Journal of Biochemistry and Molecular Biology, Vol. 34, No. 4, July 2001, pp. 342-346

JOURNAL OF
Biochemistry and Molecular Biology

Cloning and Characterization of the *Mycobacterium bovis* BCG *panB* Gene Encoding Ketopantoate Hydroxymethyltransferase

Jin Koo Kim, Kwang Dong Kim, Jong-Seok Lim, Hee Gu Lee, Sang Jae Kim[§], Sang-Hyun Cho[§], Won-Hwa Jeong[‡], In Seong Choe, Tai Wha Chung, Sang-Gi Paik[†] and Yong-Kyung Choe*

Cell Biology Laboratory, Korea Research Institute of Bioscience & Biotechnology (KRIBB), P.O. Box 115, Yusong, Taejon 305-600, [§]Korea Institute of Tuberculosis, KNTA, Seoul 137-140, [‡]Boryung Biopharma Co., Chungbuk 365-834, [†]Department of Biology, Chungnam National University, Taejon 305-764, Korea

Received 20 April 2001, Accepted 14 May 2001

The *Mycobacterium bovis* BCG *panB* gene, encoding ketopantoate hydroxymethyltransferase (KPHMT), was cloned from a λgt11 genomic library and sequenced. The DNA sequence encodes a protein that contains 281 amino acid residues (*M*, 29,337) with a high similarity to the KPHMTs. Subcloning of a 846 bp open reading frame (ORF), but not a 735 bp ORF, into the vector pUC19 led to complementation of the *panB* mutant of *Escherichia coli*. The BCG *panB* gene was overexpressed in *E. coli* and the KPHMT purified to homogeneity. The recombinant protein was further confirmed by an enzymatic assay.

Keywords: BCG, Ketopantoate hydroxymethyltransferase (KPHMT), *panB*, Pantothenic acid

Introduction

Tuberculosis, which is caused by *Mycobacterium tuberculosis*, remains one of the major health problems in developing countries. Kochi (1991) estimated that about one-third of the worlds population, or about 1.7 billion people, is infected with *M. tuberculosis*. Each year, there are over 8 million new cases of tuberculosis and 2.9 million deaths from the disease. Tuberculosis accounts for 6.7% of all deaths and over 25% of avoidable adult deaths in the developing world. The pandemic of human immunodeficiency virus (HIV) infection, the evidence of an association between tuberculosis and HIV, as well as the emergence of increased numbers of multi-drug resistant strains of *M. tuberculosis*, have caused worldwide public concern (Bloom and Murrat, 1992).

*To whom correspondence should be addressed.
Tel: 82 42-860-4184; Fax: 82-42-860-4593
E-mail: ykccoruk@mail.kribb.re.kr

Mycobacterium bovis bacille Calmette-Guerin (BCG) is the most widely used live vaccine against tuberculosis (Bloom and Fine, 1994). It is used for superficial bladder cancer immunotherapy (Morales and Nickel, 1992). Although we still do not understand the efficacy, behavior, and impact of BCG, genetically modified BCG could secret functional murine cytokines (Murray *et al.*, 1996), and elicit systemic and mucosal immunity (Langermann *et al.*, 1994a,b). Major obstacles to the development of BCG as a multi-valent vaccine vehicle, or immune modulator, are its slow growth, the little information available about molecular genetics, as well as the lack of useful genetic systems to manipulate BCG. The application of the recombinant DNA (Kim *et al.*, 1999) and monoclonal antibody technology to mycobacteria has brought forth an explosion of studies, which are investigating mycobacterial genes and proteins.

Pantothenic acid (synthesized by microorganisms and plants, but not mammals) is a vitamin of the B group and the precursor of coenzyme A. Ketopantoate hydroxymethyltransferase (KPHMT) catalyzes the first step in the biosynthesis of pantothenic acid. Although the pantothenate that contains cofactors performed important functions in lipid synthesis, little is known about the molecular biological aspect of the pantothenate pathway. The isolation of the *panB* gene that encodes the KPHMT, and the sequence of this gene, was first reported from *E. coli* (Jones *et al.*, 1993). The *panB* gene was isolated by functional complementation of a *panB* mutant strain. The identity of this gene product was confirmed by purification of the enzyme product.

In this paper, we report the cloning, sequencing, and overexpression of the *panB* gene of *M. bovis* BCG, and the polymerase chain reaction (PCR) amplification of the *panB* gene from *M. tuberculosis* of Korean tuberculosis patients. This is the first report to show the sequence of a *panB* gene from mycobacteria.

Materials and Methods

Bacterial strains, plasmids, and media The *Mycobacteriumbovis* BCG Pasteur strain 1173-P2 was grown on a Sauton medium and used to prepare the BCG genomic DNA. *Mycobacterium tuberculosis* KIT 10468 was isolated from Korean tuberculosis patients (Kim *et al.*, 1996). *Escherichia coli* Hfr3000 YA139, a *panB* mutant derivative of *E. coli* K-12, was used to identify the *panB* gene by functional complementation. The mutant was grown on a M9 minimal medium (1X M9 salt, 20 mM glucose, 1 mM $MgSO_4$, 0.01% thiamine, and 0.1 mM $CaCl_2$). *E. coli* Y1090 (Stratagene, CA, USA), DH5α, and BL21 (λDE3) were used as hosts for cloning. They were cultured in a LB broth, or on a LB agar plate that was supplemented with 100 μg/ml ampicillin. For solid media, 1.5% bacto-agar was added to the broth. Plasmid pUC19 was used as a vector for subcloning, nucleotide sequencing, and functional complementation. Plasmid pRSETB (Invitrogen, CA, USA), which has a N-terminal polyhistidine (6xHis) tag for rapid purification, was used for overexpression.

Cloning, sequencing, and functional complementation of the *panB* gene BCG-λgt11 genomic library construction and screening with a rabbit antiserum against malonamidase was performed as described previously (Kim *et al.*, 1997a). For plaque hybridization, the probe was labeled by random priming using a DIG DNA labeling kit (Boehringer Mannheim, Germany). Recombinant phage DNA was prepared from positive clones. The subsequent subcloning into pUC19 was carried out as described by Sambrook, J. and Russell, D. W. (2001). Both strands of DNA from the E213437 clone were sequenced by the dideoxy chain termination method (Sanger *et al.*, 1977) using a Pharmacia Deaza T7 sequencing kit. Analysis of the nucleotide and amino acid (aa) sequence of the insert DNA was performed using DNASIS and the BLAST program. A comparative analysis of the deduced aa sequence was carried out using the PILEUP program of the Genetics Computer Group package.

The promoterless *panB* coding region was amplified by PCR using the oligonucleotides 5'-CCC**GGATCC**AGTGATGTCTGAGC-3' (upstream primer, *Bam*HI site in bold) and 5'-CGC**GAATTC**TGGTCAGAAACTGTG-3' (downstream primer, *Eco*RI site in bold). The two oligonucleotides were also used as primers for PCR amplification with *M. tuberculosis* chromosomal DNA. The upstream primer for the ORF region of 735bp 5'-CCCGGATCCCATGCTGACGGCCT-3' (*Bam*HI site in bold) was used. After digestion with *Bam*HI and *Eco*RI, the PCR product was cloned into pUC19 and pRSETB, respectively. Transformation-competent *E. coli* Hfr3000 YA139 cells, prepared by the calcium chloride method (Sambrook, J. and Russell, D. W., 2001), were transformed with pUC19-panB, pUC19-Δ5panB, pUC19 alone, and plasmid pAL01 (pUC19-*E. coli panB*), respectively. The transformants that were grown on the LB agar plate were transferred to a M9 minimal medium. Only the complemented transformants were able to grow on this medium when incubated at 37°C.

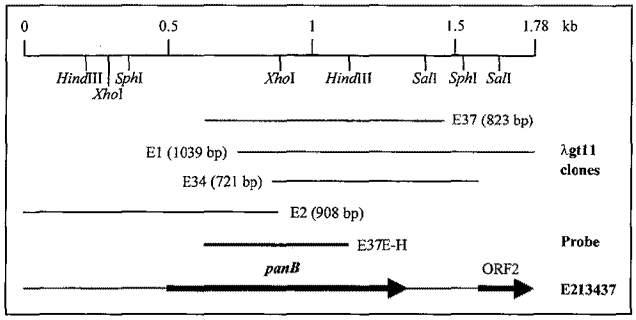

Fig. 1. Physical map and cloning strategy of the *panB* gene from *Mycobacterium bovis* BCG. The arrows indicate the direction of translation of ORFs.

Overexpression and purification The plasmid pRSETB-panB was transformed into *E. coli* BL21 (λDE3), which carried in its genome the gene encoding T7 RNA polymerase under the control of an IPTG inducible promoter. The transformant cells were grown at 37°C in 250 ml of a LB broth until an A_{600} of 0.5 (Jeong *et al.*, 2001). Then IPTG (final concentration of 1 mM) was added. After another 4 hr of growth, the cells were centrifuged and resuspended

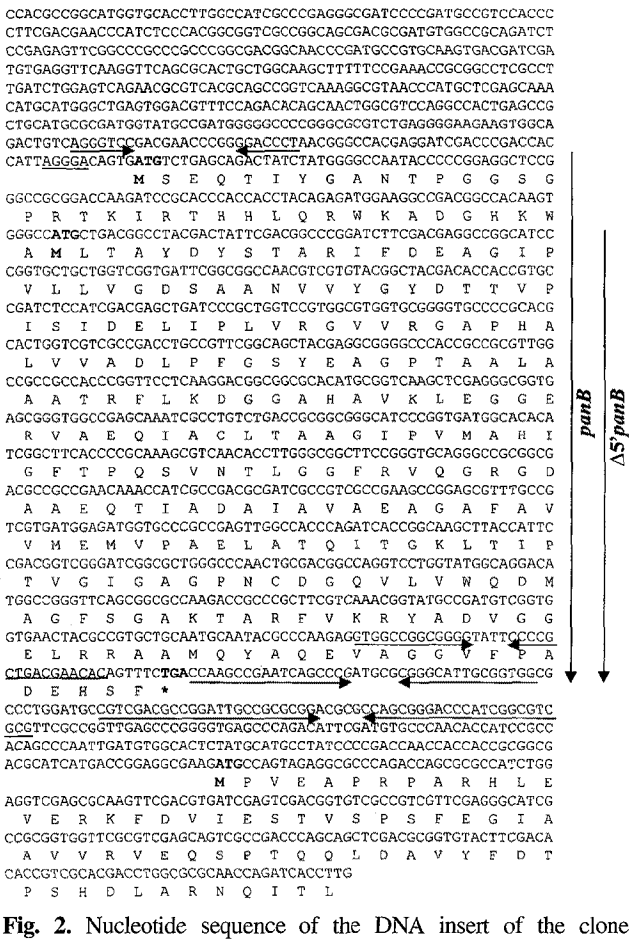

Fig. 2. Nucleotide sequence of the DNA insert of the clone E213437 encoding *Mycobacterium bovis* BCG *panB* gene and partial ORF. The ATG start codon for *panB* is bolded. An asterisk marks the stop codon. The deduced amino acid sequence is shown below the nucleotide sequence in a single letter code. The putative ribosome-binding site (RBS) for the *panB* gene is underlined. This sequence has been assigned the GenBank accession number U57435.

in 10 ml of a sonication buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM PMSF, and 10 μg/ml leupeptin). The cells were lysed by sonication and centrifuged at 12,000 × *g* for 10 min at 4°C to pellet any insoluble material. The soluble hexahistidine-tagged KPHMT was purified using TALON™ metal affinity resin (Clontech, USA), according to the manufacturers instructions. The soluble fraction was loaded on the nickel affinity column, washed with a sonication buffer, and eluted with a sonication buffer that contained 100 mM imidazole (Champreda *et al.*, 2000). The proteins were separated on 12% SDS polyacrylamide gel to estimate yields and purity. The protein bands were visualized after Coomassie blue staining.

Enzyme assays Enzyme assays for KPHMT were performed by measuring the incorporation of $H^{14}CHO$ (10 mCi/mmol; New England Nuclear Corp., USA) into ketopantoate by the method of Teller *et al.* (1976), as modified by Cronan *et al.* (1982). The assay mixture [100 mM potassium phosphate, pH 6.8, 1 mM $MgSO_4$, 0.5 mM tetrahydrofolate (Sigma Chemical Co., USA), 0.5 mM $H^{14}CHO$, and 5 mM sodium α-ketoisovalerate (Sigma)] was incubated at 37°C for 5 min to allow methylene tetrahydrofolate to form. The reaction (50 μl) was then initiated by the addition of an enzyme. After 10 min at 37°C in 100 mm screwcap tube, 0.5 ml of 0.6 M sodium acetate buffer, pH 4.5 (containing 0.4 M HCHO), and 0.4 5ml of 0.4 M dimedon (Sigma) in 50% ethanol were added. The tubes were capped, placed in boiling water for 10 min, and cooled. After 3 ml of chloroform was added, the tubes were centrifuged, and 0.2 ml of the upper phase was removed for scintillation counting. The dimedon was used to remove unreacted $H^{14}CHO$ from the water phase. Radioactivity was measured with a Beckman liquid scintillation analyzer with an efficiency of 90%. Protein was determined by a protein assay kit (Bio-Rad, USA) with bovine serum albumin as a standard.

Results and Discussion

Cloning and sequencing of the BCG *panB* gene The genomic library of *Mycobacterium bovis* BCG, strain 1173-P2, in the λgt11 phage was screened on the basis of recognition by rabbit antiserum against malonamidase (Kim *et al.*, 1997a). Several positive clones were isolated and sequenced (Kim *et al.*, 1997a,b). The nucleotide sequence analysis of the insert DNA using the DNASIS program showed that the DNA fragments of E1, E34, and E37 clones shared the same sequence, as well as having incomplete ORF (Fig. 1). To isolate the DNA fragments that contained the complete ORF, we screened the phage library with the 0.5 kb *Eco*RI-*Hind*III fragment of the E37 clone as a probe. In all of the positive clones examined, the complete nucleotide sequence of a 1,778 bp fragment was determined on both strands (Fig. 2).

Sequence analysis of the BCG *panB* gene The BCG DNA fragment, named E213437, contained two ORFs. The first ORF, which was 846 bp in length, had a 53% nucleotide sequence similarity with the sequence of the *E. coli panB* gene. The total G+C content of the cloned BCG *panB* gene was 67.2%. An analysis of the codon usage indicated a strong preference for G and C in the third base position. This pattern is similar to those of other mycobacterial genes (Andersson and Sharp, 1996). A putative ribosome-binding site (RBS, AGGGA) was located 6 bp upstream from the ATG start codon at position 495. Three inverted repeat sequences were found near the TGA termination codon (Fig. 2). Their calculated free energy values for the formation of the stem and loop structure were –21.7 kcal/mol (position 1299-1331), –26.2 kcal/mol (1341-1378), and –42.0 kcal/mol (1391-1443), respectively. In addition to these inverted sequences, a short inverted sequence (5'-CAGGGTCCGACGAACCCGG GGACCCTA-3', 427-453) was detected upstream from the

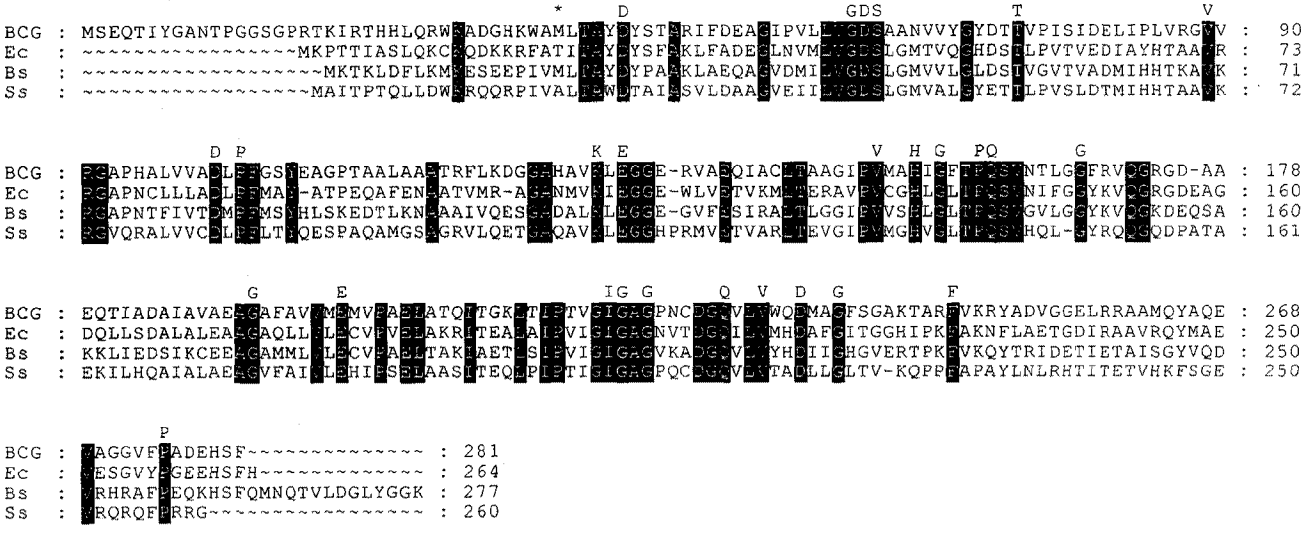

Fig. 3. Amino acid sequence alignment of KPHMT from *Mycobacterium bovis* BCG (BCG, this work), *Escherichia coli* (Ec, X65538), *Bacillus subtilis* (Bs, L47709), and *Synechocystis sp.* (Ss, D64006). The identical residues and conserved residues are shown by a single letter above the sequence and in a black box, respectively. The identical and conserved residues are chosen by alignment of 16 KPHMTs that are available from the GenBank, but only 4 sequences are shown. An asterisk marks the translation start of Δ5'panB.

Fig. 4. Functional complementation of the *panB* mutant strain *E. coli* Hfr3000 YA139 on minimal media by transformation with pUC19 (A), pAL01 (B), pUC19-Δ5'panB (C), and pUC19-panB (D).

RBS (Fig. 2). This inverted sequence showed high homology with that (5'-CAGGGTCAGACGAACCCGGGGACCCTA-3', 10800-10826, GenBank accession number L78818) from the promoter of *Mycobacterium leprae*. However, its role in transcription regulation has not yet been described. The protein that was deduced from the ORF1 nucleotide sequence was composed of 281 aa with a predicted molecular mass of 29,337 Da. The deduced aa sequence was 44%, 38%, and 44% identical to that of *panB* from *E. coli*, *Bacillus subtilis*, and *Synechocystis sp.*, respectively. An amino acid sequence alignment of the KPHMT proteins from 16 bacteria revealed that 27 amino acids are identical in all of the KPHMT proteins (Fig. 3). The second partial ORF was 195 bp long and had no homology with known sequences. The *panB* gene (U57435; *M. tuberculosis* H37Rv, Z70692) of *M. tuberculosis* KIT 10468, isolated from Korean tuberculosis patients, was shown to be identical to that of BCG.

Functional complementation of the BCG *panB* gene In order to confirm the predicted translational start codon, and the complementing region in the ORF1, two putative coding regions of ORF1 were subcloned into the vector pUC19, resulting in pUC19-panB and pUC19-Δ5'panB (Fig. 2). The plasmids were transformed into the *panB* mutant strain *E. coli* Hfr3000 YA139. Plasmid pUC19-panB complemented this mutant strain, whereas pUC19-Δ5'panB did not (Fig. 4). These results showed that the 846 bp ORF1 was the *panB* gene and N-terminal region (37 aa) of BCG KPHMT was essential for the function of this enzyme, although it had some low similarity to that region of KPHMT from other bacteria (Fig. 3).

Overexpression and purification of BCG KPHMT To evaluate the biological activity of the BCG *panB* gene product, the BCG *panB* gene was amplified with two primers

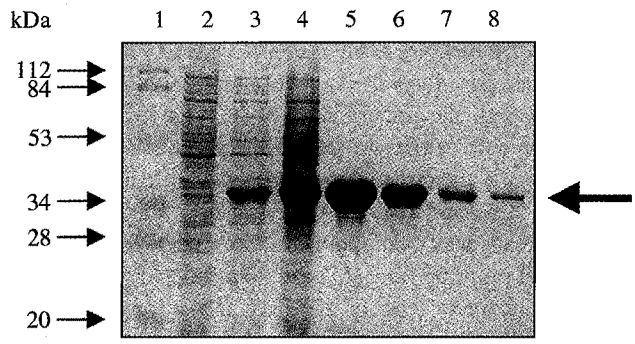

Fig. 5. SDS-PAGE of the recombinant BCG KPHMT that is overexpressed in *Escherichia coli* BL21 (λDE3). Recombinant hexahistidine-tagged KPHMT was overexpressed and purified, as described in Materials and Methods. Lane 1: prestained molecular weight marker, lane 2: total protein from *E. coli* BL21 (λDE3)/ pRSETB-panB, lane 3: total protein from *E. coli* BL21 (λDE3)/ pRSETB-panB induced with IPTG, lane 4: soluble protein from *E. coli* BL21 (λDE3)/ pRSETB-panB induced, lane 5-8: recombinant hexahistidine-tagged KPHMT from elution fractions. The position of each recombinant KPHMT is indicated on the right side by an arrow.

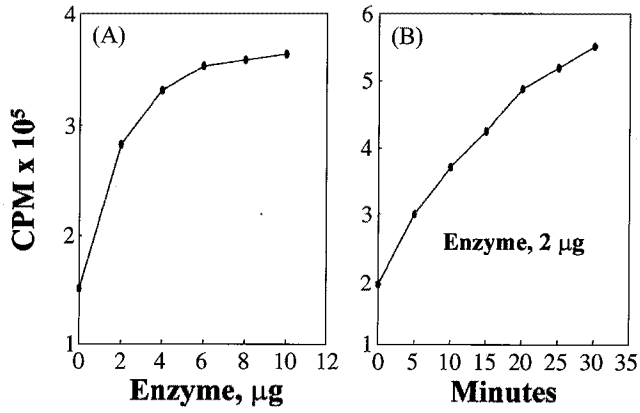

Fig. 6. BCG ketopantoate hydroxymethyltransferase activity on enzyme concentration (A) and time-course (B).

(materials and methods) using BCG chromosomal DNA and cloned into pRSETB, thereby generating pRSETB-panB. The N-terminal his-tagged KPHMT was overexpressed in *E. coli* BL21 (λDE3) and purified to homogeneity using nickel affinity resin. The expressed enzyme has a molecular mass of 35 kDa, which was larger than the 29.4 kDa that was calculated from the deduced amino acid sequence. Most of the recombinant protein was found in a soluble fraction of bacterial lysate, shown in lane 4 of Fig. 5. However, when the first fraction (Fig. 5, lane 5) was eluted, the recombinant protein was produced as an inclusion body and lost its enzyme activity completely (data not shown). This aggregation is probably due to the property of oligomerization at a high concentration. Jones *et al.* (1993) suggested that *E. coli* KPHMT is a hexamer. Thus, we used the purified protein from the second fraction (Fig. 5, lane 6) for the enzyme assay.

BCG KPHMT activity BCG KPHMT activity was determined by measuring the incorporation of $H^{14}CHO$ into ketopantoate, described in Materials and Methods. In this assay, 57.8% of the unreacted $H^{14}CHO$ was removed. The observed specific activity of BCG KPHMT was 0.019 μmol/min/mg. Fig. 6 shows the dependence of the BCG KPHMT activity on time and enzyme concentration. The amount of $H^{14}CHO$ that was incorporated into ketopantoate was proportional to time for at least 30 min, and to enzyme concentration up to 2 μg of KPHMT. These results further confirmed that the 846 bp ORF1 was the *panB* gene encoding KPHMT.

Acknowledgments We thank Dr. Alison G. Smith for the strain Hfr3000 YA139 and the plasmid pAL01.

References

Andersson, S. G. E. and Sharp, P. M. (1996) Codon usage in the *Mycobacterium tuberculosis* complex. *Microbiol.* **142**, 915-925.

Bloom, B. R. and Fine, P. E. M. (1994) The BCG experience: Implications for future vaccines against tuberculosis; in *Tuberculosis*, Bloom, B. R. (ed.), pp. 531-557, ASM Press, Washington, DC.

Bloom, B. R. and Murrat, C. J. L. (1992) Tuberculosis: Commentary on a reemergent killer. *Science* **257**, 1055-1064.

Champreda, V., Khumthong, R., Subsin, B., Angsuthanasombat, C., Panyim, S. and Katzenmeier, G. (2000) The two-component protease NS2B-NS3 of dengue virus type 2: cloning, expression in *Escherichia coli* and purification of the NS2B, NS3(pro) and NS2B-NS3 proteins. *J. Biochem. Mol. Biol.* **33**, 294-299.

Cronan, J. E., Jr., Little, K. J. and Jackowski, S. (1982) Genetic and biochemical analyses of pantothenate biosynthesis in *Escherichia coli* and *Salmonella typhimurium*. *J. Bacteriol.* **149**, 916-922.

Husson, R. N., James, B. E. and Young, R. A. (1990) Gene replacement and expression of foreign DNA in mycobacteria. *J. Bacteriol.* **172**, 519-524.

Jeong, E. and Han, O. (2001) Purification and characterization of *Bacillus subtilis* protoporphyrinogen oxidase and pre-equilibrium behavior during oxidation of protoporphyrinogen IX. *J. Biochem. Mol. Biol.* **34**, 39-42.

Jones, C. E., Brook, J. M., Buck, D., Abell, C. and Smith, A. G. (1993) Cloning and sequencing of the *Escherichia coli panB* gene, which encodes ketopantoate hydroxymethyltransferase, and overexpression of the enzyme. *J. Bacteriol.* **175**, 2125-2130.

Kim, H. J., Lee, Y. M., Hwang, J. S., Won, H. and Kim, B. H. (1999) Defining B cell epitopes of ovalbumin for the C57BL/6 mice immunized with recombinant *Mycobacterium smegmatis*. *J. Biochem. Mol. Biol.* **32**, 461-467.

Kim, J. K., Kim, J. H., Lee, S.-H., Kang, Y., Lim, J.-S., Lee, H. G., Kim, S. J., Bai, G. H., Park, Y. K., Cho, S. H., Kang, S. W., Kim, Y. S., Choe, I. S., Chung, T. W., Park, S. N., Kim, S. H., Ahn, J. S. and Choe, Y.-K. (1997a) Complete sequence of the *upp* gene encoding uracil phosphoribosyltransferase from *Mycobacterium bovis* BCG. *Biochem. Mol. Biol. Int.* **41**, 1117-1124.

Kim, J. K., Kim, J. H., Kim, S. J., Bai, G. H., Cho, S. H., Kang, S. W., Kim, Y. S., Kim, J. W., Lee, Y., Lim, J.-S., Lee, H. G., Choe, I. S., Chung, T. W., Park, S. N., Ahn, J. S. and Choe, Y.-K. (1997b) Cloning and sequencing of the *secY* gene homolog from *Mycobacterium bovis* BCG. *Biochem. Mol. Biol. Int.* **43**, 391-398.

Kim, J.-R., Kang, B.-S., Ko, J.-H., Park, J.-S., Kim, S.-J., Bai, G.-H., Chung, T.-H., Nam, K.-S., Choe, Y.-K., Choe, I.-S., Chung, T.-W., Lee, Y.-C. and Kim, C.-H. (1996) Genomic heterogeneity in clinical strains of *Mycobacterium tuberculosis*, *M. terrae* complex, *M. gordonae*, *M. avium-intracellulae* complex and *M. fortuitum* by pulsed-field gel electrophoresis. *J. Biochem. Mol. Biol.* **29**, 569-573.

Kochi, A. (1991) The global tuberculosis situation and the new control strategy of the world health organization. *Tubercle* **72**, 1-6.

Langermann, S., Palaszynski, S. R., Burlein, J. E., Koenig, S., Hanson, M. S., Briles, D. E. and Stover, C. K. (1994a) Protective humoral response against pneumococcal infection in mice elicited by recombinant bacille Calmette-Guerin vaccines expressing pneumococcal surface protein A. *J. Exp. Med.* **180**, 2277-2286.

Langermann, S., Palaszynski, S., Sadziene, A., Stover, C. K. and Koenig, S. (1994b) Systemic and mucosal immunity induced by BCG vector expressing outer-surface protein A of *Borrelia burgdorferi*. *Nature* **372**, 552-555.

Morales, A. and Nickel, J. C. (1992) Immunotherapy for superficial bladder cancer. *Urol. Clin. North Am.* **19**, 549-556.

Murray, P. J., Aldovini, A. and Young, R. A. (1996) Manipulation and potentiation of antimycobacterial immunity using recombinant bacille Calmette-Guerin strains that secrete cytokines. *Proc. Natl. Acad. Sci. USA* **93**, 934-939.

Sambrook, J. and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.