\)C.
4.4.8.1 PCR amplification of the \textit{p22} gene from \textit{Mycobacterium} species and strains

To determine if \textit{p22} was present in other \textit{M. a. paratuberculosis} strains, several isolates representing five IS900 RFLP types (Collins et al., 1990) were used in PCR analyses. The results are shown in Figure 4.18. The expected 725 base pair \textit{p22} product was amplified from all 13 isolates.

To investigate the species distribution of \textit{p22} within the genus \textit{Mycobacterium}, PCR reactions were carried out with a limited number of mycobacterial species (see Table 2.1), including two strains of \textit{M. a. paratuberculosis} and 11 strains from the \textit{M. tuberculosis} complex. Primers lpp27-fBam and lpp27-rKpn were used (see 4.3.1). Amplification of the genus-specific 16S rRNA gene (see Table 3.1 for primers) was carried out in parallel for all DNA species and served as a positive control for the reactions. Annealing temperatures of 55°C and 62°C were employed (the calculated Tm of the primers was 66°C in 50 mM NaCl). Results are shown in Figure 4.19. PCR products of the expected size for \textit{p22} (725 base pairs) were amplified from both of the \textit{M. a. paratuberculosis} isolates as well as \textit{M. scrofulaceum} and \textit{M. intracellularae} using both annealing temperatures. The only other product of the expected size for \textit{p22} was produced at 55°C, in the \textit{M. tuberculosis} complex strains (Figure 4.19 b; lanes 12 - 22), however this product was very weak and was not produced at 62°C. A major band at 250 base pairs was produced with \textit{M. terrae} at 55°C, however this product was not produced at 62°C. Various weak products were amplified from several of the other species, but none of these were produced with consistency at both annealing temperatures.

Partial sequencing of the PCR products from \textit{M. intracellularae} and \textit{M. scrofulaceum} using the primers lpp27-fBam and lpp27-rKpn, revealed that they had strong identity to \textit{p22}. The identity for \textit{M. intracellularae} was 88.5\% over 654 base pairs beginning 14 base pairs from the ATG start in \textit{p22}. The translated partial sequence was 87.6\% identical. In \textit{M. scrofulaceum} the nucleic acid identity over a similar region was 87.7\% and the translated sequence was 93.8\% identical. The genes in \textit{M. intracellularae} and \textit{M. scrofulaceum} were not identical and the partial translated products shared 85.3\% identity over 211 amino acids. The nucleic acid and translated alignments are in Appendix 7.
Figure 4.20. Southern blot analyses using a p22 probe from genomic DNA of 13 mycobacterial strains. Approximately 1 μg of each DNA was digested with Bam HI, electrophoresed and transferred to nylon membranes as described. Probe DNA was labelled by incorporation of DIG-dUTP during PCR of the p22 ORF. The blot was developed with CDP-star and exposed to film for 2 h. Lane 1, *M. bovis* BCG (Pasteur); lane 2, *M. tuberculosis* H₃₇Ra; lane 3, *M. bovis* (KML); lane 4, *M. intracellulare*; lane 5, *M. a. paratuberculosis* 316F; lane 6, *M. scrofulaceum*; lane 7, *M. gordonae*; lane 8, *M. kansasii*; lane 9, *M. phlei*; lane 10, *M. marinum*; lane 11, *M. terrae*; lane 12, *M. fortuitum*; lane 13, *M. a. paratuberculosis* ATCC 53950.
4.4.8.2 Southern blot detection of the \textit{p22} gene in \textit{Mycobacterium} species and strains

Southern blot analyses were performed with \textit{Bam} HI-digested genomic DNAs from 13 \textit{Mycobacterium} spp. The \textit{p22} ORF was used as probe and was labeled by incorporation of DIG-dUTP during PCR using the primers lpp27-fBam and lpp27-rKpn. Figure 4.20 shows the probe hybridised strongly to a single band, approximately 2,200 base pairs in length from \textit{M. a. paratuberculosis} 316F and a slightly larger band of 2,300 base pairs from \textit{M. a. paratuberculosis} ATCC 53950. The probe also hybridised weakly to a 1,000 base pair band from \textit{M. intracellularae}. Upon overnight exposure (data not shown), a weak band at approximately 10,000 base pairs was present in \textit{M. marinum}, \textit{M. terrae}, \textit{M. phlei} and \textit{M. kansasii}. Hybridisation was not detected with \textit{M. bovis}, \textit{M. tuberculosis}, or \textit{M. bovis} BCG, even though database searches revealed genes similar to \textit{p22} were present in these species.
4.5 Discussion

An *M. a. paratuberculosis* gene encoding an immunogenic exported protein of an approximate molecular mass of 22 kDa, was cloned, sequenced and expressed in *M. smegmatis* as a histidine-tagged protein. Database searches revealed P22 was similar to members of the mycobacterial LppX/LprAFG family of lipoproteins identified in *M. tuberculosis*, *M. bovis* and *M. leprae*. All contain signal sequences for export and potential lipid attachment sites (lipobox), (von Heijne, 1989). Other conserved functional elements contained in the PROSITE database (ExPASy Molecular Biology Server) were not found in the amino acid sequence of P22. Consequently, its function, like other members of this lipoprotein family, is not known. P22 appears to be constitutively expressed. Other lipoproteins have functions including uptake of nutrients, resistance to antibiotics, secretion of proteins, binding and uptake of substrates such as carbohydrates for import, and targeting to host tissues (reviewed by Sutcliffe & Russell, 1995; Navarre & Schneewind, 1999; Tjalsma et al., 1999). Other characterised Gram-positive lipoproteins have functions in cell wall biogenesis and degradation (Sutcliffe & Russell, 1995; Tjalsma et al., 1999). Database comparisons of the nucleotide and amino acid sequences of P22 did not show any significant similarities with any Gram-negative or other Gram-positive known genes or proteins that might have given more clues as to its function.

As P22 is predicted to be a lipoprotein it was postulated to be envelope-associated. To examine the cellular location of P22, culture supernatants and soluble and insoluble cell fractions after sonication were collected and used in Western blot analyses. P22 was found in all three cellular fractions when produced in *M. smegmatis* (Figures 4.7 b and 4.9 b) and *M. a. paratuberculosis* (Figure 4.17 a). This was consistent with a weakly associated envelope protein and is in agreement with other studies of the subcellular distribution of lipoproteins (Andersen et al., 1991a; Florio et al., 1997; Orme et al., 1993; Wiker et al., 1991). An example is the related 22 kDa antigen of *M. bovis* BCG. Although this protein was localised to the cell surface using flow-cytometry, it could be found in cell wall, membrane and culture filtrate fractions in Western blot analyses (Lefevre et al., 2000). Many envelope-associated proteins appear to be spontaneously released when cells are lysed (Pugsley, 1993). For example, the 34 kDa protein of *M. a. paratuberculosis* was originally isolated from soluble cell lysates but has since been localised to the cell envelope using electron microscopy with immunogold labeling (De Kesel et al., 1993). A putative exported *M. a. paratuberculosis* serine protease containing a signal sequence was also isolated in its
mature form from cell lysates (Cameron et al., 1994). In addition, a putative lipoprotein from *M. kansasii* has been found in soluble sonic extracts of *M. a. paratuberculosis* (Armoo et al., 1995). In the present study, mature P22 was also isolated from soluble sonic extracts. The proportionately larger amount of P22 present in the sonicated cell fractions of *M. a. paratuberculosis* compared to culture filtrate suggested the protein was not actively secreted to the extracellular environment. Previous assays in this lab for the activity of isocitrate dehydrogenase, an indicator of cell lysis, have shown that lysis of cells in *M. a. paratuberculosis* cultures through to stationary phase appeared to be minimal. As the cultures used in this study were mid-log phase, release of proteins due to lysis was unlikely. The presence of P22 in the culture filtrate was probably due to release from the envelope during growth. A more accurate assessment of the subcellular localisation of P22 could be achieved with immunogold labeling and electron microscopy, possibly employing the rabbit antibody raised to P22.

There was a small discrepancy in size between the molecular weight of mature recombinant P22 produced in *M. smegmatis* (23.6 kDa) based on its amino acid composition compared to that estimated from SDS-PAGE analysis (23 kDa). The marginally smaller size estimate may not be significant as SDS-PAGE molecular weight determinations are considered to be imprecise, owing to inaccuracies in measurements or aberrant migration of proteins. Aberrant migration has been reported for proteins with a high proportion of basic amino acids (Takano et al., 1988). Possession of a histidine tag may cause this. N-terminal sequencing of recombinant P22 revealed that the protein appeared to be cleaved at an alternative site from that predicted for lipoproteins. Alternative processing of lipoproteins has been previously reported for *Bacillus subtilis*, in which case, cleavage at a different site took place in the absence of the signal peptidase II (Tjalsma et al., 1999). It is possible that the lipoprotein cleavage site of P22 was not recognised by the signal peptidase II of *M. smegmatis*, resulting in alternative processing by a signal peptidase I in this host.

There was also a size discrepancy noted for native P22 from *M. a. paratuberculosis*. Mature native P22 was calculated from its amino acid composition to have a molecular weight of 22.3 kDa, however the apparent size on SDS-PAGE analysis was 24.2 kDa. This slower mobility on SDS-PAGE was confirmed by a side-by-side comparison with recombinant P22, which migrated with an apparent molecular weight of 23 kDa (Figure 4.17 b). Reduced mobility in SDS-PAGE gels has been reported for the closely related 27 kDa antigen of *M. bovis* (Bigi et al., 1997) and also for MK35, a predicted lipoprotein from *M. kansasii* (Armoo et al., 1995). Size discrepancies on Western blots are not unusual and other researchers have
considered a size difference of less than 2 kDa to be not significant (Plum & Clark-Curtiss, 1994). However, differential migration on SDS-PAGE gels has been employed to assay for the presence or absence of lipid modification of proteins (Sankaran et al., 1995). The apparent size difference may therefore reflect differences in processing of P22 between *M. a. paratuberculosis* and *M. smegmatis*. The processing predicted to occur for P22, based on the presence of the lipobox consensus in the signal sequence, would be the transfer of a diacylglycerol moiety to the cysteine residue in the lipobox consensus, followed by signal peptidase II cleavage immediately upstream of the modified cysteine residue, and final transfer of a fatty acid to the resulting N-terminal cysteine (Sankaran & Wu, 1994). This processing may result in a mature P22 lipoprotein of a size larger than that calculated for its amino acid composition. Although the presence of the lipobox consensus remains the hallmark of prediction of bacterial lipoproteins (Sankaran & Wu, 1994), more evidence would be required to confirm lipid modification. The use of metabolic labeling employing radiolabeled fatty acids that are transferred to the protein during processing in the host is a commonly used method for confirmation of lipidation (Navarre & Schneewind, 1999). Relatively few putative lipoproteins have undergone such analysis and this was not attempted for P22. N-terminal sequencing of native P22 may have provided some clues as to its form, as well, an accurate assessment of molecular weight by mass spectroscopy may have determined if post-translational modification had occurred. Such investigations would be worthwhile for better characterisation of the protein from its native host. Purification of P22 from culture filtrates for such studies might be accomplished by affinity purification using the rabbit antibody raised to P22 produced in this study.

To evaluate the relatedness of the *p22* gene to genes in other mycobacterial species, PCR and Southern blot analyses were carried out on a limited number of available strains of mycobacteria. Database searches using the *p22* ORF were also done. The ORF was found to be 99% identical to that in the *M. a. avium* TIGR database. This similarity was not surprising as genetically, *M. a. avium* and *M. a. paratuberculosis* are reported to be nearly identical (Hurley *et al.*, 1988). The major genetic difference that has been established between *M. a. paratuberculosis* and *M. a. avium* is the presence of multiple IS900 insertion elements (reviewed by Hermon-Taylor & Bull, 2002). The related lprG/lpp-27 genes of *M. tuberculosis*, *M. bovis* and *M. leprae* were also identified from database searches. Southern blot analysis using the *p22* ORF as a probe did not show any hybridisation with the *M. tuberculosis* or *M. bovis* strains used, even on extended exposure (Figure 4.20). Experimentally, more than 85% identity is considered efficient in the detection of a gene by hybridisation (Yamaguchi *et al.*, 1989), but genes sharing identities as low as 65% can be
detected by varying the conditions used (Sambrook & Russell, 2001). As p22 of M. a. paratuberculosis and lpp-27/lprG of M. tuberculosis/M. bovis share only 75% identity, the stringency conditions used in these experiments may have been too restrictive to allow hybridisation of the p22 probe.

Genes sharing low identities are more frequently isolated by low-stringency PCR (Sambrook & Russell, 2001). Consistent with this, PCR analysis appeared to be sensitive in detecting genes similar to p22 in M. intracellularae and M. scrofulaceum (Figure 4.19). Both are members of the so-called "MAIS" complex, a group of closely related mycobacteria that also includes M. a. paratuberculosis, M. a. silvaticum and M. a. avium (Wayne, 1984). Partial sequencing of these PCR products revealed the genes were approximately 88% similar to p22 of M. a. paratuberculosis. In light of this, it would be expected that these genes should have readily been detected by Southern blot analysis. Despite repeated hybridisations, only a very weak band, approximately 1,000 base pairs in length, resulted for M. intracellularae on extended exposure. Hybridisation of these same blots with the genus-specific 16S rDNA probe was successful in all of the species and strains tested, including M. intracellularae and M. scrofulaceum (data not shown). There was therefore no reason to suspect degradation of the genomic DNA present on the blots. Taken together, these results suggest both PCR and Southern blot analyses should be used when screening for species distribution of genetic elements.

There was a weak PCR product of a similar size (725 base pairs) to the p22 ORF produced with an annealing temperature of 55°C from strains of the M. tuberculosis complex. It is possible that this product was the similar lpp-27/lprG gene. Hybridisation of p22 probe to Southern blots of this product may have confirmed this. A number of inconsistent weak PCR products were produced in many of the species of mycobacteria examined. Their production was considered to be probably not significant under the low annealing temperatures used (5°C and 12°C below the calculated Tm of the primers). The addition of restriction endonuclease sites at the 5’ end of the primers may also have contributed to non-specific annealing.

Since the immunological cross-reactivity of proteins among different mycobacteria is well known, the presence of genes similar to p22 in other mycobacterial species may have implications for use of this protein as a diagnostic tool. As the genomes of M. a. avium and M. a. paratuberculosis are very similar, it will be difficult to identify useful species-specific proteins. It is however conceivable that specific peptides containing B-cell and T-cell
epitopes from P22 may exist which might allow differentiation of *M. a. paratuberculosis* infection from exposure to *M. intracellulare*, *M. scrofulaceum* and *M. bovis*. Alternatively, these similar genes may be differentially expressed in these species. The 34 kDa protein, which is not specific to *M. a. paratuberculosis*, was found to contain *M. a. paratuberculosis*-specific B-cell epitopes when portions of the protein were evaluated using rabbit anti-sera against whole sonicates of *M. a. avium* and *M. bovis* (De Kesel et al., 1993). However, when antibodies raised to a specific C-terminal peptide were used for immunohistological detection, they bound to macrophages in lesions of a horse infected with *M. a. avium* (Coetsier et al., 1998). This result may indicate screening for specificity using antisera raised to sonicates produced *in vitro* may not reflect the presence of antigen *in vivo*. Antigen p35 of *M. a. paratuberculosis* was found to cross-react with anti-*M. intracellulare* rabbit serum but not anti-*M. a. avium* serum, even though a p35 DNA probe hybridised to four *M. a. avium* strains tested (El-Zaatari et al., 1997). Antigens that are produced in greater amounts in *M. a. paratuberculosis* with respect to other mycobacteria may be useful differential diagnostic reagents. AhpC of *M. a. paratuberculosis*, which differs in only four amino acids in *M. a. avium* and shares 90% identity with *M. tuberculosis*, was shown to be highly expressed only in *M. a. paratuberculosis* cultures (Olsen, 2000a). Antisera raised to *M. a. paratuberculosis* AhpC reacted with extracts from all *M. a. paratuberculosis* isolates tested but did not react with *M. a. avium* or *M. tuberculosis*. Specifically elevated antibody responses to AhpC were found in *M. a. paratuberculosis*-infected mice but not in *M. a. avium*-infected mice, which indicated this protein was also highly expressed *in vivo* by *M. a. paratuberculosis* (Elsaghier et al., 1992). Another mycobacterial protein, MPB70 is highly expressed in *M. bovis* strains (>99%), but not in strains of *M. tuberculosis*, even though the gene is present in this species (Liebana et al., 1996). This remains a useful marker of distinction between *M. bovis* and *M. tuberculosis* (Cornet et al., 1988). Although not presented in this thesis, preliminary Western blot analysis of cell extracts of *M. intracellulare* and *M. scrofulaceum* using anti-P22 rabbit serum showed only a very weak band in these extracts compared to the strong band seen with extracts from *M. a. paratuberculosis* strain 316F. More detailed studies are needed to confirm these results and interpret their meaning. An examination of the expression of P22 and the P22-related proteins *in vitro* and *in vivo*, by mRNA studies for example, may therefore be more informative.

PCR amplification of the p22 ORF indicated that it was present in all of the cattle, sheep and deer isolates of *M. a. paratuberculosis* made available for this study (Figure 4.19). Given the varied geographical sources of these isolates, this suggested p22 is probably highly conserved in *M. a. paratuberculosis* strains. Southern blot analysis of Bam HI-digested
DNA detected hybridisation of the p22 probe to a fragment approximately 100 base pairs larger in the Neoparasec vaccine strain (strain 316F) compared to the *M. a. paratuberculosis* field isolate. It appeared this was due to differences in the sequences flanking the p22 ORF, as PCR of the p22 ORF did not show any size discrepancies. This may reflect a base change(s), resulting in an addition or removal of a *Bam* HI site, or a larger insertion(s) or deletion(s) in the flanking region(s). As genetic differences between the attenuated vaccine strain and virulent field isolates have not been previously reported it would be of interest to determine the nature of this difference.

Two of the proteins that were found to be similar to P22 of *M. a. paratuberculosis* have been investigated by other workers. The 22 kDa antigen (LppX) of *M. bovis* was shown to produce antibody responses but only low IFN-γ responses in a tuberculosis mouse model. Use of this antigen in a DNA vaccine did not confer protection upon challenge of mice with *M. tuberculosi*s (Lefevre et al., 2000). The 27 kDa antigen (Lpp-27/LprG/P27) of *M. bovis* was identified from an *M. bovis* genomic library constructed in λZAP (Bigi et al., 1997). When expressed in *E. coli*, this clone reacted on Western blot analysis with a pool of sera from *M. bovis*-infected cattle, resulting in a major band at 27 kDa as well as several other weaker bands. A size discrepancy between the theoretical size of the mature protein (21.7 kDa) and the observed size on SDS-PAGE (27 kDa) was noted for P27. The authors postulate that this may have been due to post-translational modification. They concluded that the protein was an integral membrane protein in *M. bovis*, as it was found in the detergent phase on extraction of the insoluble pellet after sonication. However, since most lipoproteins would be expected to partition to this phase, this cannot be assumed with any certainty. P27 was not detected in *M. bovis* culture supernatants, but these were subjected to trichloroacetic acid precipitation followed by resuspension to a final 20 fold concentration. In comparison, the current study used filter-concentrated culture supernatants of approximately 200 fold concentration. Although Bigi and co-workers used cultures of various ages to determine this, it was not stated how much of this preparation was used in attempts to detect the protein. Species distribution of p27 using PCR analysis showed a weak product of a similar size in *M. a. avium* and *M. a. paratuberculosis* however, none of these products were sequenced (Bigi et al., 1997). A product of a similar size to p27 was also found in *M. vaccae*, which indicates that a gene similar to p22 may exist in this species. In the study of Bigi and co-workers, PCR products were not found with *M. marinum, M. gordonae, M. aureum* or *M. scrofulaceum*. Anti-P27 rabbit sera revealed P27 was present in whole sonicates of *M. tuberculosis, M. bovis* and *M. bovis* BCG but not *M. smegmatis* or *M. a. paratuberculosis*. Other species were not examined. It would be of interest to examine if the rabbit antibody
raised to P22 produced in this study, cross-reacts with extracts from members of the *M. tuberculosis* complex. In the study of Bigi and co-workers, IFN-γ induction was shown in the three *M. bovis*-infected cattle used in the study. The *E. coli* extracts used for these studies were derived from a clone that contained not only *p27*, but also another hypothetical gene. It cannot be assumed therefore that P27 was responsible for the immunological activity observed.

In the current study, a preliminary evaluation of the capacity of P22 to induce humoral and cellular responses in a relevant host species was obtained by screening sera (immunoblots) and whole blood (IFN-γ assay) from Neoparasex-vaccinated sheep. Western blot analysis using recombinant P22 resulted in prominent bands with 10 of the 11 sera from sheep after vaccination (Figure 4.10). This demonstrated that antibody to P22 was produced as a result of vaccination. There was a faint band produced with two of the naïve sheep sera and with three of the pre-vaccination sheep sera. As these animals were kept under normal farming conditions, it is possible that contact with other environmental bacteria may have resulted in some cross-reactive antibodies to P22. Antibody to P22 was also produced in sheep vaccinated with *M. a. paratuberculosis* strain 316F culture filtrate. This indicates P22 is capable of eliciting immune responses in sheep when presented in the form of a soluble protein in adjuvant.

To investigate if antibody to P22 was present in animals with naturally acquired *M. a. paratuberculosis* infection, sera collected from infected sheep and cattle were used. Fourteen sheep from a flock in which Johne’s disease was confirmed to be present were tested. Five sheep had detectable antibody to recombinant P22 on Western blot analysis (Figure 4.11). Of the five sheep that were diagnosed with Johne’s disease on histological examination, only two had antibody to P22. It is often stated that specific antibody levels increase with progression of disease (Clarke & Little, 1996; Perez *et al.*, 1997; Waters *et al.*, 1999). However, studies on which this is based have been with responses to crude PPD antigens, and not individual proteins. A recent study investigating humoral immune responses of *M. a. paratuberculosis*-infected cattle to Johnin PPD, LAM and recombinant Hsp65 and Hsp70 heat shock proteins of *M. a. paratuberculosis* found that this pattern of increasing antibody existed only for Johnin PPD and only for IgG1. IgG2 responses were not different between clinically affected animals and asymptomatic shedders (Koets *et al.*, 2001). For the individual antigens tested, there was no increase in specific antibody in clinically diseased cattle compared to subclinical animals. In fact, IgG2 levels were significantly lower for Hsp70 in cattle with clinical disease. As IgG2 isotypes are associated with a Th1 immune
response (Abou-Zeid et al., 1997; Koets et al., 2001), this observation may be related to the waning of cell-mediated immune responses in the later stages of the disease. These results demonstrate that antibody responses may vary depending on the properties of the antigen as well as the stage of disease. It is possible that the failure of P22 to detect three of the clinically diseased sheep may have been due to decreased specific antibody production to this particular protein in later stages of the disease. There is also likely to be variation between animals in responses to antigens. Other studies in naturally infected sheep have shown that a significant proportion of clinically affected sheep with well-developed histologic lesions have yielded negative results in serological tests (Clarke & Little, 1996). In the present study it was noted that antibody to P22 was present in three sheep with no gross or histological evidence of Johne’s disease. As these sheep were in contact with known infected flock-mates, it is tempting to speculate that P22 detected subclinical infection in these animals. Alternatively, this may have been the result of immunologic cross-reactions due to exposure to other mycobacteria. Detailed microbiological, histological and immunological studies would have needed to be undertaken to confirm the status of these sheep.

Thirteen dairy cattle with confirmed subclinical *M. a. paratuberculosis* infections were tested for antibody responses to recombinant P22. Six cows that were negative on all ELISA and faecal culture tests and two cows with obvious clinical signs were also used. Eleven of the confirmed subclinical cows and both clinically affected cows had detectable antibody to P22 (Figure 4.12). One of the six cows with uniform negative tests had a weak response to P22. As the negative cows used in this study were from known infected herds, it is possible that this animal may have been in an undetected pre-clinical stage of infection. Alternatively, it may have had cross-reactive antibody to P22 from exposure to other bacteria. Subclinical animals showed a variety of strong and weak responses to P22. It was noted that three of the four cows that were positive on all ELISA and faecal culture tests also had strong antibody responses to P22. The other two cows that exhibited strong antibody responses to P22 had one negative faecal culture and at least two negative ELISA tests each. Because of the small number of animals tested, it was not possible to discern any definite pattern relating the degree of antibody response to P22 and ELISA and faecal culture test results. Of interest was the observation that four of the subclinical cows that were negative on all three ELISA tests had detectable antibody to P22. One of these cows had a very strong response to P22. This preliminary study indicated that P22 may be a candidate for future investigation as a serodiagnostic reagent for the detection of Johne’s disease. A wider screening of both uninfected and infected cattle and sheep is needed to assess the specificity and sensitivity of P22.
Serological tests have the advantage of simplicity and produce rapid results. Although few defined proteins from *M. a. paratuberculosis* have been investigated for this purpose, some preliminary studies, using small numbers of animals have indicated that specific proteins are capable of correctly diagnosing *M. a. paratuberculosis* infection and may allow increased specificities compared to commercial ELISA tests. For example, recombinant p35 produced in *E. coli* was used in a Western blot immunodetection assay and results were compared to four other available tests, including two commercial ELISAs, an agarose gel immunodiffusion test and a complement fixation test (El-Zaatari *et al.*, 1997). The p35 protein was recognised by sera from all of the 12 cattle, two goats and two sheep with advanced Johne’s disease, and by sera from 15 of 20 cattle tested with subclinical disease (confirmed by faecal culture and tissue sample). The sensitivity using p35 was greater than that of the compared tests for the same samples (El-Zaatari *et al.*, 1997). Antigen p35 was not recognised by *M. bovis* BCG-infected cattle sera or sera from the 15 cows without detectable *M. a. paratuberculosis* infection. It did however, cross-react with anti-*M. intracellularae* rabbit serum. Another antigen evaluated for use in a serological test is a C-terminal peptide from the 34 kDa protein (a362) reported to have specific B-cell epitopes with respect to other mycobacteria (De Kesel *et al.*, 1993). Sonic extracts of *E. coli* producing the recombinant polypeptide was used in an ELISA, and test sera were pre-adsorbed to sonic extracts of non-recombinant *E. coli* before being applied. The assay correctly diagnosed all 25 infected cattle at different phases of the disease and yielded negative results for the seven healthy cattle tested (De Kesel *et al.*, 1993). Examination of cross-reactivity in a wider range of environmental organisms found cross-reactivity with *Corynebacterium xerosis* (De Kesel *et al.*, 1993). When purified recombinant a362 polypeptide was used in an ELISA with 30 reference sera from *M. a. paratuberculosis*-infected cows, it detected 70% of these (Vannuffel *et al.*, 1994). This was somewhat lower than the sensitivities for several commercial ELISA tests (77-87%) that were run in parallel with the same sera. The reported specificity was 95% with respect to healthy cattle (n=175) and tuberculous cattle (n=38). AhpC, an intracellular protein isolated from *M. a. paratuberculosis* soluble sonicates has been used in Western blotting analyses (Olsen, 2000a). Antibody to AhpC was detected in four experimentally-infected, asymptomatic goats.

Development of accurate tests for the detection of Johne’s disease at all stages of infection would greatly assist in eradication programmes. Given the present knowledge that the early immune response to *M. a. paratuberculosis* infection appears to be cell-mediated, it may be more advantageous to use a reagent capable of eliciting IFN-γ secretion. Cellular immune
responses of the vaccinated and unvaccinated sheep to recombinant P22 were evaluated using a whole blood IFN-\( \gamma \) assay. P22 induced a significant level of IFN-\( \gamma \) in the vaccinated sheep groups compared to unvaccinated control groups (Figures 4.13 and 4.14). This response was generally dose-dependent. Purified P22 produced significant IFN-\( \gamma \) responses in the vaccinated sheep group when used in amounts as low as 1 \( \mu \)g/ml. Although the Ni\(^{2+}\)-affinity-enriched preparation appeared to elicit equivalent amounts of IFN-\( \gamma \) secretion with 0.32 \( \mu \)g, it was difficult to compare these two results. Variations in the IFN-\( \gamma \) responses of the vaccinated sheep to Johnin or Avian PPDs often occurred over different time intervals. It is also possible that the small amount of \( M. \) \textit{smegmatis} components present in the Ni\(^{2+}\)-affinity-enriched preparation was in part responsible for the larger IFN-\( \gamma \) response using this preparation. To investigate this possibility, the two preparations would need to be tested in parallel. Due to the high cost of the IFN-\( \gamma \) EIA assay and the limiting amounts of P22 available, this was not done.

On one occasion, three of the five unvaccinated control animals showed IFN-\( \gamma \) responses to Avian PPD, one of these (569) was notably high (Figure 4.13). Conversely, none of these animals reacted to the P22 preparation. One unvaccinated control animal (133) showed a low response to purified P22 on a separate occasion (Figure 4.14). This may suggest an overall higher specificity for P22 antigen compared to Avian PPD. Notable responses of the control animals to Avian and Johnin PPD had been previously observed on several occasions during the time of this study. This emphasized the importance of using a sufficient number of animals in each group for comparison. Since infection of domestic hoofstock with environmental mycobacteria, such as \( M. \) \textit{a. avium} and \( M. \) \textit{intracellularae} is considered to be uncommon (Thorel \textit{et al.}, 2001), responses may have been due to transient sensitisation as the animals were kept on open pasture. Other researchers have reported responses to PPD antigens in uninfected cattle (de Lisle \& Duncan, 1981; Hintz, 1981, McDonald \textit{et al.}, 1999: Olsen \& Storset, 2001). These reactions have been thought to be due to cross-reactions with other environmental organisms (Buergelt \textit{et al.}, 1977; Gilot \& Misonne, 1994) and others have attributed these reactions to innate immune responses to PPD antigens, especially in young cattle (Olsen \& Storset, 2001). This is an illustration of the poor specificity reported for these antigens in cellular tests (Buergelt \textit{et al.}, 1977; Chiodini \textit{et al.}, 1984a; de Lisle \& Duncan, 1981; Riemann \& Abbas, 1983).

The preliminary results described here suggest that further studies of P22 are warranted to investigate its usefulness as an immunodiagnostic reagent for the detection of \( M. \) \textit{a.}
paratuberculosis infection. Furthermore, its ability to induce cell-mediated immune responses indicates that this antigen may be a candidate for inclusion in a subunit vaccine against Johne’s disease. These studies illustrate the potential value of characterising new antigens from *M. a. paratuberculosis*.
General discussion and conclusions

An important first step towards the design of improved Johne's disease vaccines and diagnostic reagents is the identification and characterisation of the bacterial components involved in the host immune response. The aim of the work presented was to identify exported proteins of *M. a. paratuberculosis* that elicit IFN-γ secretion or antibody production in vaccinated or infected animals. Only a small number of individual proteins, approximately 24, have previously been isolated from *M. a. paratuberculosis* and only a portion of these have been studied for their immunological activity. *M. a. paratuberculosis* does not readily lend itself to the study of immunogenic proteins due to its fastidious growth requirements, pathogenic nature and slow growth. In this study, the novel use of PhoA fusion proteins to identify exported proteins from *M. a. paratuberculosis* has resulted in the isolation of partial gene sequences from a large number of previously unidentified proteins from this organism. This study focussed on one of the putative exported proteins identified from the library, a probable lipoprotein named P22.

The advantage of PhoA technology is the rapid selection of putative exported proteins and the direct identification of the corresponding gene sequences. Information from databases facilitated the cloning of the entire p22 ORF for the expression and purification of sufficient amounts of protein for preliminary characterisation. There are a number of technical restrictions associated with the use of the pJEM11 system. These include the necessary cloning of gene sequences in the correct frame for translation of the fusion proteins, which in turn is dependent on the number and location of Sau3A restriction endonuclease sites in the genomic DNA. Improved plasmids, allowing translation in all three reading frames would facilitate the isolation of a greater number of potential exported proteins. Also, the system relies on the ability of *M. smegmatis* or *E. coli* to recognise *M. a. paratuberculosis* promoter elements and signal sequences. Observations from this study indicated that some promoter elements from *M. a. paratuberculosis* may not have been recognised. For example, the pJEM11 construct (Eco-3) containing a protein similar to MK35 of *M. kansasii* was expressed in *E. coli* but not when it was transferred to *M. smegmatis*. The appearance of proportionately more PhoA+ colonies expressed in *M. smegmatis* than in *E. coli* indicates *M. a. paratuberculosis* promoters were not widely recognised by *E. coli*. It would be interesting to characterise the promoters of the *M. a. paratuberculosis* gene sequences cloned as there have been very few studies done regarding gene expression for this organism.
There were several proteins identified in this study that did not appear to contain signal sequences. This may have been due to the presence of segments such as transmembrane regions that allowed insertion of the PhoA reporter protein through the membrane. Despite the short-comings of the pJEM11 system, eleven putative exported proteins not previously described for *M. a. paratuberculosis* were identified, demonstrating the value of PhoA technology for the identification of exported proteins from *M. a. paratuberculosis*.

P22 was found to be related to a family of unique exported mycobacterial lipoproteins, all of which have unknown functions. P22 possessed a signal sequence and was found in culture supernatants and sonicated preparations of *M. a. paratuberculosis*, which is consistent with an exported lipoprotein. Structural and biochemical studies of the protein and generation of *M. a. paratuberculosis* p22 knock-out mutants may give an indication as to its function. Gene expression studies could be conducted, for example using RT-PCR, to investigate what conditions affect expression of *p22 in vitro and in vivo*.

P22 was expressed as a C-terminal histidine-tagged recombinant protein using the vector pMIP12 in the surrogate host *M. smegmatis*. This allowed highly reproducible purification of the recombinant protein from cell sonicates of *M. smegmatis* using Ni²⁺-affinity chromatography followed by size exclusion chromatography. A limitation of the pMIP12 expression system was the relatively low amount of P22 produced (approximately 0.05 mg per litre of culture). Expression of protein from pMIP12 is governed by the constitutive mycobacterial promoter *pblaF*+. High-expression *E. coli* plasmids, such as the glutathione S-transferase system (pGEX vectors, Pharmacia Biotech Inc.), are capable of producing very large amounts of recombinant protein (up to 10 mg per litre), owing to inducible and highly efficient promoter elements. Several studies with mycobacterial proteins have reported a lack of T-cell stimulation with proteins produced in such *E. coli* systems compared to these same proteins produced in mycobacteria (Abou-Zeid *et al.*, 1997; Roche *et al.*, 1996; Triccas *et al.*, 1996; Rosenkrands *et al.*, 1999). The authors of these papers postulate that incorrect folding, improper intramolecular associations or lack of proper post-translational modifications may be responsible for the lack of activity. In two of these studies, conformational differences were confirmed in the proteins when expressed in different hosts (Triccas *et al.*, 1996; Rosenkrands *et al.*, 1999). Other studies have found no difference in T-cell responses using *E. coli* or native versions of proteins (Colangeli *et al.*, 1998; Sorensen *et al.*, 1995). These conflicting reports led to the decision to express P22 in *M. smegmatis*. Although not presented in this thesis, expression of several genes cloned from this study, including *p22*, resulted in insoluble protein products when expressed from plasmids in *E. coli*, suggesting
improper folding of the proteins. Although *M. smegmatis* is more closely related to *M. a. paratuberculosis* than *E. coli* is, differences in protein production may still occur. Expression of P22 in *M. smegmatis* resulted in the production of immunologically active protein, however the difference in apparent size between recombinant P22 and native P22, indicated differences in post-translational modification of the protein may have occurred between *M. smegmatis* and *M. a. paratuberculosis*. The nature of the size differences and the immunologic activities of recombinant P22 versus native P22 were not investigated due to the limited amounts of P22 available and lack of purified native P22. Immunological comparisons between P22 expressed in *E. coli*, *M. smegmatis* and from the native host would be worthwhile.

Increased amounts of P22 would allow further studies such as investigation of cell-mediated immune responses of naturally infected animals to P22 and immunisation of sheep followed by challenge studies. Western blot analyses indicated that P22 might be useful as a serodiagnostic reagent. Although this type of test is valuable for research purposes, it is impractical for large-scale field studies. Furthermore, Western blots are not easily quantitated. Larger amounts of purified, immunoreactive P22 would also enable a P22-ELISA to be evaluated as a diagnostic tool. P22 could also be tested for use as a reagent in an IFN-γ EIA test.

The presence of *p22* in *M. a. avium* and of similar genes in *M. intracellularae* and *M. scrofulaceum* suggests that immunological cross-reactions with P22 could occur. Although infections of sheep and cattle with these species are uncommon, immunological sensitisation through contact with them in the environment is possible. A relevant approach to evaluate the specificity of P22 would be to screen herds or flocks assessed to be free of *M. a. paratuberculosis* infection. If cross-reactions were found to occur, it may be possible to use epitope mapping of P22 to produce peptide antigens to improve specificity.

P22 was shown to elicit IFN-γ production in blood from vaccinated sheep. Very few other proteins from *M. a. paratuberculosis* have been investigated for their ability to induce IFN-γ responses in infected or vaccinated animals (Gilot et al., 1992; Burrells et al., 1995; Olsen et al., 2000a; Olsen & Storset, 2001). The A36 antigen complex, which contains the 34 kDa protein, among other proteins, was shown to produce T-cell proliferation from the lymph nodes of A36-immunised mice and delayed-type hypersensitivity reactions in *M. a. paratuberculosis*-infected rabbits (Gilot et al., 1992). Antigen P30 was shown to induce IFN-γ responses in whole blood from two experimentally infected sheep (Burrells et al., 1995; Olsen et al., 2000a; Olsen & Storset, 2001).
AhpC and AhpD were both shown to induce IFN-γ responses in whole blood from three experimentally infected goats (Olsen, 2000a). A 14 kDa protein (MMP14) has been reported to induce IFN-γ secretion in eight experimentally infected calves and also induced IFN-γ secretion in several of the uninfected control calves (Olsen & Storset, 2001). The authors postulate this antigen may induce innate IFN-γ secretion, although they were unable to confirm the cell types that were the source of the IFN-γ production. To date, no M. a. paratuberculosis purified antigens have been used as a vaccine.

Although the characteristics of antigens that elicit protective immunity is not known, those proteins that are capable of inducing high levels of IFN-γ secretion have generally been selected for inclusion in novel mycobacterial vaccine preparations. Several proteins have shown promise as subunit vaccines for other mycobacterial diseases in animal models when presented in conjunction with various adjuvants such as DNA and IL-12. Some of these include Ag85A (Baldwin et al., 1998; Huygen et al., 1996; Li et al., 1999; Lozes et al., 1997; Tange et al., 2000), ESAT-6 (Brandt et al., 2000; Li et al., 1999), the 38 kDa protein (Zhu et al., 1997) and the 30 kDa protein (Sharma et al., 1999) of M. tuberculosis, the Hsp65 and the 36 kDa protein of M. leprae (Lowrie et al., 1999; Tascon et al., 1996) and the 35 kDa antigen of M. a. avium (Martin et al., 2000). Culture filtrates have also shown protection against M. tuberculosis (Andersen, 1994; Hubbard et al., 1992; Horwitz et al., 1995; Bosio & Orne, 1998; Lindblad et al., 1997) and M. a. avium infections (Silva et al., 2000). However, high levels of IFN-γ production in vitro do not necessarily correlate to protection in vivo. This has been illustrated in several studies. Immunisation with DNA encoding the 19 kDa lipoprotein (Abou-Zeid et al., 1997), and AhpC protein of M. tuberculosis were both found to stimulate antibody responses rather than a cell-mediated immune response against the proteins in mice and resulted in no protection against subsequent challenge (Erb et al., 1998). In fact, immunisation with the 19 kDa protein worsened the course of experimental infection (Abou-Zeid et al., 1997). In light of this, it would be useful to investigate the profile of other cytokines in addition to IFN-γ that might be induced by P22. For example, IL-12, IL-2 and IL-18, which are associated with cell-mediated immune reactions and cytokines such as IL-4, and IL-5 that are associated with antibody production (reviewed by Jankovic et al., 2001). It would also be of interest to determine the source of the IFN-γ produced by stimulation with P22 and if cytotoxic activity was induced by P22. Studies with lymphocyte populations from M. bovis-infected cows have shown exogenously added proteinaceous antigens can elicit significant amounts of IFN-γ production in CD4+ T-cells and in CD8+ T-cells (Lébana et al., 1999). Both these cell populations have been shown to have important roles in mycobacterial
infections, although the relative contribution to protection from these cell types has not been determined. In view of the evidence gathered through numerous studies to support the role of cell-mediated immune responses in the protection against Johne's disease, it is of great interest to investigate the factors that influence the shift from cell-mediated immune responses to humoral responses during the course of this disease. This information may facilitate the development of vaccines that boost the protective responses and downregulate the responses that serve to initiate disease.

As stated in a recent review article, "Paratuberculosis is now considered a cosmopolitan disease emerging from the shadow of other more dramatic and costly animal diseases," (Tessema et al., 2001). Clearly, more information on immunogenic proteins of *M. a. paratuberculosis* is needed to facilitate the development of new diagnostic tests and vaccines. The studies presented in this thesis have contributed to the general knowledge of *M. a. paratuberculosis*. The genetic approach taken to search for immunogenic proteins by selecting for exported proteins using PhoA technology, has resulted in the identification of an immunogenic 22 kDa protein not previously described for *M. a. paratuberculosis*. 
### Appendix 1

#### Commonly used solutions

**PBS**
- 136 mM NaCl
- 2.6 mM KCl
- 4 mM Na$_2$HPO$_4$
- 1.8 mM KH$_2$PO$_4$
- pH 7.4

**Transfer buffer**
- 25 mM Tris base
- 193 mM glycine
- 20% (v/v) methanol
- 0.015% (w/v) SDS

**Blocking solution**
- 20 mM Tris.Cl (pH 7.4)
- 100 mM NaCl
- 0.1% (v/v) Tween 20
- 5% (w/v) skim milk

**Wash buffer**
- 20 mM Tris.Cl (pH 7.4)
- 100 mM NaCl
- 0.1% (v/v) Tween 20

**TE buffer**
- 10 mM Tris.Cl
- 1 mM EDTA
- pH 8.0

**TBS**
- 10 mM Tris.Cl
- 150 mM NaCl
- pH 7.4

**Luria Bertani (LB) media**
- 1% (w/v) tryptone (Difco, USA)
- 0.5% (w/v) yeast extract (Difco, USA)
- 0.5% (w/v) NaCl
- For plates, add agar to 1.5% (w/v)

**TAE Buffer**
- 40 mM Trisacetate
- 1 mM EDTA
- pH 8.0

**20 x SSC**
- 3 M NaCl
- 300 mM sodium citrate
- pH 7.0

**5 x DNA loading dye**
- 0.1% (w/v) bromophenol blue
- 0.25% (w/v) xylene cyanol
- 30% (w/v) glycerol

**2 x SDS PAGE loading buffer**
- 0.125 M Tris.Cl (pH6.8)
- 4% (w/v) SDS
- 20% (v/v) glycerol
- 10% (v/v) β-mercaptoethanol
- 0.2% (w/v) bromophenol blue

**Start buffer**
- 20 mM Na$_2$HPO$_4$
- 500 mM NaCl
- pH 7.4

**Loading buffer**
- Start buffer plus
- 15 mM imidazole
- pH 7.4

**Modified 7H9 for M. smegmatis**
- 47 g/L Middlebrook 7H9 (Difco, USA)
- 0.2% (v/v) glycerol
- 0.05% (v/v) Tween 80

Prepare broth with 7H9, glycerol and Tween 80. Autoclave. Cool to 55°C. Add a sterile, concentrated solution of glucose to 2 g/L.
Modified 7H9 for *M. a. paratuberculosis*
4.7 g/L Middlebrook 7H9 (Difco, USA)
1 mg/L Mycobactin J (Allied Monitor, USA)
0.2% (v/v) glycerol
1% (v/v) Middlebrook ADC enrichment supplement (Difco, USA)
2 g/L glucose
Prepare broth with 7H9, Mycobactin J and glycerol. Autoclave. Cool to 55°C.
Add a sterile, concentrated solution of glucose to 2 g/L.

7H10 agar
19 g/L Middlebrook 7H10 (Difco, USA)
0.5% (v/v) glycerol
10% (v/v) Middlebrook OADC enrichment supplement (Difco, USA)
Prepare agar with 7H10 and glycerol. Autoclave. Cool to 55°C. Add OADC.
For culture of *M. a. paratuberculosis*, add 1 mg/L Mycobactin J prior to autoclaving.

Sauton’s media
Modified from Sauton (1912)
For 1 L of liquid media
0.5 g K₂HPO₄
0.5 g MgSO₄
4.0 g asparagine
0.05 g ammonium ferric citrate
2.0 g citric acid crystals
1.0 ml glycerol
1.0 ml Tween 80
pH to 6.8-7.2 and autoclave. Add 1.0 ml of a 1% solution of filter-sterilised ZnSO₄.
Appendix 2  Diagnosis of Johne’s disease and results for “Limestone Downs” sheep

The criteria used for the diagnosis of animals for Johne’s disease by necropsy and histology, were based on those published in the Manual of Standards for Diagnostic Tests and Vaccines (Office International des Epizooties, 2000).

The mucosa and serosa of the jejunum and ileum was inspected for evidence of thickening and corrugation. Multiple samples of intestinal wall, ileocaecal valve and mesenteric lymph node were collected into 10% buffered formalin for subsequent staining. Haematoxylin and eosin-stained sections were examined for pathognomonic lesions consisting of diffuse or focal infiltration of the lamina propria, submucosa and the cortex of the mesenteric lymph nodes with large, pale-staining clusters of epitheloid cells and multinucleated Langhans’ giant cells. Diagnosis was positive if all of the above criteria were present. Suspicious foci found on histopathology were stained with Ziehl/Neelsen’s method and examined for the presence of acid-fast bacilli located in clumps or singly within the lesion. Diagnosis was considered positive if this was found, otherwise animals were categorised as suspect. Acid-fast bacilli are usually, but not invariably, found.

Results for the individual animals used in this study are shown below.

<table>
<thead>
<tr>
<th>Sheep #</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>positive, severe lesions</td>
</tr>
<tr>
<td>44</td>
<td>positive, severe lesions</td>
</tr>
<tr>
<td>36</td>
<td>positive, severe lesions</td>
</tr>
<tr>
<td>48</td>
<td>positive, moderate lesions</td>
</tr>
<tr>
<td>23</td>
<td>positive, mild lesions</td>
</tr>
<tr>
<td>26</td>
<td>Suspect (single cluster of epitheloid cells in mesenteric lymph node, but no organisms visible)</td>
</tr>
<tr>
<td>25</td>
<td>negative</td>
</tr>
<tr>
<td>43</td>
<td>“</td>
</tr>
<tr>
<td>13</td>
<td>“</td>
</tr>
<tr>
<td>2</td>
<td>“</td>
</tr>
<tr>
<td>6</td>
<td>“</td>
</tr>
<tr>
<td>51</td>
<td>“</td>
</tr>
<tr>
<td>32</td>
<td>“</td>
</tr>
<tr>
<td>27</td>
<td>not examined</td>
</tr>
</tbody>
</table>
Appendix 3

Most significant protein database alignments obtained
using the translated DNA segments fused to phoA

Identity searches were done through the NCBI BLAST server. Red arrows indicate signal peptide cleavage sites, predicted by SignalP or through experimental evidence. Identities refers to the number of identical amino acids. Positives refers to identities plus similar amino acids.

**Eco-3**

>gi|7478265|pir|F70934 probable lpgN protein - Mycobacterium tuberculosis (strain H37Rv)

Expect = 2e-26, Identities 81/137 (59%), Positives 103/137 (75%)

<table>
<thead>
<tr>
<th>PTB</th>
<th>Sbjct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MKHTLATISVALSLA-LVGCGHEHKSG-CDTSSSTSTSTSTSTSTATSSSTAPAQAQT 59</td>
</tr>
<tr>
<td>60</td>
<td>IADYVKONHISETPVHRCDPG-PNVDLPVPACMG-LNQQGIVGIVQSPAPDPDPEPT 117</td>
</tr>
<tr>
<td>118</td>
<td>VSALFKLIGVDPAK</td>
</tr>
</tbody>
</table>

**Eco-4**

>`gi|1174333|emb|CAA64678.1| (X95394) w-amino-transferase-like protein [Agrobacterium tumefaciens]

Expect = 6.9, Identities 16/33 (48%), Positives = 19/33 (57%)

<table>
<thead>
<tr>
<th>PTB</th>
<th>Sbjct</th>
</tr>
</thead>
<tbody>
<tr>
<td>73</td>
<td>GAAQPPHSAGRRNSSMPVVRAGTVGAAASRR 105</td>
</tr>
<tr>
<td>21</td>
<td>QPPVE-----FQVFSŁQARATRGGLLLTALGGLIIGVLTVTA IPTVGISSGILLAYIDS 74</td>
</tr>
</tbody>
</table>

**P TB -1**

>`emb|CAC29589.1| (AL583917) putative membrane protein [Mycobacterium leprae]

Expect = 1e-22, Identities 59/95 (61%), Positives 70/95 (72%)

<table>
<thead>
<tr>
<th>PTB</th>
<th>Sbjct</th>
</tr>
</thead>
<tbody>
<tr>
<td>81</td>
<td>KPVETDAASFLWPRSRQDL-ARRALLLTALGGLIAGLVTAPAQGSGGRL----LDA 136</td>
</tr>
<tr>
<td>21</td>
<td>QPVVE------FQVFSŁQARATRGGLLLTALGGLIIGVLTVTA IPTVGISSGILLAYIDS 74</td>
</tr>
</tbody>
</table>

**P TB -4**

>`gi|13375559|gb|AAK20393.1|AF334164.1 (AF334164) 34 kDa protein [Mycobacterium avium subsp. paratuberculosis]

Expect = 6e-36, Identities 130/130 (100%), Positives 130/130 (100%)

<table>
<thead>
<tr>
<th>PTB</th>
<th>Sbjct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MTSFGGPSGYPQAQQS GYTAGATPSFAKDDFKSLPLYNIAVVALFAYAAYLNFPT 60</td>
</tr>
<tr>
<td>1</td>
<td>MTSFGGPSGYPQAQQS GYTAGATPSFAKDDFKSLPLYNIAVVALFAYAAYLNFPT 60</td>
</tr>
</tbody>
</table>
>gi|15607544|ref|NP_214917.1| (NC_000962) mmpS1 [Mycobacterium tuberculosis H37Rv] Expect = 7.7, Identities = 23/55 (41%), Positives = 25/55 (44%)

>gi|6225700|sp|P95212|MMS1_MYCTU Putative membrane protein MMS1 Expect = 3e-51, Identities 112/112 (100%), Positives 112/112 (100%)

>gi|8439555|gb|AAF74988.1|AAF139058_1 (AF139058) serine esterase cutinase [Mycobacterium avium] Expect = 4.5, Identities 13/24 (54%), Positives 15/24 (62%)

>gi|13092756|emb|CAC30065.1| (AL583918) putative lipoprotein [Mycobacterium leprae] Expect = 3e-70, Identities 170/178 (95%), Positives 170/178 (95%)

>gi|1817698|emb|CAB6572.1| (284724) SodC [Mycobacterium tuberculosis] Expect = 3e-70, Identities 170/178 (95%), Positives 170/178 (95%)

>gi|1870631|prob|SodC protein - Mycobacterium tuberculosis (strain H37Rv) Expect = 3e-70, Identities 170/178 (95%), Positives 170/178 (95%)
pTB-20
no significant similarities found

pTB-26
no significant similarities found

pTB-73
>gi|560008|gb|AAC77832.1| (U13174) putative basolateral Na-K-2Cl cotransporter [Mus musculus]
Expect = 0.89, Identities = 23/57 (40%), Positives = 30/57 (52%)
pTB: 94 RRRQSQLQGHHXGCVQRQART--QQCHFGGARKNGTSPGCYPARKASC--PQQHHF 147
RRRG H + G++ T G+Q G GG RG GTP++G + GP PQ F
Sbjct: 25RRRGDP--ARHGRASSQEDATTAGRQAG--GGVREEGTPTAAGDGGLGRPLGPTPSQRG 78

Eco-24/pTB-88
>gi|3242295|emb|CAC17756.1| (AL021957) hypothetical protein Rv2190c [Mycobacterium tuberculosis]
Expect = 7e-27, Identities 70/107 (68%), Positives 89/107 (83%)
pTB: 178 RPIAGAJASLALCGELAGSVQualDADDALKINELRSQAETTEAMHS AQQLDELNA 237
RG AS + +G+LA +V ADPADDALKINELRSQAETTEAMHS AQQLDELN A
Sbjct: 15RSGFFPASFTVSSVLANWL ADPADDALKINELRSQAETTEAMHS AQQLDELNA 74

pTB-106
>gi|13092840|emb|CAC30200.1| (AL583919) putative D-alanyl-D-alanine carboxypeptidase [Mycobacterium leprae]
Expect = 1e-74, Identities 156/207 (75%), Positives 178/207 (85%)
pTB: 6 ALVRSASCIALLALFLAAAFAPVSG---TASPLGRAEPQAEPAPAPNCYKVNTPAPAVDSS 63
+ +R S +CLAA F P TA P +G AEP A PNA P CPY +V TPAVDS 61
Sbjct: 2SFLRSTACLAAAVFAIVGPVAALDLPTAVPVG---TASPLGRAEPQAEPAPAPNSCGIVTDAS 61

pTB: 64 VPTACGDPMPVPKCVDPGVDAWLDLAGCQVGAAPNTYPPLGPDVSVEAEMLVLADLSPQVIAAK 123
VP G +FPMPALVP KVPV GL SL SGGVAP++TPP+PGD+SAEVMLVADL+SNGIVIAA+
Sbjct: 62 VPTQUNPMPMPVATPVPQVGAANSCIVOATDPPPVPGWLDAEMLVADLSPQVIAAA 121

pTB: 124 DPHGHRPASITKVL+ANA+INTLNLK+V GTR+DAAEGTVGV +G OG TQNL HLG 183
DPHGRHRPASITKVL+ANA+INTLNLK+V GTR+DAAEGTVGV +G OG TQNL HLG 122
Sbjct: 122DPHGRHRPASITKVL+ANA+INTLNLK+V GTR+DAAEGTVGV +G OG TQNL HLG 181

pTB: 184 LLMHSQGANDAHALAMQLGOMQGALEKI 210
LLM SQNAH+LAMQLGOMQGAL+KI
Sbjct: 182 LLMSQGANDAHALAMQLGOMQGALEKI 208
pTB-107
>gi|13431768|sp|P72001|PKNE_MYCTU probable serine/threonine-protein kinase Pkne Mycobacterium tuberculosis (strain H37Rv)
Expect = 9e-7, Identities 41/93 (43%), Positives 51/93 (55%)

pTB : 66 RYMAPERFG--TDDHARRLLVGLRIHECLTGASRDS-DLLEEQLNAHNTARRGVGTAP 122
+YMAPERF+ T A + LHECL/TG+ R DS + + +HL G P
Sbjct: 180 KYNAPERSFNEVTTYRADYIALVCVLHIETGAPFYRADASAGTLVSSHL----MGPIPOQP 235

pTB : 123 EVRP----FDAVIAMKDPERRYQSVTEL 150
+RP FDAV+ARKMAK PE RY S + L
Sbjct: 236 SAIROPERPFAVVARMAKKPEDRYASAGDL 268

pTB-162
>gi|7478364|pir|B70662 probable membrane protein - Mycobacterium tuberculosis (strain H37Rv)
Expect = 4e-29, Identities 73/153 (47%), Positives 105/153 (67%)

pTB : 1 MRVVRLLFAAVLTLTDLGIVAPAGAAQPPTKTPD/DHI/TDSNRLVT/DSDRAAVSSAIDRLY 60
MR+VRL +TIL AGLL PAGA QPP +L+++TD+ VLT S R AV++A+DRLY
Sbjct: 1 MRVVRLLGMVL/TILAAAGLLGPAGAAQPFPRLSNV/NAGVLTSQGRTAUTAVDRLYA 60

pTB : 61 DRHDLVAVYDYFDNPKNFVWADRTPRESGCMRDURVLAVAAINTKFLTNPPQPLSNETP 120
DR I-LVVYV+NF+ NWA RT S +G+ D LLAVAT + + + +
Sbjct: 61 DRRIRLVVYVENSSQALNWAQRTTRTSELGNYDALLAVATTGREAYFLVSPASSGV 120

pTB : 121 ADEINLSRNKIGPASRAW/ATAAAADDL 153
+++++R + I PA+ +++ AA+AAA+GL+
Sbjct: 121 EGQVDNVRRYQIEPALHDGDYSGAA VAAALN 153

pTB-166
>gi|2829577|sp|P71882|LPPP_MYCTU putative lipoprotein Lppp precursor Mycobacterium tuberculosis (strain H37Rv)
Expect = 3e-06, Identities 27/43 (63%), Positives 33/43 (77%)

pTB : 45 LLIVMVVLALGCGGG/KPPAPFQARPDYCKASKDGEPTD/TQVAI 87
L+ +++ L ASC MKPP + P+TCK SDGTAD/TQVAI
Sbjct: 19 LALIVVLGAAGA/VKPPETRTPSPPPNTCKDSDGTAD/TQVAI 61

pTB-179
sp|O10530|CISY_MYCTU citrate synthase 1 >gi|1314035|emb|CA97402 (Z 73101) gltA2 [Mycobacterium tuberculosis]
Expect = e-102, Identities 188/209 (89%), Positives 195/209 (92%)

pTB : 13 GYPBQLAQEKSTLFYLVSYLLTLYELPDD/KELEFTKQKLNLMTGKDDG/TPQFNNAH 72
GYPI+QLEKSTF+EV YLLTLYELP DQLA+PT RQQ SIGNALKDDG/TPQFNNAH
Sbjct: 73 GYPIFDQLEKSFVLCYLLTLYELPDD/DLAQFTQRIQFTHLMHDKDDG/TQFNNAH 132

pTB : 73 PFMVLSASNVNLASAYPSDLMDETD/EVELSTLMLAKLT/TAYAYKSGVQPFLYPDN 132
PFMVLSSSNVLSAY/SAYD/LD/LDMD VELSTKLLALKLTYA YAYKSGVQPFLYPDN 192

pTB : 133 KLISVENGFLRTTFQPDFLPEPQYQADPEV/RAL/MMLLILILILHDEHQNCSTSTVRLVGSSQANLF 192
L+I/VENFLR+FGDFPQYQAPDFPVEV/RAL/LMLILILILHDEHQNCSTSTVRLVGSSQANLF 252
Sbjct: 193 SLTV/NFG/TLTFQPDFLPEPQYQAPDFPVEV/RAL/LMLILILILHDEHQNCSTSTVRLVGSSQANLF 252
**pTB - 184**

>gi|7476409|pir| |C70549 hypothetical protein Rv0560c - *Mycobacterium tuberculosis* (strain H37Rv) Expect = 3e-39, Identities 87/122 (71%), Positives 96/122 (78%)

**pTB - 187**

>gi|8928200|sp|Q9X781|LYTBJCLE lytB protein homologue *M. leprae* Expect = 5e-22, Identities 51/61 (83%), Positives 55/61 (89%)

**pTB - 190**

>gi|1709567|sp|O10505|PANB_MYCTU 3-methly-2-oxobutanoate hydroxymethyltransferase (ketopantoate hydromethyltransferase) *Mycobacterium tuberculosis* (strain H37Rv) Expect = 8e-22, Identities 84/169 (49%), Positives 91/169 (53%)

**pTB - 191**

>gi|13633834|sp|Q9CBK1|RL11_MYCLE 50S ribosomal protein L11

gi|13093572|emb|CAC30859.1| (AL583923) 50S ribosomal protein L11 [*Mycobacterium leprae*] Expect = 1e-57, Identities 122/129 (94%), Positives 123/129 (95%)

**pTB - 192**

>gi|13633834|sp|Q9CBK1|RL11_MYCLE 50S ribosomal protein L11

>gi|13093572|emb|CAC30859.1| (AL583923) 50S ribosomal protein L11 [*Mycobacterium leprae*] Expect = 1e-57, Identities 122/129 (94%), Positives 123/129 (95%)
pTB-201

>gi|6225864|sp|P96393|PPCK_MYCTU phosphoenolpyruvate carboxykinase [GTP] (phosphoenolpyruvate carboxylase) (PEPCK) Mycobacterium tuberculosis (strain H37Rv)
Expect = e-103, Identities 178/213 (83%), Positives 192/213 (89%)

pTB-235

>gi|2492801|sp|O50668|YM98_MYCTU hypothetical 34.9 kDa protein Rv2298
Mycobacterium tuberculosis (strain H37Rv)
Expect = 1e-10, Identities 51/132 (38%), Positives 64/132 (47%)
**pTB-371**

sp|007184|DXS_MYCTU probable 1-deoxyxylulose-5-phosphate synthase (DXP synthase)  
*Mycobacterium tuberculosis* (strain H37Rv)

Expect = 1e-74, Identities 138/148 (93%), Positives 144/148 (97%)

---

**pTB : 2**

RELAAEIREFLIHVATCGHGLPGNCVIVNTVALHRVFDSPHDFIIDTG/GEAY/VMML 61
RELAAEIREFLIHVATCGHGLPGNLGQV +TA+LHRVFDSPHDFIIDTG/GEAY/VMML

Sbjct: 20 RELAAEIREFLIHVATCGHGLPGNCVIVNTVALHRVFDSPHDFIIDTG/GEAY/VMML 79

---

**pTB : 62**

TGRAHEFSLRRKGGLGSRSPESSEHDWVSHASAAALSYADGLAKAFETGEHRNHFHV

Sbjct: 62 TGRAHEFSLRRKGGLGSRSPESSEHDWVSHASAAALSYADGLAKAFETGEHRNHFHV 121

---

**pTB : 122**

AVVCGDA/TGMYC/MYCTU/9AA/SAPRVI 149

AVVCGDA/TGMYC/MYCTU/9AA/SAPRVI

Sbjct: 122 AVVCGDA/TGMYC/MYCTU/9AA/SAPRVI 167

---

**pTB-413**

>gi|2578375|emb|CA15451.1| (AL008609) 6-phosphogluconate dehydrogenase  
[Mycobacterium leprae]

Expect = 2e-32, Identities 74/85 (87%), Positives 75/85 (88%)

---

**pTB : 211**

T QAQGTVGLAVMGSNIARNFARKYTVLRRSIAKTDALLKEHDEKFRVCTIET 270

Sbjct: 11 TAIAQGTVGLAVMGSNIARNFARKYTVLRRSIAKTDALLKEHDEKFRVCTIET 70

---

**pTB : 271**

FLDALEKPRRVLIDVAKITPDIV 295

Sbjct: 71 FLAALQTPRRVLIDVAKITPDIV 95

---

**pTB-433**

>gi|7445701|pir|D70873 probable ABC transporter ATP binding protein  
*Mycobacterium tuberculosis* (strain H37Rv)

Expect = 3e-87, Identities 181/207 (87%), Positives 193/207 (92%)

---

**pTB : 1**

VITATDLEV-RARILSPDPGDLTVQGDRGILVRRNGRKGKTTLILARLAGETEPTITYT 59

Sbjct: 1 VITATDLEV-RARILSPDPGDLTVQGDRGILVRRNGRKGKTTLILARLAGETEPTITYT 60

---

**pTB : 60**

ARG/CEG/DPQ/KEW/DLVLAR/DHSVGLSVHLTV/LDLEK/QMALAEVA/DDAR/RAIR 119

Sbjct: 61 ARG/CEG/DPQ/KEW/DLVLAR/DHSVGLSVHLTV/LDLEK/QMALAEVA/DDAR/RAIR 120

---

**pTB : 120**

RYQ/LEF/VAL/CGYAESAEARICASLGLPER/TVHVR/LSQ/RRV/VELARILFFA 179

Sbjct: 121 RYQ/LEF/VAL/CGYAESAEARICASLGLPER/TVHVR/LSQ/RRV/VELARILFFA 180

---

**pTB : 180**

EGGAGSSTTLQMLDEPIHLDDSI 206

Sbjct: 181 EGGAGSSTTLQMLDEPIHLDDSI 204
**pTB-437**

>gi|5031428|gb|AAD38165.1|AF152394_3 (AF152394) F420-dependent glucose-6-phosphate dehydrogenase [Mycobacterium avium]

Expect = 3e-43, Identities 83/86 (96%), Positives 86/86 (99%)

**pTB : 4**

```
AQAFATTACYLPGRIFVGCYTQAENIAAGQEMPEFKKFRARVRESVRLMRELWGAD 63
```

Sbjct : 87

```
AQAFATTACYLPGRIFVGCYTQAENIAAGQEMPEFKKFRARVRESVRLMRELWGAD 146
```

**pTB : 64**

```
RVDFGEYRLKGASYIDVFDGQVIY 89
```

Sbjct : 147

```
RVDFGEYRLKGASYIDVFDGQVIY 172
```

**pTB-458**

>gi|6066632|emb|CAC58280.1 | (AL121849) putative tetR family transcriptional regulator [Streptomyces coelicolor]

Expect = 7e-17, Identities 51/103 (49%), Positives 70/103 (67%)

**pTB : 58**

```
QPVRSDAARNREALIEVATRLFAAAAGGDEPSLRLIARFAGVGGTLFRHFPTPREALVAE 117
```

Sbjct : 7

```
+P+R+DA RNR++ + A R+F+ G + L IAREAVGQ GIL+R+FFPREAL+EA
```

**pTB : 118**

```
VYQDDQVRRLEGADQCLNLASSAPPQAMRRVDMLFDMLATKHGM 160
```

Sbjct : 65

```
AYRNEVARLDSVPLLALLPAEALRAWTKRFDYATAKLKM 107
```

**pTB-470**

>gi|13092566|emb|CAC29730.1 | (AL583917) putative integral membrane peptidase [Mycobacterium leprae]

Expect = 4e-64, Identities 133/203 (66%), Positives 155/206 (74%)

**pTB : 95**

```
QQELGLNLKGNADTNFQK---GHHQISKRRSTVNA--ERKTRSGSARSSQGSLQGKLLVY 150
```

Sbjct : 59

```
EQQVRDLKSNQNTENSDKITKPYTGQLTLYNLQDKGAKMTVQVSAGQWTTLLLHY 118
```

**pTB : 151**

```
VLPLLILWGLSYLALAA---PFGFGSKVAQTLKMDKTFADAVGVEDAELVEY 206
```

Sbjct : 119

```
MLPLLMLVALFVMSSPMKQGGYRFFGKSRALKQDIKMTTFADAVGVEDAELVEY 178
```

**pTB : 207**

```
KDFQLNPGRHYHGKAPGQVYLGPPGQKTGTLLLRAVAGEAVGPFTTISGSDFVEMFGV 266
```

Sbjct : 179

```
KDFQLNPISRYGAKPGQVYLGPPGQKTGTLLLRAVAGEAVGPFTTISGSDFVEMFGV 238
```

**pTB : 267**

```
VGASRVRDLDFQAKQNNPFCITFVDEI 292
```

Sbjct : 239

```
VGASRVRDLDFQAKQNNPFCITFVDEI 264
```

**pTB-494**

emb|CAAl7409| (AL021932) hypothetical protein Rv0452 [Mycobacterium tuberculosis]

Expect = 1e-32, Identities 79/103 (76%), Positives 86/103 (82%)

**pTB : 5**

```
RVFRYTHMCVMVLAQFLRQARTEEKKKRQAEALVEAR+VALGETOAASVTLTAVARRAGI 64
```

Sbjct : 22

```
RQPSYKSGYVRPLAVQLGFRQARTEEKRQQAEALVEAR+LALTGAASVTLTAVRAGI 81
```

127
YOH9_MTCTU putative ATPase Rv3679 Mycobacterium tuberculosis (strain H37Rv)

Expect = 3e-50, Identities 129/141 (86%), Positives 129/141 (90%)
Appendix 4  

*M. a. paratuberculosis* **p22** and *M. bovis**  **lpp-27**  
(lprG) alignment (Fasta 3)
<table>
<thead>
<tr>
<th>p22-</th>
<th>GCCCGAGGCGCCGCAAAATCAACCGTCAGAGACGACGATCCGGATCAGCGGGAACGTCTCTCT</th>
<th>lpp27-</th>
<th>AGCCGAAGGGCGGATACCATTCAAGCCAGACCTTCGATCAGCGGGAAGGTAT</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>460 470 480 490 500 510</td>
<td>520 530 540 550 560 570</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p22-</td>
<td>GGCAGCCGCGGTGAACAGATCATGCGCGATTCAGCCGCAACCCAGGCGGAGCCGAC</td>
<td>lpp27-</td>
<td>GCCAGCGCGGTGAACAGATCATGCGCGATTCAGCCGCAACCCAGGCGGAGCCGAC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>580 590 600 610 620 630</td>
<td></td>
<td>980 990 1000 1010 1020 1030</td>
</tr>
<tr>
<td></td>
<td>p22-</td>
<td>CGTGCAGGACGCACGAGCCGGAACCGCCGCAGCCACCTTCGTCGGCGTACAGGAGCTC</td>
<td>lpp27-</td>
<td>CGTGCAGGACGCACGAGCCGGAACCGCCGCAGCCACCTTCGTCGGCGTACAGGAGCTC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>640 650 660 670 680 690</td>
<td></td>
<td>1040 1050 1060 1070 1080 1090</td>
</tr>
<tr>
<td></td>
<td>p22-</td>
<td>CGGGAATTCCGTGCAGGTGAGCCTGCGTGTCAATTGGGCGAGCCAGGTTCAGGTTCAGC</td>
<td>lpp27-</td>
<td>CGGGAATTCCGTGCAGGTGAGCCTGCGTGTCAATTGGGCGAGCCAGGTTCAGGTTCAGC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>700</td>
<td></td>
<td>1100 1110</td>
</tr>
<tr>
<td></td>
<td>p22-</td>
<td>CCCGGTGAGCTCGTGA</td>
<td>lpp27-</td>
<td>CCCGGTGAGCTCGTGA</td>
</tr>
</tbody>
</table>
## Raw data for IFN-γ assay Figure 4.13

<table>
<thead>
<tr>
<th>#</th>
<th>PBS</th>
<th>Avian PPD 12.5 µg</th>
<th>P22 2.0 µg</th>
<th>P22 0.64 µg</th>
<th>P22 0.32 µg</th>
<th>ConA 20 µg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A&lt;sub&gt;450 nm&lt;/sub&gt;</td>
<td>Avg</td>
<td>A&lt;sub&gt;450 nm&lt;/sub&gt;</td>
<td>Avg</td>
<td>A&lt;sub&gt;450 nm&lt;/sub&gt;</td>
<td>Avg</td>
</tr>
<tr>
<td>124</td>
<td>0.077</td>
<td>0.076</td>
<td>0.454</td>
<td>0.485</td>
<td>0.192</td>
<td>0.149</td>
</tr>
<tr>
<td>127</td>
<td>0.081</td>
<td>0.081</td>
<td>1.328</td>
<td>1.283</td>
<td>0.455</td>
<td>0.398</td>
</tr>
<tr>
<td>129</td>
<td>0.070</td>
<td>0.076</td>
<td>2.142</td>
<td>1.870</td>
<td>0.510</td>
<td>0.445</td>
</tr>
<tr>
<td>131</td>
<td>0.118</td>
<td>0.119</td>
<td>0.678</td>
<td>0.707</td>
<td>0.550</td>
<td>0.363</td>
</tr>
<tr>
<td>132</td>
<td>0.066</td>
<td>0.074</td>
<td>0.710</td>
<td>0.726</td>
<td>0.252</td>
<td>0.237</td>
</tr>
<tr>
<td>136</td>
<td>0.101</td>
<td>0.090</td>
<td>0.292</td>
<td>0.288</td>
<td>0.117</td>
<td>0.126</td>
</tr>
<tr>
<td>507</td>
<td>0.125</td>
<td>0.123</td>
<td>0.687</td>
<td>0.709</td>
<td>0.249</td>
<td>0.189</td>
</tr>
<tr>
<td>578</td>
<td>0.152</td>
<td>0.159</td>
<td>2.388</td>
<td>2.416</td>
<td>0.948</td>
<td>0.608</td>
</tr>
<tr>
<td>587</td>
<td>0.054</td>
<td>0.057</td>
<td>0.943</td>
<td>0.984</td>
<td>0.369</td>
<td>0.197</td>
</tr>
<tr>
<td>128</td>
<td>0.063</td>
<td>0.074</td>
<td>0.195</td>
<td>0.197</td>
<td>0.063</td>
<td>0.057</td>
</tr>
<tr>
<td>133</td>
<td>0.183</td>
<td>0.139</td>
<td>0.312</td>
<td>0.291</td>
<td>0.063</td>
<td>0.057</td>
</tr>
<tr>
<td>137</td>
<td>0.062</td>
<td>0.065</td>
<td>0.117</td>
<td>0.111</td>
<td>0.103</td>
<td>0.072</td>
</tr>
<tr>
<td>569</td>
<td>0.065</td>
<td>0.061</td>
<td>0.752</td>
<td>0.751</td>
<td>0.108</td>
<td>0.068</td>
</tr>
<tr>
<td>599</td>
<td>0.063</td>
<td>0.063</td>
<td>0.053</td>
<td>0.053</td>
<td>0.059</td>
<td>0.053</td>
</tr>
</tbody>
</table>
## Appendix 6  Raw data for IFN-γ assay Figure 4.14

<table>
<thead>
<tr>
<th></th>
<th>PBS</th>
<th>Johnin PPD 12.5 μg</th>
<th></th>
<th>P22 1 μg</th>
<th></th>
<th>P22 5 μg</th>
<th></th>
<th>ConA 20 μg</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ave</td>
<td>Ave</td>
<td>Ave</td>
<td>Ave</td>
<td>Ave</td>
<td>Ave</td>
<td>Ave</td>
<td>Ave</td>
<td>Ave</td>
</tr>
<tr>
<td>124</td>
<td>0.073</td>
<td>0.078</td>
<td>0.607</td>
<td>0.640</td>
<td>0.242</td>
<td>0.242</td>
<td>0.069</td>
<td>0.069</td>
<td>0.139</td>
</tr>
<tr>
<td></td>
<td>0.082</td>
<td></td>
<td>0.672</td>
<td></td>
<td>0.242</td>
<td></td>
<td>0.069</td>
<td></td>
<td></td>
</tr>
<tr>
<td>127</td>
<td>0.056</td>
<td>0.055</td>
<td>0.708</td>
<td>0.703</td>
<td>0.270</td>
<td>0.281</td>
<td>0.567</td>
<td>0.588</td>
<td>0.176</td>
</tr>
<tr>
<td></td>
<td>0.053</td>
<td></td>
<td>0.698</td>
<td></td>
<td>0.292</td>
<td></td>
<td>0.609</td>
<td></td>
<td></td>
</tr>
<tr>
<td>129</td>
<td>0.061</td>
<td>0.058</td>
<td>0.730</td>
<td>0.769</td>
<td>0.097</td>
<td>0.096</td>
<td>0.219</td>
<td>0.237</td>
<td>0.125</td>
</tr>
<tr>
<td></td>
<td>0.054</td>
<td></td>
<td>0.808</td>
<td></td>
<td>0.095</td>
<td></td>
<td>0.255</td>
<td></td>
<td></td>
</tr>
<tr>
<td>131</td>
<td>0.053</td>
<td>0.121</td>
<td>1.543</td>
<td>1.350</td>
<td>0.255</td>
<td>0.247</td>
<td>ND</td>
<td>0.155</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.189</td>
<td></td>
<td>1.166</td>
<td></td>
<td>0.239</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>132</td>
<td>0.058</td>
<td>0.138</td>
<td>2.271</td>
<td>2.040</td>
<td>0.590</td>
<td>0.628</td>
<td>ND</td>
<td>0.103</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.218</td>
<td></td>
<td>1.800</td>
<td></td>
<td>0.665</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>587</td>
<td>0.077</td>
<td>0.081</td>
<td>0.590</td>
<td>0.608</td>
<td>0.213</td>
<td>0.208</td>
<td>ND</td>
<td>0.181</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.084</td>
<td></td>
<td>0.625</td>
<td></td>
<td>0.203</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>578</td>
<td>0.056</td>
<td>0.057</td>
<td>1.466</td>
<td>1.503</td>
<td>0.090</td>
<td>0.085</td>
<td>ND</td>
<td>0.970</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.058</td>
<td></td>
<td>1.540</td>
<td></td>
<td>0.079</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>587</td>
<td>0.060</td>
<td>0.059</td>
<td>1.128</td>
<td>1.113</td>
<td>0.093</td>
<td>0.091</td>
<td>ND</td>
<td>0.079</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.057</td>
<td></td>
<td>1.098</td>
<td></td>
<td>0.088</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>128</td>
<td>0.049</td>
<td>0.049</td>
<td>0.140</td>
<td>0.140</td>
<td>0.055</td>
<td>0.055</td>
<td>0.060</td>
<td>0.070</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td>0.049</td>
<td></td>
<td>0.140</td>
<td></td>
<td>0.055</td>
<td></td>
<td>0.079</td>
<td></td>
<td></td>
</tr>
<tr>
<td>133</td>
<td>0.073</td>
<td>0.087</td>
<td>0.125</td>
<td>0.110</td>
<td>0.135</td>
<td>0.152</td>
<td>0.319</td>
<td>0.279</td>
<td>0.151</td>
</tr>
<tr>
<td></td>
<td>0.100</td>
<td></td>
<td>0.095</td>
<td></td>
<td>0.148</td>
<td></td>
<td>0.238</td>
<td></td>
<td></td>
</tr>
<tr>
<td>137</td>
<td>0.051</td>
<td>0.052</td>
<td>0.073</td>
<td>0.069</td>
<td>0.054</td>
<td>0.054</td>
<td>0.072</td>
<td>0.075</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td>0.053</td>
<td></td>
<td>0.064</td>
<td></td>
<td>0.054</td>
<td></td>
<td>0.077</td>
<td></td>
<td></td>
</tr>
<tr>
<td>569</td>
<td>0.060</td>
<td>0.057</td>
<td>0.131</td>
<td>0.142</td>
<td>0.082</td>
<td>0.081</td>
<td>ND</td>
<td>0.113</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.054</td>
<td></td>
<td>0.153</td>
<td></td>
<td>0.080</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>599</td>
<td>0.069</td>
<td>0.066</td>
<td>0.066</td>
<td>0.064</td>
<td>0.088</td>
<td>0.084</td>
<td>ND</td>
<td>0.139</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.062</td>
<td></td>
<td>0.062</td>
<td></td>
<td>0.079</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ND** = not done
Appendix 7 Alignment of DNA and protein sequences of *M. a. paratuberculosis* p22 with partial sequences of *M. intracellularae* and *M. scrofulaceum*

<table>
<thead>
<tr>
<th></th>
<th>M. ptb p22</th>
<th>M. intracellularae</th>
<th>M. scrofulaceum</th>
</tr>
</thead>
<tbody>
<tr>
<td>A'roCAGACCCGCCGCCGCCTATCGGCCGTITl'CGCTCCCTGACCCTCGCCACCGCCTI'GA:</td>
<td>GCAGAGCGGT96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCCGCCTATCGGCC.'GTCC'Ir�CTGACCCTCGCCACCGCCTI'GATCGCCGCTGTI'CGTCGGCAGAGCGGC83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>* * ** * * ** **** *** * *GCATCCCTGACCCTCGCCACCGCCTI'GATCGCCGGCTGTI'CGTCGGCAGAGCAGC159</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCGCCGCTGCCCGACCCCACCAGCGGTCAAGCA GTCGGCCGACGCGACCAA GAA CGTCAAGAGCG'roCACC'IGG'roCTCAGCA TCCA GGGCAAG192</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCACCACTCCCCGACGGCACGACCTroGTCAAGCA GTCGGCCGACGCCACCAA GGCCGTCAAGAGTG'CCTCGTGCTCAGCATCCAGGGCAAG179</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATCTCCGGGCTGCCCATCAAGACGCTGACCGG'lGACCTCACCACCGACGCCGGCCACCGCCGCGAAGGGCAACGCCACGATCACCC'IGGGCGGCTCG288</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GACATCGACGCCAAC'ITCGTCGTCGTCGACGGCACCC'mI'ACGCCACCCTCACCCCGAACAAG'roGA GCGACTICGGCAAGGCGTCCATCTAC384</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGAGCACGATCCGGA TCA GCGGGAACGTCTCGGCGGACGCGG'mAACAAGA TCA 'IGCCGCAGTI'CAACGCCACCCAGCCGG'roCCGAGCACCG'ro576</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGGGTCCAGGAGACCGGCGACCACCAGC'IGGTI'CA GGCCAACC'lGCAGAAGAGCTCCGGGAA�CG'roCAGG'roACGC'IGTCGAA TroGGGCGAG654</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**133**

Identify to p22

88.5 (579/654 base pairs)

87.7 (573/653 base pairs)
M. ptb P22

1 MQTRRLSAVFASLTLATALLGCSSGSKSGBPAPDPTSLVKQSDA7SVN/VSLQQLSTQKISGLPITKTLGDIITPAPAATAKGNATTLOG 95

M. intracellularae

1 RLAVLGLSTLAVALLGCSSGSKSGBPAPDPTTLVQSDATKVKQSADALTQKISGLPITKTLGDIITPAPAATAKGNATTLOG 90

M. scrofulaceum

1 ***** ASITLAVTALLGCSSGSKSSAPAPDATTLVQSDATAKVSVL/QKLISGLPITKTLGDIITTPAPAATAKGNATTLOG 84

************ *** *** ** *** ********** **.** **. *** ******** ****** ********

M. ptb P22

96 SDIDANFVVVDQTLATTLTPKWSDFGKASDLYTDSLVQNEGQLGNALANFSN4AK4G4E4T4F44Q5S5TIR5S5Q5S5N5V5Q5N4P54F44N4A4T4S5

M. intracellularae

91 SDI4DNFVVVDQTLATTLTPKWSDFGKASDLYTDSLVQNEGQLGNALANFS4AKG4E4T4F44Q5S5TIR5S5Q5S5N5V5Q5N4P54N4T5S5

M. scrofulaceum

90 SDI4DNFVVVDQTLATTLTPKWSDFGKASDLYTDSLVQNEGQLGNALANFS4AKG4E4T4F44Q5S5TIR5S5Q5S5N5V5Q5N4P54N4T5S5

************ **.** **. *** ********** **.** **. *** ******** ****** ********

M. ptb P22

191 TV/4Q5/V4D4H4Q4L4Q4A4N4Q4S4Q4S4N4S4V4T4K4P4V4S4S4 235 % identity to P22

M. intracellularae

186 TV/4Q5/V4D4H4Q4L4Q4A4N4Q4S4Q4S4N4S4V4T4K4P4V4S4S4 210 87.6 (191/218 amino acids)

M. scrofulaceum

180 TV/4Q5/V4D4H4Q4L4Q4A4N4Q4S4Q4S4N4S4V4T4K4P4V4S4S4 211 93.8 (190/211 amino acids)

***.****.***.*****.***** *

* represent identical residues
.
represent similar amino acids
References


References


References


References


References


References


References


References


References


References


References


References