Cloning and expression in *Escherichia coli* of DNA encoding a 60 kDa stress protein of *Mycobacterium paratuberculosis*, the causative agent of Johne's disease

Angela Colston, Ian McConnell and Raymond Bujdoso†

Author for correspondence: Raymond Bujdoso. Tel: +44 223 337 655. Fax: +44 223 337 610.

Department of Veterinary Pathology, University of Edinburgh, Summerhall, Edinburgh EH9 1QH, UK

Polymerase chain reaction (PCR) was used to generate DNA encoding a 60 kDa stress protein of *Mycobacterium paratuberculosis* using primers complementary to sequences at the 5' and 3' ends of 60 kDa stress protein genes (encoding the '65 kDa antigens') of *M. leprae* and *M. tuberculosis*. The predicted PCR product of 1·8 kb contained the entire coding sequence of an *M. paratuberculosis* 60 kDa stress protein, with non-coding regions of 124 bp and 1 bp at the 5' and 3' ends, respectively. DNA encoding the entire ORF for the 60 kDa stress protein, as well as thrombin and Factor Xa proteolytic cleavage sites, was ligated into the bacterial expression vector pGEX-2T and used to transform *Escherichia coli* strain JM83. Transformed bacteria, induced by IPTG, expressed an 85 kDa fusion protein comprising glutathione S-transferase (GST) and *M. paratuberculosis* 60 kDa stress protein. This fusion protein was purified by adsorption to glutathione-agarose beads and shown to cross-react in Western blot analysis with an anti-mycobacterial 60 kDa stress protein monoclonal antibody. Recombinant *M. paratuberculosis* 60 kDa stress protein was liberated from GST by proteolytic cleavage with either thrombin or Factor Xa enzyme. Authenticity of liberated recombinant stress protein was confirmed by N-terminal amino acid sequencing.

Keywords: *M. paratuberculosis*, hsp60, stress protein, Johne's disease

INTRODUCTION

Mycobacterium paratuberculosis (re-named *M. avium* subsp. *paratuberculosis*) causes Johne's disease, resulting in a chronic granulomatous inflammation through gastrointestinal infection of macrophages and epithelioid cells (Buergeit *et al.*, 1978; Chiodini *et al.*, 1984). An important element in the study of the immune response to *M. paratuberculosis* is the *in vitro* analysis of the receptor repertoire and functional capabilities of ovine T cells which are specific for *M. paratuberculosis* antigens. Other investigations of related mycobacterial species, including *M. tuberculosis* and *M. bovis*, have identified the bacterial 60 kDa stress protein (hsp60 or '65 kDa antigen') as a major target of the host immune response to these

pathogens (Shinnick *et al.*, 1987; Young *et al.*, 1988). It seems paradoxical that mycobacterial 60 kDa stress proteins are immunologically dominant antigens (Cohen, 1991), because the amino acid identity between prokaryotic and eukaryotic 60 kDa stress proteins is in the range of 40–50 %, conferring upon these mycobacterial proteins the potential to generate autoimmune disease. This paradox may be resolved if the immunological dominance of the mycobacterial 60 kDa proteins arises from the host's immune response to its own 60 kDa stress protein, presented by cells stressed by stimuli including infection (Janeway *et al.*, 1988). In this situation, the production of *M. paratuberculosis* 60 kDa stress protein within infected macrophages may affect the host stress response as a consequence of excessive production and presentation of conserved 60 kDa stress protein epitopes. Such immune reactivity may contribute to the pathogenesis of Johne's Disease.

The utility of sheep, a natural host for *M. paratuberculosis*, to test this hypothesis has become a reality in recent years

†Present address: Department of Clinical Veterinary Medicine, University of Cambridge, Madingley Road, Cambridge CB2 0ES, UK.

Abbreviation: GST, glutathione S-transferase.

The EMBL accession number for the nucleotide sequence reported in this paper is X74518.

through the development of a panel of monoclonal antibodies that define the major populations of T cells in this species (Mackay *et al.*, 1986). A major limitation, however, has been the lack of *M. paratuberculosis* 60 kDa stress protein for use in immunological assays to assess its potential role in the pathogenesis of Johne's disease. Here, we describe the successful generation of recombinant *M. paratuberculosis* 60 kDa stress protein using the *Escherichia coli* expression system pGEX-2T. The availability of recombinant *M. paratuberculosis* 60 kDa stress protein will allow a more thorough analysis of the contribution of this antigen to Johne's Disease.

METHODS

Bacteria and plasmids. *Escherichia coli* strains JM101 and JM83 were used as recipients of the plasmids pTZ18 and pTZ19R (Pharmacia). The expression vector pGEX-2T (Smith & Johnson, 1988) was a gift from Dr D. Smith, University of Edinburgh. In the pGEX-2T expression system, the protein of interest is expressed as a fusion protein with the glutathione γ -transferase (GST) of *Schistosoma japonicum* (Smith & Johnson, 1988). An important feature of this GST-fusion protein is that it is easily purified from bacterial lysate by a simple affinity purification step, involving binding of the GST moiety to its specific substrate, glutathione, immobilized on agarose beads (Smith & Johnson, 1988).

Antiserum. Monoclonal antibody IIH9 (IgG₁) (Gillis & Buchanan, 1982) which recognises the 60 kDa stress protein of *M. leprae*, was a gift from Dr T. M. Shinnick (Centers for Disease Control, Atlanta, GA, USA).

Detection of *M. paratuberculosis*-specific IS900 insertion sequence. *M. paratuberculosis* (strain 18) genomic DNA was a kind gift from the Central Veterinary Laboratory, Weybridge, Surrey, UK. The authenticity of this DNA was verified by a polymerase chain reaction (PCR) to detect the *M. paratuberculosis*-specific IS900 insertion sequence (Vary *et al.*, 1990) using a 5' primer comprising 5'-CCGCTAATTGAGAGATGCGATTGG-3' and a 3' primer comprising 5'-AATCAACTCCAGCAGCGCGCCTCG-3'. PCR was carried out as described by Saiki (1989) with 0.1 nM of each primer and a previously determined optimum concentration of MgCl₂ and DMSO. PCR products were electrophoresed in 1.5% (w/v) agarose gel and either visualized by ethidium bromide staining and UV-transillumination, or transferred onto nitrocellulose membrane for Southern blot analysis. Membranes were prehybridized in 7% (w/v) SDS, 0.5 M phosphate buffer (Na₂HPO₄/NaH₂PO₄) pH 7, 1% BSA for 4–6 h at 68 °C. The prehybridization fluid was replenished and a ³²P-labelled IS900 oligonucleotide 5'-AGGTTGTGCCACAACCACCTCCGTA-3', which was internal to the two PCR primers, added and incubated at 68 °C overnight. The filter was washed in 2 × SSC and 0.5% SDS,

with a final wash in 0.1 × SSC and 0.5% SDS, dried, wrapped in cling film and exposed to Kodak XAR film.

Cloning of DNA encoding 60 kDa stress protein. *M. paratuberculosis* (strain 18) genomic DNA, verified as described above, was used as substrate in a PCR using primers which contained common sequences found in the 60 kDa stress protein of *M. leprae* (Mehra *et al.*, 1986) and *M. tuberculosis* (Shinnick, 1987). The 5' primer sequence was 5'-AGTGCT/CAGGTCGGGACGGTGAG-3' and the 3' primer sequence 5'-GTCAGAAG/ATCCATA/GCCACCCATG-3'. The 1.8 kb reaction product was blunt-end-ligated into the phagemids pTZ19R and pTZ18R and rescued by transformation of *E. coli* JM101. The 1.8 kb inserts of two independent transformants were fully sequenced by the dideoxy method (Sanger *et al.*, 1977) using internal primers, and through the generation of subclones obtained by restriction digest. One transformant designated MPHSP60A was used as substrate in a PCR to generate DNA encoding the entire putative ORF of the 60 kDa stress protein gene starting at nucleotide 125. The full primer sequences used to generate the 1.6 kb DNA fragment are shown in Fig. 1.

The 1.6 kb PCR product was ligated into the unique *Bam*HI site of the bacterial expression vector pGEX-2T, and rescued by transformation of JM83 bacteria. Transformants grown on LB-ampicillin agar plates were selected by restriction digestion and double-stranded-sequencing through the 5' vector-insert junction of pGEX-2T for those which contained the 60 kDa stress protein DNA in the correct orientation, and with a maintained reading frame for fusion protein expression. Positive clones were grown in small-scale liquid culture using LB supplemented with 150 µg ampicillin ml⁻¹. Following induction with IPTG, crude bacterial lysates were analysed by SDS-PAGE. One transformant, designated MP2211B, was selected for large-scale culture. A transformant with reversed insert was also selected as an appropriate control.

Purification of recombinant 60 kDa stress protein. The fusion protein comprising GST-stress protein was routinely purified from 5 l IPTG-induced culture of MP2211B in LB plus 150 µg ampicillin ml⁻¹ in Erlenmeyer flasks on an orbital shaker at 37 °C. Pelleted bacteria were resuspended in PBS and lysed by sonication in the presence of the protease inhibitors aprotinin, antipain, chymostatin, leupeptin, pepstatin (all at 2.5 µg inhibitor ml⁻¹) and PMSF (0.2 mM) (all from Sigma). The bacterial lysate was made up to 1% (w/v) in Triton X-100, clarified by centrifugation at 10000 g, and rotated with glutathione-agarose beads [about 30 ml supernatant (150 mg dried beads)⁻¹] at room temperature for 30 min. The beads had previously been washed in PBS (0.136 M NaCl, 2.6 M KCl, 8.1 mM Na₂HPO₄, 1.4 mM KH₂PO₄, pH 7.2) plus 1% Triton X-100 (PBS-Triton). The glutathione-agarose beads were washed three times with PBS-Triton and twice with protease cleavage buffer [50 mM Tris/HCl, 1 mM CaCl₂ (2.5 mM CaCl₂ with thrombin) 100 mM NaCl, 0.05% deoxycholate, pH 7.5]. Either thrombin (Sigma) at 25 units ml⁻¹ or Factor Xa enzyme (Boehringer Mannheim) at 0.2% (w/w) was added for 1 h at room temperature, after which time *p*-(amidinophenyl)methanesulphonyl fluoride (*p*-APSMF) (Sigma) was added to a final concentration of 80 µg ml⁻¹ to halt the reaction. Beads were pelleted by centrifugation, and clarified supernatant was retained along with two further supernatants collected from two washes of the beads with 50 mM Tris/HCl, 100 mM NaCl, pH 7.5. The three supernatants were pooled and dialysed into 20 mM Tris/HCl, pH 8. Recombinant *M. paratuberculosis* 60 kDa stress protein was further purified by fractionation on an anion (DEAE-5PW) HPLC column (10 × 95 mm). Material was eluted

*Bam*HI Factor Xa site First 22 bp of 60 kDa ORF

5' primer: 5'-GAGTTCGGGATCCATAGAGGTAGAAATGGCCAAGACAATTGCGTACG-3'

*Bam*HI Last 19 bp of 60 kDa ORF

3' primer: 5'-TTACACCGGGATCCTTCTATCAGTCAGAGTCCATGCCACCCATG-3'

Stop

Fig. 1. Primer sequences used to generate the 1.6 kb DNA fragment by PCR.

with a 0–300 mM gradient of NaCl generated over 40 min at a flow rate of 1 ml min⁻¹. Two millilitre fractions were collected and analysed as described below. Fractions containing purified stress protein were pooled, dialysed into PBS and the protein concentration estimated by BioRad protein assay. Purified recombinant 60 kDa stress protein was stored at -70 °C at a final concentration of 1 mg ml⁻¹.

Protein analysis. Protein preparations were analysed by SDS-PAGE under reducing conditions (Laemmli, 1970). Total proteins were detected by staining gels with 0.25% (w/v) Coomassie Brilliant Blue G-250 in 20% (v/v) methanol, 5% (v/v) acetic acid. For Western Blot analysis, proteins were electrophoresed as above, then transferred to nitrocellulose membranes (Hybond-C, Amersham) using a semi-dry electroblotter (Kyhse-Anderson, 1984). After blocking in PBS plus 5% (w/v) dried milk powder (Marvel), blots were probed by incubation overnight in specific antibody and washed five times in PBS plus 1% dried milk powder plus 0.1% Tween 80 (PBS-MP) over 25 min. Blots were incubated for 1 h at room temperature with affinity-purified anti-mouse IgG-biotin conjugate (Sigma), washed with PBS-MP as before, and incubated for 1 h at room temperature with streptavidin-alkaline phosphatase conjugate (Boehringer Mannheim). Blots were washed again before development with 330 µg nitroblue tetrazolium ml⁻¹ (Sigma) and 165 µg 5-bromo-4-chloro-3-indolyl phosphate ml⁻¹ (Sigma) in 0.1 M Tris/HCl, pH 9.5, 100 mM NaCl and 5 mM MgCl₂ (Pluzek & Ramlau, 1988).

N-terminal sequence analysis. Proteins were separated by SDS-PAGE and electroblotted, as described above, onto polyvinylidene difluoride membranes (Problott transfer membranes, Applied Biosystems). The proteins of interest were identified by Coomassie staining, and sequenced with an Applied Biosystems 470A protein sequencer.

RESULTS

Verification of *M. paratuberculosis* genomic DNA

The strategy to clone and express *M. paratuberculosis* 60 kDa stress protein was to use PCR to amplify DNA encoding this protein from genomic *M. paratuberculosis* DNA. To verify that the DNA to be used was derived from *M. paratuberculosis*, a PCR was carried out to detect the *M. paratuberculosis*-specific insertion sequence IS900. PCR primers specific to the 5' and 3' ends of the IS900 coding sequence were used to amplify a fragment of the gene which was subsequently detected by Southern blot with a ³²P-labelled primer internal to those used for the PCR. Fig. 2 clearly shows the presence of the predicted 229 bp IS900 PCR product, verifying that the genomic DNA was that of *M. paratuberculosis*.

Cloning and sequence of DNA encoding the 60 kDa stress protein

DNA encoding the 60 kDa stress protein was generated by PCR using *M. paratuberculosis* genomic DNA as template DNA with primers comprising common sequences found at the 5' and 3' ends of 60 kDa stress protein genes of *M. leprae* (Mehra *et al.*, 1986) and *M. tuberculosis* (Shinnick, 1987). The predicted PCR product of 1.8 kb was cloned into the phagemids pTZ18R and pTZ19R, and used to transform *E. coli* strain JM101.

Insert DNA from two independent clones, MPHSP60A and MPHSP60B, was fully sequenced in both directions using internal sequencing primers and a subcloning strategy as shown in Fig. 3. Both clones exhibited exactly the same sequence (EMBL accession number X74518). The 1.8 kb fragment had a relatively high G + C content of 68.2 mol%, which is similar to that of DNA from other mycobacteria. The 1.8 kb DNA fragment contains an ORF beginning at ATG (nucleotide residue 125–127) which codes for a methionine residue. The ORF proceeds for 1623 bp and terminates at the stop codon TGA (nucleotides 1748–1750). This ORF encodes a potential protein of 541 amino acids, with a predicted molecular mass of 58 kDa in the absence of any glycosylation.

Generation of recombinant 60 kDa stress protein

DNA encoding the entire ORF of the predicted *M. paratuberculosis* 60 kDa stress protein was generated by PCR using plasmid MPHSP60A as substrate and with primers flanked with *Bam*HI sites. The predicted 1.6 kb product was ligated into the unique *Bam*HI site of the

Fig. 2. Verification of *M. paratuberculosis* genomic DNA. Aliquots (10 µl) of a PCR carried out with primers specific for a 229 bp fragment of IS900 and different template DNA were electrophoresed in a 1.5% agarose gel and visualized by ethidium bromide staining (lanes 1–3) or Southern blotting (lanes 4–6) using a ³²P-labelled IS900 probe internal to the PCR primers. Lanes 1 and 4, no substrate DNA; lanes 2 and 5, *M. paratuberculosis* (strain 18) DNA; lanes 3 and 6, *E. coli* DNA. Markers are *Eco*RI/*Hind*III-digested λDNA.

Fig. 3. Partial restriction map of *M. paratuberculosis* 60 kDa stress protein DNA (coding region hatched) indicating the restriction enzymes used to generate subclones (continuous black lines) for sequence analysis.

Fig. 4. pGEX-2T expression vector for the production of recombinant *M. paratuberculosis* 60 kDa stress protein.

bacterial expression vector pGEX-2T (Fig. 4). Recombinant clones in *E. coli* JM83 were sequenced through the insertion site of the vector to select those with the correct orientation and maintenance of reading frame. One transformant, MP2211B, was chosen for further work. A crude extract of MP2211B grown in liquid culture with IPTG-induction was analysed by SDS-PAGE. Lysate from this transformant showed the presence of a prominent band of molecular mass 85 kDa (Fig. 5, lane 3) which correlates with the predicted size of the GST-60 kDa fusion protein, and which was not present in lysates of untransformed JM83 (Fig. 5, lane 1). Lysates of JM83 transformed with pGEX-2T containing the ORF in the antisense direction showed a unique protein band of approximately 87 kDa (Fig. 5, lane 2). Evidence

that the 85 kDa protein present in the MP2211B lysate was indeed the GST-60 kDa fusion protein was shown by its retention on glutathione-agarose beads (Fig. 5, lane 4). Further evidence that the 85 kDa protein was the predicted GST-60 kDa fusion protein, was obtained by Western blot analysis using monoclonal antibody IIH9 which was raised to *M. leprae* hsp60 (Gillis & Buchanan, 1982). This monoclonal antibody reacted with an 85 kDa protein present in lysate derived from MP2211B, and also with the material from this lysate that was adsorbed, and subsequently eluted, from the glutathione-agarose beads (Fig. 5, lanes 7 and 8, respectively). Monoclonal IIH9 did not react with lysate from untransformed JM83, or from bacteria transformed with pGEX-2T containing PCR product in the reverse direction (Fig. 5, lanes 5 and 6, respectively).

The plasmid encoding GST-60 kDa stress protein was engineered to contain both thrombin and Factor Xa proteolytic cleavage sites to allow the separation of the 60 kDa stress protein from GST. Fig. 6 shows that efficient cleavage of fusion protein adsorbed onto glutathione-agarose beads occurred in the presence of Factor Xa or thrombin enzyme. Factor Xa liberated three proteins, of molecular mass 57, 53 and 27 kDa (Fig. 6, lane 2). Only the 57 and 53 kDa bands reacted with monoclonal antibody IIH9 (Fig. 6, lane 5), suggesting the 27 kDa band was GST. Cleavage of the fusion protein with thrombin released a single protein of 58 kDa (Fig. 6, lane 3) which reacted with monoclonal antibody IIH9 (Fig. 6, lane 6).

N-terminal sequence analysis of the 57 kDa Factor Xa cleavage product revealed the N-terminus of *M. paratuberculosis* 60 kDa stress protein predicted from the nucleic acid sequence (Fig. 7). The five amino acids at the N-terminus of the 53 kDa Factor Xa cleavage product were NVVLE, which corresponded to amino acid position 36–40 of the authentic 60 kDa stress protein predicted from the nucleic acid sequence. N-terminal amino acid sequence analysis of the thrombin cleavage

Fig. 5. Expression of GST-stress protein fusion protein. SDS-PAGE (10% acrylamide) was performed on lysates of IPTG-induced pGEX-2T transformed bacteria and the gels visualized by Coomassie Blue (lanes 1–4) or by Western blot (lanes 5–8) using monoclonal antibody IIH9. Lanes 1 and 5, untransformed JM83; 2 and 6, JM83 transformed with MP2211B (insert reversed); 3 and 7, JM83 transformed with clone MP2211B; 4 and 8, glutathione-agarose bead-adsorbed material from JM83 transformed with clone MP2211B.

Fig. 6. Cleavage of *M. paratuberculosis* 60 kDa stress protein from GST. Aliquots (100 μ l) of glutathione-agarose beads treated with lysate from IPTG-induced bacterial clone MP2211B were cleaved with different proteolytic enzymes for 30 min at room temperature. Beads were pelleted by centrifugation and cleavage products within the supernatant analysed by 10% acrylamide SDS-PAGE. Beads treated with no enzyme were boiled in sample buffer prior to pelleting of beads. Gels were visualized by Coomassie Blue (lanes 1–3) or Western blotting (lanes 4–6) using monoclonal antibody 11H9. Lanes 1 and 4, no proteolytic enzyme; lanes 2 and 5, Factor Xa (1 mg ml⁻¹); 3 and 6, thrombin (25 units ml⁻¹).

Fig. 7. N-terminal sequence of fusion protein proteolytic cleavage products.

product was consistent with predicted additional amino acids between the thrombin cleavage site and the authentic N-terminus of the native molecule.

Routinely, thrombin was used to cleave the 60 kDa stress protein from fusion protein adsorbed to glutathione-agarose beads, and the protein was subsequently purified by ion-exchange HPLC. The recombinant *M. paratuberculosis* hsp60 was routinely > 95% pure as judged by silver stained SDS-PAGE with a final yield of approximately 2 mg l⁻¹ of original bacterial culture.

DISCUSSION

Research into the immunopathogenesis of *M. paratuberculosis* infection (Johne's disease) in ruminants has gained new impetus with the resurgence in interest of a role for this organism in the pathogenesis of Crohn's disease in humans (Seldenrijk *et al.*, 1990; Moss *et al.*, 1991). As a consequence, there is a requirement for the generation of defined reagents, such as recombinant antigens, which will allow a more detailed examination of the immunopathogenesis of these two diseases. Here we have described the cloning and expression of *M. paratuberculosis* 60 kDa stress protein.

Because of the relatedness of mycobacterial species it was important to establish that the DNA to be used as PCR template was that of *M. paratuberculosis*. This was achieved through the use of a PCR to detect the *M. paratuberculosis*-specific insertion sequence IS900 (Green *et al.*, 1989) found in all *M. paratuberculosis* strains so far studied. IS900 is not present in members of the closely related *M. avium*-*M. intracellulare* (MAI) complex (Vary *et al.*, 1990; McFadden *et al.*, 1987), whose genomes have been shown to differ in sequence by only 1–2% from that of *M. paratuberculosis*. Using PCR primers, and an internal oligonucleotide probe, all specific for IS900, the authenticity of *M. paratuberculosis* DNA was confirmed. The source of mycobacterial genomic DNA used here was *M. paratuberculosis* strain 18. Strain 18 is an atypical strain of *M. paratuberculosis* that is now considered to be a strain of *M. avium* (Chiodini, 1993). In studies by others, *M. paratuberculosis* strain 18 did not contain the IS900 insertion sequence but did contain the related sequence IS901 (Kunze *et al.*, 1992). The presence of the IS900 insertion sequence in the genomic DNA used in our studies reported here suggests that this material originates from an authentic *M. paratuberculosis* strain. A PCR using primers which contained common sequences found in the 60 kDa stress protein of *M. tuberculosis* and *M. leprae*, was

used to derive the DNA clone MPHSP60A which potentially encoded the 60 kDa stress protein of *M. paratuberculosis*.

Sequence analysis of clone MPHSP60A revealed a G + C content of 68.2 mol%, which is consistent with the high G + C content reported for *M. paratuberculosis* genomic DNA. This relatively high G + C content reflects the bias in codon usage towards a G or C at codon position 3, which is a possible mechanism for conservation of genetic material (Ikemura, 1985). The nucleotide sequence of clone MPHSP60A shows two possible start codons: a GTG at nucleotide position 2–4, and an ATG at nucleotide position 125–127. It is likely that the ATG is an authentic start signal because 12 bp upstream of this codon there is a putative Shine–Dalgarno sequence (5 out of 5 nucleotide match with GGAGG). Five lines of evidence strongly suggest that clone MPHSP60A encodes a hsp60 from *M. paratuberculosis*. Firstly, the PCR primers designed from the *M. tuberculosis* and *M. leprae* hsp60 gene sequences gave a product of the predicted length (1.8 kb). Furthermore, the 1623 bp ORF in clone MPHSP60A is predicted to result in a 541 amino acid polypeptide with a molecular mass of approximately 56.4 kDa, similar to that of other hsp60 proteins, including the 65 kDa antigens from the three related mycobacteria, *M. bovis* (Thole *et al.*, 1987), *M. leprae* (Mehra *et al.*, 1986), and *M. tuberculosis* (Shinnick, 1987). Secondly, the degree of amino acid sequence identity between the ORF in clone MPHSP60A and the mycobacterial hsp60 (65 kDa antigen) homologues is about 94%. Comparison of the mycobacterial hsp60 protein sequences reveals that the differences between *M. paratuberculosis* and *M. leprae* are evenly spread throughout the sequence whereas the differences between *M. paratuberculosis* and *M. tuberculosis* are clustered at the C-termini (data not shown). Thirdly, nucleotide residues 316–658 of clone MPHSP60A are identical to the sequence of a 343 bp PCR product generated from *M. paratuberculosis* DNA by Hance *et al.* (1989) using primers complementary to *M. bovis* 60 kDa stress protein. Fourth, the expressed ORF of clone MPHSP60A reacted in a Western blot analysis with monoclonal antibody IIH9, which was raised to *M. leprae* 60 kDa stress protein (65 kDa antigen). Fifth, the sequence of MPHSP60A at nucleotides 1196–1234 encodes the consensus sequence YXXXXLXERXAKL which is thought to form part, or all, of the ATPase site of some stress proteins (Luis *et al.*, 1990).

The mycobacterial hsp60s are homologues to other stress proteins and chaperonins, such as the GroEL protein of *E. coli* (Hemmingsen *et al.*, 1988), the human P1 mitochondrial protein (Jindal *et al.*, 1989) and the ribulose 1,5-bisphosphate carboxylase large-subunit-binding-protein of chloroplasts (Hemmingsen *et al.*, 1988). In the case of *E. coli*, the GroEL protein is expressed from an operon along with the 10 kDa GroES homologue (Hemmingsen *et al.*, 1988). This is not the case for the hsp60 (65 kDa antigen) genes in mycobacteria, although *M. leprae* (Rinke de Wit *et al.*, 1992) and *M. tuberculosis* (Kong *et al.*, 1993) do have, at another locus, an operon containing an hsp60 isoform (60% identity at the protein level with the

65 kDa antigen) and a GroES homologue. The nomenclature used to denote the hsp60 homologues in the literature is variable and confused. A clarification of the terms and a call for an agreed terminology has been presented (Hunt *et al.*, 1994). We prefer the term hsp60.

It was decided that the expressed recombinant *M. paratuberculosis* hsp60, generated here, would have as its N-terminus the methionine encoded by nucleotides 125–127 of clone MPHSP60A, because the resultant protein would be homologous with other sequenced hsp60s, and in particular, the first 12 predicted residues of the expressed protein would have complete identity with published N-terminal amino acid sequence of purified *M. bovis* 65 kDa antigen (Thole *et al.*, 1987). The GST-*M. paratuberculosis* hsp60 fusion protein expressed from the pGEX-2T vector was the most predominant cellular protein in *E. coli* transformed with this construct. The fusion protein was efficiently adsorbed onto glutathione-agarose beads and after a single round of affinity-purification a relatively pure preparation of GST-hsp60 was obtained. Site-specific proteolytic cleavage of fusion protein attached to the glutathione-agarose beads was found to be the most effective method for isolation of soluble recombinant hsp60. Thrombin cleavage of GST-hsp60 fusion protein resulted in the release of a single protein of 58 kDa. N-terminal sequence analysis predicted this to be authentic *M. paratuberculosis* hsp60 with additional amino acids at its N-terminus. During prolonged incubation with thrombin, smaller products were observed which were detected with monoclonal antibody IIH9. These were presumed to be degradation products of the recombinant hsp60 and were prevented by the addition of the irreversible thrombin inhibitor *p*-(aminophenyl)methanesulphonyl fluoride (Laura *et al.*, 1980) immediately after cleavage reactions. Factor Xa treatment of GST-hsp60 fusion protein resulted in authentic recombinant hsp60 and a 53 kDa cleavage product of this protein. The 53 kDa cleavage product had the sequence NVVLE as its first five amino acids, which are predicted from the DNA sequence to be amino acid residues 36–40 of native *M. paratuberculosis* hsp60. Amino acid residues 32–45 of native *M. paratuberculosis* hsp60 are predicted to be PKGR. This sequence has matches at two out of four residues to the reportedly specific Factor Xa cleavage site IEGR (Nagai *et al.*, 1984). Other Factor Xa cleavage sites include the prothrombin sequence IDGR (Magnusson *et al.*, 1975), the HIV p24 sequence ISPR (Gilmour *et al.*, 1989) and more recently the IL-1 α sequence IKPR and IL-1 β sequence IEEK (Fiskerstrand *et al.*, 1992). These, and the PKGR reported here, appear to be less specific, or relaxed, Factor Xa cleavage sites.

Mycobacterial hsp60s and their mammalian homologues have approximately 50% identity and contain significant stretches of complete sequence identity, interspersed with regions of variability (Hemmingsen *et al.*, 1988; Jindal *et al.*, 1989). The mycobacterial hsp60s may therefore be viewed as comprising mycobacterium-specific epitopes (non-conserved epitopes) and epitopes that resemble host hsp60 epitopes (conserved epitopes). During an immune response to a mycobacterial infection, mycobacterial

hsp60s will be processed by infected cells, and both non-conserved and conserved epitopes of the hsp60s will be presented to the host immune system, along with host epitopes. Recognition of self stress proteins, possibly by $\gamma\delta$ T cells (Janeway *et al.*, 1988), may be a natural mechanism by which the host signals to its immune system the presence of an infected, or stressed cell. Initial infection of host cells with *M. paratuberculosis* could therefore elicit an immune response initially targeted on self stress protein, which becomes directed onto mycobacterial hsp60. In the continued presence of mycobacteria, constant stimulation of T cells specific for hsp60 epitopes may result in excessive cytotoxicity and tissue damage. The generation of recombinant *M. paratuberculosis* hsp60 described here, and the relatively high level of $\gamma\delta$ T cells in the sheep species, will allow us to investigate this proposal.

ACKNOWLEDGEMENTS

This work was supported by an AFRC Project Grant (AG15/634). Angela Colston was supported by an AFRC Veterinary Schools Fellowship. We thank Paul Hunt for useful and stimulating discussion.

REFERENCES

- Buergelt, C. D., Hall, C., McKentee, K. & Duncan, J. R. (1978). Pathological evaluation of paratuberculosis in naturally infected cattle. *Vet Pathol* **15**, 196–207.
- Chiodini, R. J. (1993). Abolish *Mycobacterium paratuberculosis* Strain 18. *J Clin Microbiol* **31**, 1956–1957.
- Chiodini, R. J., Van Kruinigen, H. J. & Merkal, R. S. (1984). Ruminant paratuberculosis (Johne's disease): current status and future prospects. *Cornell Vet* **74**, 218–262.
- Cohen, I. R. (1991). Autoimmunity to chaperonins in the pathogenesis of arthritis and diabetes. *Annu Rev Immunol* **9**, 567–589.
- Fiskerstrand, C. E., Roy, D. J., Green, I. & Sargan, D. R. (1992). Cloning, expression and characterisation of ovine interleukins-1 α and β . *Cytokine* **4**, 418–428.
- Gillis, T. P. & Buchanan, T. M. (1982). Production and partial characterisation of monoclonal antibodies to *M. leprae*. *Infect Immun* **37**, 354–361.
- Gilmour, J. E. M., Senior, J. M., Burns, N. R., Esnouf, M. P., Gull, K., Kingsman, S. M., Kingsman, A. J. & Adam, S. E. (1989). A novel method for the purification of HIV-1 p24 protein from hybrid Ty virus-like particles (Ty-VLPs). *AIDS* **3**, 717–723.
- Green, E. P., Tizard, M. L. V., Moss, M. T., Thompson, J., Winterbourne, D. J., McFadden, J. J. & Hermon-Taylor, J. (1989). Sequence and characteristics of IS900, an insertion element identified in a human Crohn's-disease isolate of *Mycobacterium paratuberculosis*. *Nucl Acids Res* **17**, 9063–9073.
- Hance, A. J., Grandchamp, B., Levyfiebault, V., Lecossier, D., Rauzier, J., Bocart, D. & Gicquel, B. (1989). Detection and identification of mycobacteria by amplification of mycobacterial DNA. *Mol Microbiol* **3**, 843–849.
- Hemmingsen, S. M., Woolford, C., Vandervies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W. & Ellis, R. J. (1988). Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature* **333**, 330–334.
- Hunt, P., Colston, A. & Bujdoso, R. (1994). Nomenclature of Mycobacterial stress proteins (65 kDa antigens) and other members of the hsp60 family. *Trends Microbiol* **2**, 298.
- Ikemura, T. (1985). Codon usage and transfer-RNA content in unicellular and multicellular organisms. *Mol Biol Evol* **2**, 13–34.
- Janeway, C. A., Jones, B. & Hayday, A. (1988). Specificity and function of T cells bearing $\gamma\delta$ receptors. *Immunol Today* **9**, 73–76.
- Jindal, S., Dudani, A. K., Singh, B., Harley, C. B. & Gupta, R. S. (1989). Primary structure of a human mitochondrial protein homologous to the bacterial and plant chaperonins and to the 65-kilodalton mycobacterial antigen. *Mol Cell Biol* **9**, 2279–2283.
- Kong, T. H., Coates, A. R. M., Butcher, P. D., Hickman, C. J. & Shinnick, T. M. (1993). *Mycobacterium tuberculosis* expresses two chaperonin-60 homologs. *Proc Natl Acad Sci USA* **90**, 2608–2612.
- Kunze, Z. M., Portaels, F., McFadden, J. J. (1992). Biologically distinct subtypes of *Mycobacterium avium* differ in possession of insertion sequence IS901. *J Clin Microbiol* **30**, 2366–2372.
- Kyhse-Anderson, J. (1984). Electrophoretic blotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polyacrylamide to nitrocellulose. *J Biochem Biophys Meth* **10**, 203–209.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
- Laura, R., Robinson, D. J. & Bing, D. H. (1980). *p*-(Amidinophenyl)methanesulphonyl fluoride, an irreversible inhibitor of serine proteases. *Biochemistry* **19**, 4859–4864.
- Luis, A. M., Alconada, A. & Cuezva, J. M. (1990). The alpha-regulatory subunit of the mitochondrial F1-ATPase complex is a heat-shock protein – identification of two highly conserved amino acid sequences among the α subunits and molecular chaperones. *J Biol Chem* **265**, 7713–7716.
- Mackay, C. R., Maddox, J. F. & Brandon, M. R. (1986). Three distinct subpopulations of sheep T lymphocytes. *Eur J Immunol* **16**, 19–25.
- Magnusson, S., Petersen, T. E., Sottrup-Jensen, L. & Cleays, H. (1975). In *Proteases and Biological Control*, pp. 123–149. Edited by E. Reich, D. Rifkin & E. Shaw. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- McFadden, J. J., Butcher, P. D., Thompson, J., Chiodini, R. & Hermon-Taylor, J. (1987). The use of DNA probes identifying restriction-length polymorphisms to examine the *Mycobacterium avium* complex. *Mol Microbiol* **1**, 283–291.
- Mehra, V., Sweetser, D. & Young, R. A. (1986). Efficient mapping of protein antigenic determinants. *Proc Natl Acad Sci USA* **83**, 7013–7017.
- Moss, M. T., Green, E. P., Tizard, M. L., Malik, Z. P. & Hermon-Taylor, J. (1991). Specific detection of *Mycobacterium paratuberculosis* by DNA hybridization with a fragment of the insertion element IS900. *Gut* **32**, 395–398.
- Nagai, K. & Thogersen, H. C. (1984). Generation of β -globin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli*. *Nature* **309**, 810–812.
- Pluzek, K. J. & Ramlau, J. (1988). Alkaline phosphatase labelled reagents. In *Handbook of Immunoblotting of Proteins* vol. 1, *Technical Descriptions*, pp. 177–188. Edited by O. J. Bjerrum & N. H. H. Heegaard. Boca Raton, FL: CRC Press.
- Rinke de Wit, T. F., Bekelie, S., Osland, A., Miko, T. L., Hermans, P. W. M., van Soolong, D., Drijfhout, J. W., Schoningh, R., Janson, A. A. & Thole, J. E. R. (1992). Mycobacteria contain two GroEL genes – the second *Mycobacterium leprae* GroEL gene is arranged in an operon with GroES. *Mol Microbiol* **6**, 1995–2007.
- Saiki, R. K. (1989). The design and optimisation of the PCR. In

PCR Technology, pp. 25–30. Edited by H. A. Erlich. New York: Stockton Press.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5477.

Seldenrijk, C. A., Drexhage, H. A., Meuwissen, S. G. M. & Meijer, C. J. L. M. (1990). T-cellular immune reactions (in macrophage inhibition-factor assay) against *Mycobacterium paratuberculosis*, *Mycobacterium kansasii*, *Mycobacterium tuberculosis*, and *Mycobacterium avium* in patients with chronic inflammatory bowel disease. *Gut* **31**, 529–535.

Shinnick, T. M. (1987). The 65-kilodalton antigen of *Mycobacterium tuberculosis*. *J Bacteriol* **169**, 1080–1088.

Shinnick, T. M., Sweetser, D., Thole, T., Vanembden, J. & Young, R. A. (1987). The etiologic agents of leprosy and tuberculosis share an immunoreactive protein antigen with the vaccine strain *Mycobacterium bovis* BCG. *Infect Immun* **55**, 1932–1935.

Smith, D. B. & Johnson, K. S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**, 31–40.

Thole, J. E. R., Keulen, W. J., Kolk, A. H. J., Groothuis, D. G., Berwald, I. G., Tiesjema, R. H. & Vanembden, J. D. A. (1987). Characterization, sequence determination, and immunogenicity of a 64-kilodalton protein of *Mycobacterium bovis* BCG expressed in *Escherichia coli* K-12. *Infect Immun* **55**, 1466–1475.

Vary, P. H., Andersen, P. R., Green, E., Hermon-Taylor, J. & McFadden, J. J. (1990). Use of highly specific DNA probes and the polymerase chain-reaction to detect *Mycobacterium paratuberculosis* in Johne's disease. *J Clin Microbiol* **28**, 933–937.

Young, D., Lathigra, R., Hendrix, R., Sweetser, D. & Young, R. A. (1988). Stress proteins are immune targets in leprosy and tuberculosis. *Proc Natl Acad Sci USA* **85**, 4267–4270.

Received 13 May 1994; revised 11 July 1994; accepted 16 August 1994.