Cloning and Sequence Analysis of the 10 kDa Antigen Gene of *Mycobacterium tuberculosis*

By PAUL N. BAIRD, LUCINDA M. C. HALL AND
ANTHONY R. M. COATES*

Department of Medical Microbiology, The London Hospital Medical College, Turner Street,
London E1 2AD, UK

(Received 19 September 1988; revised 14 November 1988; accepted 13 December 1988)

The gene encoding a major protein antigen of *Mycobacterium tuberculosis* has been cloned and sequenced using oligonucleotide probes derived from the N-terminal sequence of the analogous protein from *Mycobacterium bovis* BCG. The gene analysis revealed a sequence encoding a protein of 99 amino acid residues, with a molecular mass of 10·7 kDa. Computer prediction suggests that the protein contains three T-cell-determined epitopes (of which one has been demonstrated experimentally) and three B-cell-determined epitopes. The protein sequence was homologous to two prokaryote heat-shock proteins and the gene possessed heat-shock-like promoter sequences upstream of the initiation codon. A hairpin loop identified in the upstream sequence may also be important in regulation of the gene.

INTRODUCTION

Tuberculosis which is caused by *Mycobacterium tuberculosis* is still a major cause of mortality and morbidity. The World Health Organization estimates that there are three million deaths from the disease annually, mainly in developing countries (IUAT/WHO Study Group, 1982).

Protection against tuberculosis in adolescents has been achieved in some countries by vaccination with *Mycobacterium bovis* Bacillus Calmette Guérin (BCG) (Hart & Sutherland, 1977), although it does not seem to affect the prevalence of the disease in a country (Sutherland, 1981). In other countries, for example India, BCG vaccination has proved ineffective (Tuberculosis Prevention Trial, Madras, 1980). A further problem with the BCG vaccine is that it can cause fatal disseminated disease, although this is very rare (Lotte *et al.*, 1978). If protective antigens of *Mycobacterium* species could be identified, the inadequacies of BCG vaccination could be better understood. As yet, it is unknown which antigens of any mycobacterial species induce immune resistance.

Immunity against *M. tuberculosis* is T cell mediated (Lefford, 1975; Hahn & Kaufman, 1981). Thus an important criterion in seeking protective antigens should be that they induce T-cell-mediated immunity. The best-characterized group of *M. tuberculosis* antigens that bear epitopes recognized by T lymphocytes (T cell epitopes) (Kingston *et al.*, 1987; Young *et al.*, 1986; Kadival *et al.*, 1987; Oftung *et al.*, 1987) are those defined by the 'TB' series of monoclonal antibodies (mAbs), (Coates *et al.*, 1981). The gene for one of these antigens, of molecular mass 65 kDa, has been cloned (Young *et al.*, 1985) and sequenced (Shinnick, 1987): the antigen appears to be a heat-shock protein.

Apart from those antigens defined by the 'TB' series of mAbs, a 10 kDa antigen from *M. bovis* BCG has been purified with the mAb SA12 and its N-terminal sequence of 20 amino acids determined (Minden *et al.*, 1984). This purified 10 kDa antigen induces T-cell-mediated delayed-type hypersensitivity in guinea-pigs vaccinated with *M. bovis* BCG (Minden *et al.*,

---

Abbreviation: mAb, monoclonal antibody.

1984). A cross-reacting antigen of 10 kDa is also present in *M. tuberculosis* (Minden *et al.*, 1984) and is detected in immunoblots by sera from patients with pulmonary tuberculosis (Coates *et al.*, 1989). Sera from tuberculosis patients compete with mAb SA12 for this antigen (A. Guy, personal communication). Taking these factors into account, this antigen may play a role in protection against tuberculosis.

In this paper we report the cloning and nucleotide sequence determination of the gene encoding the 10 kDa antigen of *M. tuberculosis*, and compare the gene flanking sequences with those of other mycobacterial genes.

METHODS

Bacteria and bacteriophage strains. *M. tuberculosis* H37Rv, obtained from the Trudeau Institute (Saranac Lake, New York, USA), was used for all mycobacterial preparations. *Escherichia coli* strain JM105 was the host organism for transformation. M13 phages mp18 and mp19 (Norlander *et al.*, 1983) were used for cloning and the determination of the nucleotide sequence.

Cultures. *M. tuberculosis* H37Rv was grown on Middlebrook's 7H11 (Difco) agar plates for 6–8 weeks at 37 °C, or until large colonies were visible. *E. coli* JM105 was cultured on 2 × YT plates and in 2 × YT broth (Amersham M13 Cloning and Sequencing Handbook).

Mycobacterial DNA preparation. Colonies taken from Middlebrook's 7H11 plates were washed twice and resuspended in 10 ml STE (10 mM-Tris/HCl 100 mM-NaCl, 1 mM-EDTA, pH 8.0). DNA was extracted by the addition of 10000 units of lipase type II (Sigma) and incubation for 90–120 min at 37 °C; then 1 ml 20% SDS was added and the mixture agitated for 2 h at 37 °C, followed by the addition of 2.5 ml chloroform, after which the mixture was left shaking overnight at 37 °C. DNA was then purified by multiple extractions with phenol/chloroform/isoamyl alcohol (25:24:1, by vol.) and precipitated with ethanol.

Restriction enzymes. These were obtained from Northumbria Biologicals or Boehringer Mannheim (BCL) and were used according to the manufacturer's instructions.

Restriction mapping. A restriction map of the 2.25 kb *Sal*I fragment using both single and double digests of *Bam*HI, *Sal*I, *Xho*I, *Eco*RI, *Sma*I, *Pst*I, *Hind*III, *Kpn*I, and *Taq*I was constructed.

Oligonucleotide preparations. A 60mer DNA oligonucleotide probe based on the N-terminal amino acid sequence of *M. bovis* BCG (Minden *et al.*, 1984) was constructed; this probe should recognize the 5' end of the 10 kDa antigen gene of *M. tuberculosis*. Three synthetic primers were made to various regions of the sequence, primer 1 (5' GACACCGCCAAGGAGAAGCC 3'), primer 2 (5' GGCTTCTCCTTGGCGGTGTC 3') and primer 3 (5' AGGCTAGTCCGTAACGCATG 3') (Fig. 3a). Probe and primers were synthesized on a New Brunswick Cyclone DNA synthesizer using phosphoramidite chemistry at the Department of Medical Microbiology, London Hospital Medical College, and used without further purification.

Labelling of probe. The 60mer oligonucleotide used was 5' end-labelled using T4 polynucleotide kinase (Pharmacia) and [γ -³²P]ATP (Amersham) to a specific activity of between 10⁶ and 10⁷ c.p.m. (μ g DNA)⁻¹. Unincorporated [γ -³²P]ATP was removed by the use of a Sephadex G-50 column equilibrated in TE pH 8.0, as described by Maniatis *et al.* (1982).

Agarose gel electrophoresis, hybridization and autoradiography. Electrophoresis of *M. tuberculosis* H37Rv was performed on 1% (w/v) horizontal agarose gels in Tris/borate/EDTA buffer, pH 8.0 (Maniatis *et al.*, 1982). Several tracks of each digest were run and one track from each digest was then transferred to a nylon membrane (Hybond-N; Amersham) by the method of Southern (1975). Filters were then prehybridized and probed at 52 °C in 6 × SSC, 0.5% SDS, 5 × Denhardt's solution and 100 μ g denatured salmon sperm DNA ml⁻¹ according to Maniatis *et al.* (1982) [1 × SSC is 0.15 M-NaCl, 0.015 M-trisodium citrate, pH 7.0; 1 × Denhardt's solution is 0.02% Ficoll (Pharmacia), 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin (fraction V)]. Filters were then washed twice in 2 × SSC, followed by two more washes in 2 × SSC/0.1% SDS, for 15 min each, at 52 °C. Kodak X-AR5 film was exposed to the dried filters overnight with a CAWO intensifying screen at -70 °C.

Regions of homology located on the Southern blot after probing were then used to identify the corresponding regions remaining in the tracks of the agarose gel. These regions were excised from the agarose, electroeluted and purified on DE-52 according to the method of Maniatis *et al.* (1982).

Cloning and DNA sequencing. The regions containing the 600 bp *Sau*3A fragment and the 2.25 kb DNA *Sal*I fragment, purified as described above, were ligated with either a *Bam*HI-digested or a *Sal*I-digested M13mp18 bacteriophage respectively and used to transform *E. coli* JM105 according to the Amersham M13 Cloning and Sequencing Handbook. Clones from sublibraries were taken from overnight plates as duplicate plaque lifts onto nylon filters, the phage lysed and DNA bound to the filters (Maniatis *et al.*, 1982). Positive clones were identified by hybridization with the labelled 60mer DNA probe as described above. Where necessary, secondary screening was conducted until pure plaques containing DNA complementary to the oligonucleotide probe were obtained. Phage DNA from positive plaques was purified from 1.5 ml cultures and sequenced by the chain-termination method (Sanger *et al.*, 1977), using either Klenow polymerase (BCL) or Sequenase (United States Biochemical

- Fig. 1. (a, b) N-terminal amino acid sequences of the 10 kDa antigens of (a) *M. bovis* BCG (Minden *et al.*, 1984) and (b) *M. tuberculosis* H37Rv. Dashes indicate conserved residues. (c) 60mer DNA probes synthesized, based on the N-terminal sequence of the 10 kDa antigen of *M. bovis* BCG, but representing the N-terminal sequence of *M. tuberculosis* H37Rv. (d) Nucleotide sequence of the N-terminus of the 10 kDa gene of *M. tuberculosis*. Dashes indicate conserved residues.

Computer analysis. Computer programs were run either on an Olivetti M24 (Compseq program, University of Geneva, 1985) or on an IBM PS2/model 50 (Staden-Plus software, Amersham, UK 1987), computer. T cell predictions were made using an Orion computer with the algorithm of Rothbard & Taylor (1988) or with a modified version of the algorithm of DeLisi & Berzofsky (1985) using a sample length of 7 amino acids. Hydrophathy values of amino acid residues used in T cell predictions were determined by the method of Kyte & Doolittle (1982).

Cloning of the 10 kDa antigen gene

Southern transfers of restricted *M. tuberculosis* genomic DNA were probed with the 60mer oligonucleotide. A 0.6 kb *Sau*3A fragment and a 2.25 kb *Sal*I fragment (Fig. 2) were excised from untransferred tracks of the gel, purified and cloned directly into M13 bacteriophage (Wei & Surzycki, 1986). Phages containing the fragment of interest were detected in a plaque lift assay. The clone containing the 0.6 kb *Sau*3A fragment was sequenced using an M13 universal primer. This fragment contained the 5' end of the 10 kDa gene encoding the N-terminal part of the protein, but was truncated at the 3' end by a *Bam*HI site in the coding region (Fig. 3a). The sequence of the gene and adjoining regions was completed on the 2.25 kb *Sal*I fragment using three synthetic primers (Fig. 3a).

The coding region of the gene (Fig. 3*b*) is predicted to begin with the initiation codon GTG at position 1 and end with the termination codon TAG, 300 bp later (Fig. 3*b*). The sequence should encode 99 amino acids, starting with alanine, corresponding to a protein of 10673 Da; this is in good agreement with previous estimates of 10 kDa (Minden *et al.*, 1984) and 12 kDa (Engers *et*

Fig. 2. Agarose gel electrophoresis and Southern transfers of restriction-enzyme-digested DNAs from *M. tuberculosis* H37Rv, hybridized with 60mer probes which represent the N-terminus of the 10 kDa antigen of *M. tuberculosis* H37Rv. Lanes 1 and 4, 1 kb yeast molecular size ladder; lanes 2 and 3, *Sal*I-digested H37Rv chromosomal DNA; lanes 5 and 6, *Sau*3A-digested H37Rv chromosomal DNA. Lanes 1, 2, 4 and 5, agarose gel electrophoresis of the DNAs; lanes 3 and 6, blots with ^{32} P-labelled 60mer probe. The sizes of bands homologous to the probe are shown on the right.

al., 1986). Homology was found between this protein and two heat-shock proteins (Baird *et al.*, 1988), the *groES* protein of *E. coli* (Hemmingsen *et al.*, 1988) and the *htpA* gene product of *Coxiella burnetii* (Vodkin & Williams, 1988). This finding suggests that 10 kDa heat-shock proteins are widely conserved amongst bacteria. Whilst the biological significance of such proteins in *M. tuberculosis* is still unknown, it has been shown in other organisms, including *C. burnetii*, that stress or heat will induce their production (Vodkin & Williams, 1988). As a potential candidate for a vaccine, highly conserved proteins might protect against a range of bacterial infections. However, the implication that the 65 kDa antigen of mycobacteria is involved in autoimmune arthritis (Res *et al.*, 1988) should be borne in mind when considering heat-shock proteins as vaccines.

Fig. 3 (on facing page). (a) Restriction map of the 2.25 kb *Sal*I fragment (top thin line), containing 10 kDa antigen gene of *M. tuberculosis* H37Rv (thick line). Synthetic primers 1–3 are marked ×. M13 universal primer is unlabelled. Direction of sequencing is indicated by arrows. Restriction sites: Sa, *Sal*I; X, *Xho*I; B, *Bam*HI; S, *Sma*I. The *Bam*HI site within the coding region corresponds to the *Sau*3A site at the 3' end of the initial *Sau*3A fragment cloned. (b) Complete nucleotide sequence of the 10 kDa antigen gene of *M. tuberculosis* H37Rv. Numbering starts from the initiation codon (GTG). Nucleotides underlined with dashes show homology with the –35 sequence in *E. coli*. Sequences underlined with wavy lines (~~~~) show homology with the –10 sequence of *E. coli*. SD indicates a Shine–Dalgarno sequence. The arrows indicate regions of dyad symmetry. The predicted amino acid sequence coded by this gene is numbered from alanine at position 1. Sequences labelled B indicate predicted B cell epitope sites (Hopp & Woods, 1981). Sequences labelled (///-///-) indicate predicted T cell epitope sites.

(a)

(b)

TACCGGTTTCGCCTATGACGGGACGTCGTGACCGAATGACCTGGTACCGACACGCTGGCAA - 223

CCAGAAGCAAGGGGCGCCCTTAGAGTGTCAGCACTCTCATGTATAGAGTGCTAGATGGCAA - 163

TCGGCTAACCCTGCGTCGGCACCCGCGACGACGGGCCGAGGGCGCGGACGAATGCCGAA - 103

TCACCTGGTAATTTCGGACGGTTCGGCACGCCCGGACCGACCGCCAACCTCCGGTCCGGGC - 43

GAGCGTCCCGGGCTCTGATCCAAATAGTGGAGGGCTCCAATCGTGGCGAAGGTGAACATC 18

SD

LysProLeuGluAspLysIleLeuValGlnAlaAsnGluAlaGluThrThrThrAlaSer 25

AAGCCACTCGAGGACAAGATTCTCGTGCAGGCCAACGAGGCCGAGACCACGACCGCGTCC 78

///-///-///-// /-///-///-///-///

B B B B B B

GlyLeuValIleProAspThrAlaLysGluLysProGlnGluGlyThrValValAlaVal 45

GGTCTGGTCATTCTGACACCGCCAAGGAGAAGCCGCAGGAGGGCACCGTCGTTGCCGTC 138

B B B B B B

GlyProGlyArgTrpAspGluAspGlyGluLysArgIleProLeuAspValAlaGluGly 65

GGCCCTGGCCGGTGGGACGAGGACGGCGAGAAGCGGATCCCGCTGGACGTTGCGGAGGGT 198

B B B B B B ///-///-///-///-///-///-///-///

AspThrValIleTyrSerLysTyrGlyGlyThrGluIleLysTyrAsnGlyGluGluTyr 85

GACACCGTCATCTACAGCAAGTACGGCGGCACCGAGATCAAGTACAACGGCGAGGAATAC 258

LeuIleLeuSerAlaArgAspValLeuAlaValValSerLys*** 99

CTGATCCTGTCGGCACGCGACGTGCTGGCCGTCGTTTCCAAGTAGTAGAGCGTGTTCGGG 318

(300)

CCCCGGCGATCCCCGTGCTCACCACGGTGATTTCGGGGCGGCATGCGTTACGGACTAGC 378

CTGCGTAGAGA 389

The first 20 amino acids of the protein are identical to the N-terminal sequence from *M. bovis* BCG (Minden *et al.*, 1984) except for a glutamic acid to glutamine change at amino acid 15. Our sequence is also in agreement with the first 15 amino acids of an N-terminal amino acid sequence described by Patarroyo *et al.* (1986), for an *M. tuberculosis* antigen (MTP-13), but from position 16 the sequences diverge.

Prediction of T and B cell epitopes

The method of Rothbard & Taylor (1988) predicted three possible T cell sites located at amino acid positions 6–9 (Lys-Pro-Leu-Glu), 11–15 (Lys-Ile-Leu-Val-Gln) and 57–64 (Arg-Ile-Pro-Leu-Asp-Val-Ala-Glu) (Fig. 3*b*). The latter site is composed of two overlapping motifs: 57–61 (Arg-Ile-Pro-Leu-Asp) and 61–64 (Asp-Val-Ala-Glu), (Fig. 3*b*). The algorithm used identifies the core of any potential T cell epitope, but it should be noted that the complete T cell site is normally 8–12 residues in length. Rothbard & Taylor (1988) include threonine as a hydrophobic residue in T cell epitope prediction, whereas Kyte & Doolittle (1982) assigned this residue a hydropathy value of -0.7 , which suggests that it is not hydrophobic. If threonine was included in the hydrophobic core, then another motif at amino acid positions 31–34 (Asp-Thr-Ala-Lys) would also be predicted. No epitopes could be defined using the modified DeLisi & Berzofsky (1985) algorithm. The predicted T cell epitopes at amino acid positions 6–9 and 11–15 overlap with a synthetic peptide (BCG-a) (Minden *et al.*, 1986) which elicited a T-cell-mediated response in the skin of BCG-vaccinated guinea pigs. This peptide comprised the first 13 amino acids of the 10 kDa antigen of *M. bovis* BCG and was antigenic only when polymerized.

The method of Hopp & Woods (1981) predicted three B cell epitope sites at amino acid positions 6–11 (Lys-Pro-Leu-Glu-Asp-Lys), 34–39 (Lys-Glu-Lys-Pro-Gln-Glu), and 51–56 (Asp-Glu-Asp-Gly-Glu-Lys) (Fig. 3*b*), of which 51–56 is the most hydrophilic. Minden *et al.* (1986) showed that a synthetic peptide derived from the N-terminal 13 amino acids of the 10 kDa antigen which overlaps with our predicted 6–11 site could elicit an antibody response in rabbits. This suggests that the N-terminus of the protein possesses a B cell epitope.

Features of the non-coding region

Sequences upstream of the proposed initiation codon at +1 (Fig. 3*b*) were compared with sequences involved in initiation of transcription and translation in *E. coli*. At position -14 , there is a perfect match with the *E. coli* ribosome-binding site GGAGG (Shine & Dalgarno, 1974). Three sequence motifs at positions beginning -191 , -182 , and -180 have four bases out of six homologous with the -10 or TATAAT box sequence of *E. coli* (Pribnow, 1975). At position -203 , there is a three-out-of-three base homology with the highly conserved TTG of the -35 sequence found in *E. coli* (Hawley & McClure, 1983). This comparison suggests that the start of transcription should occur around 170 bp upstream of the initiation codon.

Examination of the upstream sequences of three other mycobacterial genes encoding the 65 kDa antigen from different species (Fig. 4) shows that each has a potential RNA polymerase binding site at 160–190 bp before the proposed initiation codon at a similar distance to that of the 10 kDa antigen gene. A second set of -35 - and -10 -like sequences occurs 26 bp upstream of the predicted coding region in these 65 kDa antigen genes, but not in the 10 kDa antigen gene (Fig. 4). The 65 kDa antigen, like the 10 kDa antigen, has homology to heat-shock proteins of other organisms (Young *et al.*, 1988), so it is of interest to compare the upstream sequences of these genes. In *E. coli* heat-shock genes, a relatively long 5' leader is also found in the *groE* operon, whilst the *dnaK* operon gene has three sites for initiation of transcription, at 115, 40 and 19 nucleotides upstream of the ATG (Gowing *et al.*, 1985). However, a consensus established for the RNA polymerase binding site of heat shock promoters in *E. coli* differs from non-heat-shock promoters, particularly in the -10 region (Cowing *et al.*, 1985). A match of five bases out of eight with this -10 consensus is present at the proposed RNA polymerase binding site of the *M. tuberculosis* 10 kDa antigen gene (Fig. 5). To determine the significance of such consensus sequences in mycobacteria, transcription studies are clearly required.

As well as the similarities to *E. coli* upstream sequences, the mycobacterial genes have other shared sequence elements. Firstly, a pentamer of GATCC is found, which lies close to the

	Homology to RNA polymerase binding site	Homology to RNA polymerase binding site	GATCC	SD	Met
<i>M. tuberculosis</i> 10 kDa	TTGACT-17-TAGACT-----148	TTGACT-17-TAGACT-----148	-----GATCC-7-GGAGG-----9	-----GTG	
<i>M. tuberculosis</i> 65 kDa*	TTGCAC-18-AAGAAAT---133	TTGCAC-18-AAGAAAT---133	---TTTCAT-4-GATCC---GGAGG---12	---ATG	
<i>M. bovis</i> 64 kDa†	TTGCAC-18-AAGAAAT---133	TTGCAC-18-AAGAAAT---133	---TTTCAT-4-GATCC---GGAGG---12	---ATG	
<i>M. leprae</i> 65 kDa‡	TTGCAC-23-AATGATCC-105	TTGCAC-23-AATGATCC-105	---CTTCAT-9	---GGAGG---12	---ATG
<i>E. coli</i> consensus§		TTGACA--15-21--TATAAT	-----GGAGG--12-18-----ATG		

Fig. 4. Comparison of mycobacterial upstream regulatory sequences showing sequence similarity with prokaryotic promoter consensus sequences. SD indicates Shine-Dalgarno site. Met indicates initiation codon. *Shinnick (1987). †Thole *et al.* (1987). ‡Mehra *et al.* (1986). §Hawley & McClure (1983).

Fig. 5

Fig. 6. Comparison of upstream heat-shock promoter sequence of the 10 kDa antigen gene of *M. tuberculosis* with the homologous *groES* gene promoter and heat-shock consensus sequence. *Hemmingsen *et al.* (1988). †Cowing *et al.* (1985).

Fig. 7. Comparison of upstream hairpin loop structures located 5' to the gene: (1) *M. leprae* 65 kDa, (2) *M. tuberculosis* 10 kDa, (3) *M. tuberculosis* 65 kDa and *M. bovis* 64 kDa (same sequence). Numbering at the apex of the loop indicates central position. Numbering on stems is in relation to the stem of (1).

Shine-Dalgarno sequence in the 10 kDa and 65 kDa antigen genes of *M. tuberculosis*, the 64 kDa antigen gene of *M. bovis* BCG, and further upstream in the 65 kDa antigen gene of *M. leprae* (Fig. 4). Secondly, in all of these genes potential hairpin loop structures sharing a common stem sequence were found 5' to the start of the coding regions (Fig. 6). Such conservation in sequence and structure may suggest a regulatory role for these elements.

Downstream of the 10 kDa antigen gene in the 3' flanking region, an additional 88 bases have been sequenced. An inverted repeat occurs at positions 316 and 350 which is G/C rich and is likely to be involved in transcription termination (Rosenberg & Court, 1979).

In conclusion, this paper describes for the first time the entire DNA sequence of the gene encoding the 10 kDa antigen of *M. tuberculosis*. The derived protein sequence contains one known T cell epitope and an additional predicted T cell site. It is not known whether this antigen will be protective, but knowledge of the sequence will enable large quantities of either synthetic or recombinant antigen to be produced for studies in animals.

We thank Dr P. Minden for BCG-a and MAB SA12, Mr J. Cookson (Hill Centre, London Hospital Medical College) for the computer searches of T cell epitopes and American Cyanamid for financial support.

REFERENCES

- BAIRD, P. N., HALL, L. M. C. & COATES, A. R. M. (1988). A major protein antigen from *Mycobacterium tuberculosis* which is homologous to the heat shock proteins groES from *E. coli* and the htpA gene product of *Coxiella burnetii*. *Nucleic Acids Research* **16**, 9047.
- BRADLEY, S. G. (1973). Relationships among mycobacteria and nocardiae based upon deoxyribonucleic acid reassociation. *Journal of Bacteriology* **113**, 645-651.
- COATES, A. R. M., HEWITT, J., ALLEN, B. W., IVANJI, J. & MITCHISON, D. A. (1981). Antigenic diversity of *Mycobacterium tuberculosis* and *Mycobacterium bovis* detected by means of monoclonal antibodies. *Lancet* **ii**, 167-169.
- COATES, A. R. M., NICOLAI, H., PALLAN, M., GUY, A., CHAPARAS, S. D. & MITCHISON, D. A. (1989). The 45 kilodalton molecule of *M. tuberculosis* identified by immunoblotting and monoclonal antibodies as antigenic in patients with tuberculosis. *British Journal of Experimental Pathology* (in the Press).
- COLLINS, D. M. & DE LISLE, G. W. (1985). DNA restriction endonuclease analysis of *Mycobacterium tuberculosis bovis* and other members of the tuberculosis complex. *Journal of Clinical Microbiology* **21**, 562-564.
- COWING, D. W., BARDWELL, J. C. A., CRAIG, E. A., WOOLFORD, C., HENDRIX, R. W. & GROSS, C. A. (1985). Consensus sequence for *Escherichia coli* heat shock gene promoters. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 2679-2683.
- DELISI, C. & BERZOFKY, J. A. (1985). T-cell antigenic sites tend to be amphipathic structures. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 7048-7052.
- ENGERS, H. D., HOUBA, V., BENNEDSEN, J., BUCHANAN, T. M., CHAPARAS, S. D., KADIVAL, G., CLOSE, O., DAVID, J. R., VAN EMBDEN, J. D. A., GODAL, T., MUSTAFA, S. A., IVANJI, J., YOUNG, D. B., KAUFMANN, S. H. E., KHOMENKO, A. G., KOLK, A. H. J., KUBIN, M., LOUIS, J. A., MINDEN, P., SHINNICK, T. M., TRNKA, L. & YOUNG, R. A. (1986). Results of a World Health Organisation-sponsored workshop to characterise antigens recognised by mycobacterium-specific monoclonal antibodies. *Infection and Immunity* **51**, 718-720.
- HAHN, H. & KAUFMANN, S. H. E. (1981). The role of cell-mediated immunity in bacterial infections. *Reviews of Infectious Diseases* **3**, 1221-1250.
- HART, P. D. & SUTHERLAND, I. (1977). BCG and vole bacillus vaccines in the prevention of tuberculosis in adolescence and early adult life. *British Medical Journal* **2**, 293-295.
- HAWLEY, D. K. & MCCLURE, W. R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Research* **8**, 2237-2255.
- HEMMINGSEN, S. M., WOOLFORD, C., VAN DER VIES, S. M., TILLY, K., DEENIS, D. T., GEORGEPOULOS, C. P., HENDRIX, R. W. & ELLIS, J. (1988). Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature, London* **333**, 330-334.
- HOPP, T. P. & WOODS, K. R. (1981). Prediction of protein antigenic determinants from amino acid sequences. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 3824-3828.
- IMAEDA, T. (1985). Deoxyribonucleic acid relatedness among selected strains of *Mycobacterium tuberculosis*, *Mycobacterium bovis*, *Mycobacterium bovis* BCG, *Mycobacterium microti*, and *Mycobacterium africanum*. *International Journal of Systematic Bacteriology* **35**, 147-150.
- IUAT/WHO STUDY GROUP (1982). Tuberculosis control. *Tubercle* **63**, 157-169.
- KADIVAL, G. V., CHAPARAS, S. D. & HUSSONG, D. (1987). Characterization of serologic and cell-mediated reactivity of a 38 kDa antigen isolated from *Mycobacterium tuberculosis*. *Journal of Immunology* **139**, 2447-2451.
- KINGSTON, A. E., SALGAME, P. R., MITCHISON, N. A. & COLSTON, M. J. (1987). Immunological activity of a 14-kilodalton recombinant protein of *Mycobacterium tuberculosis* H37Rv. *Infection and Immunity* **55**, 3149-3154.

- KYTE, J. & DOOLITTLE, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157**, 105–132.
- LEFFORD, M. J. (1975). Transfer of adoptive immunity to tuberculosis in mice. *Infection and Immunity* **11**, 1174–1181.
- LOTTE, A., WASZ-HOCKERT, O., POISSON, N. & DUMITRESCU, D. (1978). Complications induced by BCG vaccination: a retrospective study. *Bulletin of the International Union against Tuberculosis* **53**, 114–116.
- MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- MEHRA, V., SWEETSER, D. & YOUNG, R. (1986). Efficient mapping of protein antigenic determinants. *Proceedings of the National Academy of Sciences of the United States of America* **83**, 7013–7017.
- MINDEN, P., KELLEHER, P. J., FREED, J. H., NIELSEN, L. D., BRENNAN, P. J., MCPHERON, L. & MCCLATCHY, J. K. (1984). Immunological evaluation of a component isolated from *Mycobacterium bovis* BCG with a monoclonal antibody to *M. bovis* BCG. *Infection and Immunity* **46**, 519–525.
- MINDEN, P., HOUGHTEN, R. A., SPEAR, J. R. & SHINNICK, T. M. (1986). A chemically synthesized peptide which elicits humoral and cellular immune responses to mycobacterial antigens. *Infection and Immunity* **53**, 560–564.
- NORRANDER, J., KEMPE, T. & MESSING, J. (1983). Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* **26**, 101–106.
- OFTUNG, F., MUSTAFA, A. S., HUSSON, R., YOUNG, R. A. & GODAL, T. (1987). Human T cell clones recognise two abundant *Mycobacterium tuberculosis* protein antigens expressed in *Escherichia coli*. *Journal of Immunology* **138**, 927–931.
- PATARROYO, M. E., PARRA, C., PINILLA, C., DEL PORTILLO, P., LOZADA, D., ORAMAS, M., TORRES, M., CLAUDIO, P., RAMIREZ, C., FAJARDO, N., CRUZ, N. & JIMENEZ, C. (1986). Immunogenic synthetic peptides against *Mycobacterium tuberculosis*. In *Vaccines 86*, pp. 219–224. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- PRIBNOW, D. (1975). Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. *Proceedings of the National Academy of Sciences of the United States of America* **72**, 784–788.
- RES, P. C. M., SCHAAR, C. G., BREEDVELD, F. C., VAN EDEN, W., VAN EMBDEN, J. D. A., COHEN, I. R. & DE VRIES, R. R. P. (1988). Synovial fluid T cell reactivity against 65 kD heat shock protein of mycobacteria in early chronic arthritis. *Lancet* **ii**, 478–480.
- ROSENBERG, M. & COURT, D. (1979). Regulatory sequences involved in the promotion and termination of RNA transcription. *Annual Review of Genetics* **13**, 319–353.
- ROTHBARD, J. B. & TAYLOR, W. R. (1988). A sequence pattern common to T cell epitopes. *EMBO Journal* **7**, 93–100.
- SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463–5467.
- SHINE, J. & DALGARNO, L. (1974). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences of the United States of America* **71**, 1342–1346.
- SHINNICK, T. M. (1987). The 65-kilodalton antigen of *Mycobacterium tuberculosis*. *Journal of Bacteriology* **169**, 1080–1088.
- SOUTHERN, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503–517.
- SUTHERLAND, I. (1981). The epidemiology of tuberculosis – is prevention better than cure? *Bulletin of the International Union against Tuberculosis* **56**, 127–134.
- THOLE, J. E. R., KEULEN, W. J., DE BRUYN, J., KOLK, A. H. J., GROOTHUIS, D. G., BERWALD, L. G., TIESJEMA, R. H. & VAN EMBDEN, J. D. A. (1987). Characterization, sequence determination, and immunogenicity of a 64-kilodalton protein of *Mycobacterium bovis* BCG expressed in *Escherichia coli* K-12. *Infection and Immunity* **55**, 1466–1475.
- TUBERCULOSIS PREVENTION TRIAL, MADRAS (1980). Trial of BCG vaccines in South India for tuberculosis prevention. *Indian Journal of Medical Research* **72** (Suppl.), 1–74.
- VODKIN, M. H. & WILLIAMS, J. C. (1988). A heat shock operon in *Coxiella burnetii* produces a major antigen homologous to a protein in both mycobacteria and *Escherichia coli*. *Journal of Bacteriology* **170**, 1227–1234.
- WEI, Y.-G. & SURZYCKI, S. J. (1986). Screening recombinant clones containing sequences homologous to *Escherichia coli* genes using single-stranded bacteriophage vector. *Gene* **48**, 251–256.
- YOUNG, D., KENT, L., REES, A., LAMB, J. & IVANJI, J. (1986). Immunological activity of a 38-kilodalton protein purified from *Mycobacterium tuberculosis*. *Infection and Immunity* **54**, 177–183.
- YOUNG, D., LATHIGRA, R., HENDRIX, R., SWEETSER, D. & YOUNG, R. A. (1988). Stress proteins are immune targets in leprosy and tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 4267–4270.
- YOUNG, R. A., BLOOM, B. R., GROSSKINSKY, C. M., IVANJI, J., THOMAS, D. & DAVIS, R. W. (1985). Dissection of *Mycobacterium tuberculosis* antigens using recombinant DNA. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 2583–2587.