JOURNAL OF BACTERIOLOGY, May 1997, p. 2900–2906 Vol. 179, No. 9
0021-9193/97/$04.00+0

Three Different Putative Phosphate Transport Receptors Are Encoded by the *Mycobacterium tuberculosis* Genome and Are Present at the Surface of *Mycobacterium bovis* BCG

PHILIPPE LEFÈVRE,[1] MARTINE BRAIBANT,[1] LUCAS DE WIT,[1] MICHAEL KALAI,[1] DETLEF RÖEPER,[2] JOACHIM GRÖTZINGER,[2] JEAN-PIERRE DELVILLE,[3] PRISKA PEIRS,[1] JOSETTE OOMS,[1] KRIS HUYGEN,[1] AND JEAN CONTENT[1]*

Department of Virology, Institut Pasteur, B-1180 Brussels,[1] and Hopital Erasme, Département d'Immunologie, B-1070 Brussels,[3] Belgium, and Institut für Biochemie, Rheinisch-Westfälisch Technische Hochschule Aachen, D-52057 Aachen, Germany[2]

Received 12 August 1996/Accepted 6 February 1997

A gene encoding a protein homologous to the periplasmic ABC phosphate binding receptor PstS from *Escherichia coli* was cloned and sequenced from a λgt11 library of *Mycobacterium tuberculosis* by screening with monoclonal antibody 2A1-2. Its degree of similarity to the *E. coli* PstS is comparable to those of the previously described *M. tuberculosis* phosphate binding protein pab (Ag78, Ag5, or 38-kDa protein) and another *M. tuberculosis* protein which we identified recently. We suggest that the three *M. tuberculosis* proteins share a similar function and could be named PstS-1, PstS-2, and PstS-3, respectively. Molecular modeling of their three-dimensional structures using the structure of the *E. coli* PstS as a template and their inducibility by phosphate starvation support this view. Recombinant PstS-2 and PstS-3 were produced and purified by affinity chromatography. With PstS-1, these proteins were used to demonstrate the specificity of three groups of monoclonal antibodies. Using these antibodies in flow cytometry and immunoblotting analyses, we demonstrate that the three genes are expressed and their protein products are present and accessible at the mycobacterial surface as well as in its culture filtrate. Together with the *M. tuberculosis* genes encoding homologs of the PstA, PstB, and PstC components we cloned before, the present data suggest that at least one, and possibly several, related and functional ABC phosphate transporters exist in mycobacteria. It is hypothesized that the mycobacterial gene duplications presented here may be a subtle adaptation of intracellular pathogens to phosphate starvation in their alternating growth environments.

The diseases caused by mycobacteria are still important sources of morbidity and mortality in the world today. The World Health Organization has established that there are about 10 million new cases of tuberculosis each year and that tuberculosis is responsible for at least 3 million deaths annually (5, 16, 26, 40). The efficacy of the live attenuated strain of *Mycobacterium bovis* BCG, which is presently the only available tuberculosis vaccine, varies considerably from one population to another. This may explain the effort invested by many laboratories in identifying new antigens (51) that may be used to stimulate an efficient immune defense response against the pathogen or provide new and more rapid tools for its identification. The first generation of antigens to be identified was mostly stress proteins (with molecular masses of 10, 65, and 70 kDa) selected with monoclonal antibodies (MAbs) raised in BALB/c mice immunized with killed mycobacteria (16, 51). Immunization of mice with mycobacterial culture filtrates (CF) (29) or with live *Mycobacterium tuberculosis* (11) or *M. bovis* BCG (23) resulted in different humoral responses directed against other distinct antigens. The response obtained was dependent on the mouse strain. C57BL/6 mice infected with live BCG vaccine were found to react preferentially against a 40-kDa protein (Ag88) and a 37-38-kDa doublet protein, whereas BALB/c mice reacted essentially against the 65-kDa heat shock protein and against the 30-32-kDa fibronectin-binding antigen 85 complex (23).

* Corresponding author. Mailing address: Department of Virology, Institut Pasteur, 642 rue Engeland, B-1180 Brussels, Belgium. Phone: 32 2 373 34 16. Fax: 32 2 373 32 79. E-mail: jcontent@ben.vub.ac.be.

In this paper, we report on the cloning and sequencing of an *M. tuberculosis* gene by using the MAb 2A1-2, which reacts with the 37-38-kDa doublet protein from *M. bovis* BCG CF (22). The deduced amino acid sequence encoded by this gene shows a high degree of similarity with PstS (PhoS), the high-affinity phosphate binding subunit of the *E. coli* Pst system (31, 42). Since another mycobacterial PstS-like protein was cloned before (2), we decided to call this latter protein PstS-1 instead of pab, Ag78, Ag5, or 38-kDa protein and the gene we cloned and its product *pstS-2* and PstS-2, respectively.

We have previously shown that the *M. tuberculosis pstS-1* gene belongs to a gene cluster including three other genes called *pstB*, *pstC-1*, and *pstA-2* (see Fig. 1, cluster 1) due to their similarity to the corresponding genes of the *E. coli* Pst system (10). In addition, by genomic walking upstream of the *M. tuberculosis pstA-1* gene (8), the first mycobacterial homolog of *E. coli pstA*, we identified another gene similar to *pstC* and a third *pstS*-like gene, which we designated *pstC-2* and *pstS-3*, respectively, and which constitute cluster 2 (9). Cluster 3 encloses *pstS-2*, which is described here, and the *MbK* gene (33a). All of these genes were found on a continuous DNA stretch of the *M. tuberculosis* genome (see Fig. 1).

Here, we report on the production and purification of the recombinant PstS-2 and PstS-3 proteins (rPstS-2 and rPstS-3, respectively). Using these proteins and recombinant PstS-1 (rPstS-1), we show that a group of MAbs reacting with antigens from the CF of *M. bovis* BCG previously known as the 38-kDa protein (pab) (2), 37-38-kDa protein, and the 40-kDa protein (Ag88) (22) can also react with and distinguish between the products of the *pstS-1*, *-2*, and *-3* genes. Finally, using these

MAbs for flow cytometry, we have demonstrated that these PstS proteins are also presented and are accessible on the mycobacterial surface.

MATERIALS AND METHODS

Screening of λgt11 *M. tuberculosis* library with a MAb. The λgt11 *M. tuberculosis* genomic DNA library was kindly provided by R. A. Young (52). Screening was performed with MAb 2A1-2 (directed against a 37-38-kDa doublet protein [22]) as previously described (7, 24). Crude lysates from selected λgt11 recombinant lysogens were prepared as previously described (7) and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on 8 to 16% gradient gels (NOVEX) and immunoblotting with MAb 2A1-2 as previously described (22).

Large-scale preparation of phage DNA, subcloning, and DNA sequencing. Large-scale preparation of phage DNA was carried out according to standard procedure (34). The R1 clone (see Fig. 1) was digested with *Eco*RI, and a 1.1-kb fragment was subcloned into the *Eco*RI site of pBluescribe-M13+. DNA sequencing was carried out by Sanger's technique (35). Double-stranded plasmid DNA was sequenced on both strands with the T7 sequencing kit (Pharmacia). GC-rich compressed regions were sequenced with *Taq* DNA polymerase (Promega) in the presence of 7-deaza-dGTP.

Production and purification of PstS-2 and PstS-3 recombinant proteins. The gene fragment coding for the mature PstS-2 protein was generated by digestion of the 1.1-kb *Eco*RI fragment of R1 pBluescribe-M13+ with *Sph*I and *Eco*RI and subcloned into pRSET-B (Invitrogen) in frame with a His_6 encoding headpiece. For PstS-3, a fragment of 1,079 bp was amplified for 25 PCR cycles with the Pfu polymerase (Stratagene) with the oligonucleotides 5′-CGCGGATCCT-CTGTG GTAACGACGACAATGTGACC (forward primer) and 5′-CGCGGA-TCCCG TCAACCTCAGATCAGG (reverse primer). The amplification product was then digested with *Bgl*II and inserted in pRSET-C (Invitrogen) in frame with a His_6 encoding headpiece. Competent *E. coli* JM109 (DE3) (Promega) was transformed with the two constructions, and transformants were selected on Luria-Bertani plates supplemented with 50 μg of ampicillin/ml.

For protein expression, cells were grown in SOB medium with 50 μg of ampicillin/ml to an optical density at 600 nm (OD_{600}) of 0.45 and then induced with 1 mM isopropyl-β-D-thioglycolate for 3 h at 37°C. The recombinant proteins, recovered in major quantities in the insoluble fraction (inclusion bodies), were purified by metal affinity chromatography on a Ni-nitrilotriacetic acid resin column (Qiagen) under denaturing conditions. Purity of the purified recombinant proteins was examined by Coomassie staining of an SDS–15% PAGE electrophoresis gel. For immunoblot analysis, nitrocellulose filters were incubated with the specified MAbs and revealed with the Protoblot Western Blot AP system (Promega) according to the manufacturer's instructions.

Phosphate starvation. Prior to culturing the cells, the Sauton medium (36) was supplemented with K_2HPO_4 to a final concentration of 3 g/liter (instead of 0.5 g/liter in the original medium). *M. bovis* BCG was inoculated into this modified Sauton medium and grown at 37°C under agitation. When the culture had reached an OD_{460} of 0.3, two samples were harvested by centrifugation at 6,000 × *g*, washed, and resuspended in the same volume of either modified Sauton medium or Sauton medium prepared without phosphate. The culture was continued for 24 h, and the bacterial lysates were prepared as previously described (3).

Computer analysis. Computer-aided analyses of the nucleic acid and deduced amino acid sequences were performed with the DNA Strider program (32) and the Genetics Computer Group program (14) of the BEN (Belgian EMBnet Node) network facility. Homology search in the protein sequence data banks was greatly facilitated by use of the EMBL Blitz server, which utilizes the Smith and Waterman algorithm (39) on a MPsrch program, and the NCBI Blast programs (1, 20). Probable sequencing errors and open reading frames (ORFs) were detected with GenMark 2.1 (6).

Molecular modeling. The three-dimensional models of PstS-1, -2, and -3 were built with the X-ray structure of the *E. coli* PstS (protein data bank accession code, 1bpb) as a template. The side chains of residues in the *E. coli* PstS were exchanged for those of the *M. tuberculosis* proteins according to the sequence alignments (see Fig. 6) by using a database search approach. Preliminary inspection of the protein cores showed fairly good packing of the molecules without unfavorable contacts of side chain atoms, indicating a good accommodation of the *M. tuberculosis* sequences into the *E. coli* PstS structure. Nevertheless, the structures were energy minimized over 500 steps by using a steepest-descent algorithm to reach an energetically low conformation. Molecular modeling and graphical representation were carried out by using the programs WHATIF (48) and GRASP (33), respectively, on an Indigo 2 SGI (Silicon Graphics Inc., Mountain View, Calif.).

Flow cytometry analysis. Flow cytometry analysis was performed on a Facscan (Becton Dickinson) driven by Cellquest software. Green fluorescence was studied with a 530/30-nm bandpass filter. Photomultiplier tube pulses were amplified logarithmically. Ten-thousand events were measured and stored in list mode data files. Before each sample run, the flow cytometer performances were monitored by using fluorescent microspheres (Calibrite Fluorescent kit, Becton Dickinson). The fluorescence was calibrated with beads of known fluorescence activity ranging from 6.3×10^4 to 1.41×10^6 equivalent soluble molecules of fluorochrome (Flow Cytometry Standards). Bacteria were gated with their morphological properties (forward scatter-side scatter) set on logarithmic mode. The mean fluorescence intensity of the related populations of bacteria was calculated from histograms and expressed in arbitrary units corresponding to an intensity channel number ranging from 0 to 1023. For the staining procedure with MAbs, *M. bovis* BCG and *E. coli* (JM109, Promega) cultures were grown in Sauton medium supplemented with 0.025% Triton X-100 and Luria-Bertani medium, respectively, to an OD_{600} of 0.7. Cells were harvested by centrifugation at 3,000 × *g* at 4°C for 5 min and resuspended in 50 μl of a buffer containing 10% fetal calf serum and 1% bovine serum albumin in normal saline solution (buffer L). Bacteria were incubated for 1 h at 4°C, with permanent agitation, in the presence of one of the following MAbs: HBT12, HYT28, 2A1-2, 2C1-5, or 2F8-3. Anti-human CD3 MAb and anti-human CD4 MAb (Becton Dickinson) were used as negative controls. Cells were washed three times with buffer L and stained (by incubation for 1 h at 4°C in a final volume of 100 μl) with the secondary fluorescein isothiocyanate-conjugated anti-mouse antibody (Dako, Glostrup, Denmark). After being washed, the cell pellets were resuspended in normal saline for cytometry analysis.

Nucleotide sequence accession number. The DNA sequence of the *pstS-2* gene has been deposited in the European Molecular Biology Laboratory database under accession number Z48056.

RESULTS

Cloning and sequencing of *M. tuberculosis pstS-2* gene. Screening a λgt11 *M. tuberculosis* expression library with MAb 2A1-2, we isolated the recombinant phage R1 (Fig. 1). Analysis of crude lysates of *E. coli* lysogens of this recombinant phage revealed a fusion protein of ~150 kDa that was recognized by MAb 2A1-2 in a Western blot assay, suggesting that the clone contained a relatively large part of the 37-38-kDa protein (data not shown).

To complete the gene sequence, we screened the same λgt11 *M. tuberculosis* library with a fragment from the right extremity of R1 (Fig. 1). We isolated clone B1, which displayed an overlap of about 900 bp with the R1 fragment sequence (Fig. 1) (10). The nucleotide sequence (EMBL Z48056) corresponding to the overlap between R1 and B1 shows that among the several potential in-frame initiation codons, only one GTG was preceded, 12 nucleotides upstream, by a purine-rich, potential ribosome binding site. It was therefore chosen as the most likely start codon of the R1-B1 ORF, extending to a final TGA codon at position 1111. The overall base composition was 61% guanines and cytosines with a predominance of these bases at the third position of each codon. The deduced amino acid sequence of the complete R1-B1 ORF corresponds to a protein of 370 residues which is highly similar to the *M. tuberculosis* phosphate binding component PstS-1 (77% similarity, 29.5% identity) (2). We therefore designated this gene *pstS-2* and the protein it encodes, PstS-2. Shortly after, we identified by genomic walking upstream of *pstA-1* (8) another gene, similar to *pstS-1* and *pstS-2*, which we named *pstS-3* (28 and 65.5% identity; 73 and 94% similarity with PstS-1 and PstS-2, respectively) (Fig. 1) (9). Examination of the deduced amino acid sequences of PstS-1 (2), PstS-2, and PstS-3 revealed they contain a typical 20- to 25-amino-acid hydrophobic signal sequence and the consensus amino acid sequence for a predicted prokaryotic membrane lipoprotein lipid attachment site: {DERK}$_{(6)}$–[LIVMFWSTAG]$_{(2)}$–[LIVMFYSTAG-CQ]–[AGS]–C (4) (see Fig. 6). Interestingly, PstS-1 was recently shown to be a lipoprotein by metabolic labeling with fatty acid precursors and Triton X-114 phase separation (42).

Expression of PstS-2 and PstS-3 in *E. coli* and their recognition by specific MAbs. In order to show that these genes are coding for the predicted PstS proteins, and to find out whether they are recognized by available MAbs directed against *M. bovis* BCG CF antigens, we expressed PstS-2 and PstS-3 in recombinant form in *E. coli*. To prevent PstS-2 and PstS-3 from being exported to the periplasm and anchored to the cytoplas-

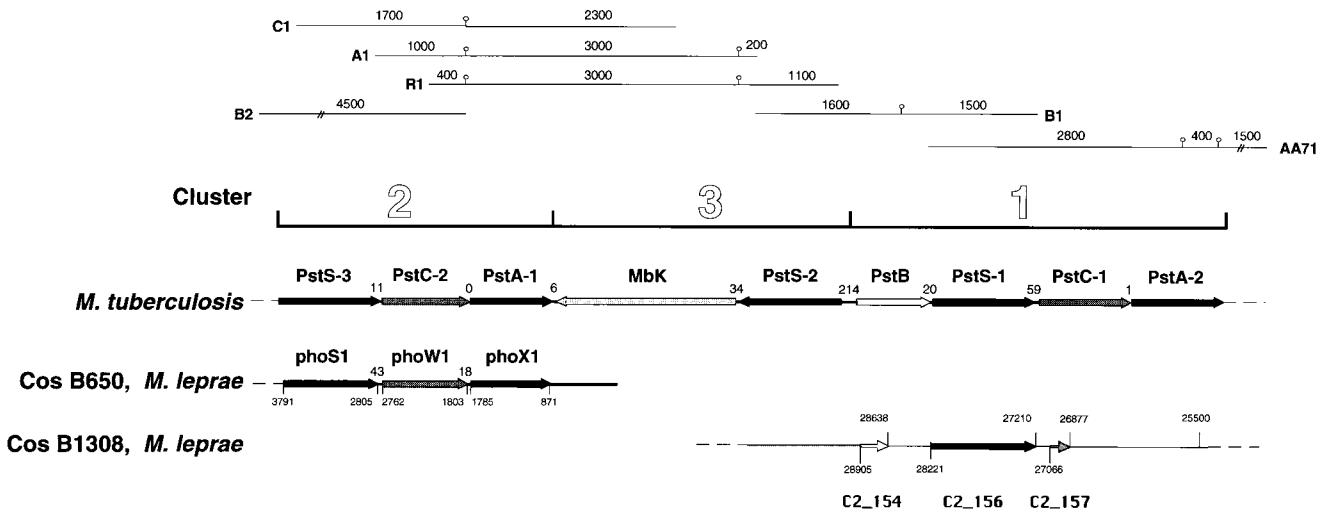

FIG. 1. Map of overlapping *M. tuberculosis*-λgt11 clones C1, A1 (8, 9), R1 (this study), B2 (9), B1 (10), and AA71 (2). The arrows indicate the orientation of ORFs. The drawing shows the locations of the new *pstS* genes relative to the components of other phosphate permeases which are organized in three clusters as indicated (8–10). The *M. leprae* ORFs found in cosmids B650 and B1308 and which are homologous to the *M. tuberculosis* proteins are indicated. Hairpins represent *Eco*RI restriction sites.

mic membrane by a lipoyl-amino acid moiety we chose to produce these recombinant proteins without their putative signal sequence. Recombinant PstS-2 and PstS-3 were expressed as His-tagged fusion proteins and purified. Coomassie blue staining of SDS-PAGE gels of the purified proteins rPstS-2 and rPstS-3 has shown that they were both expressed well and had the expected molecular mass. An example of purification of rPstS-2 is shown in Fig. 2A. Previous studies have shown that MAbs HBT12 and HYT28, 2A1-2, and 2F8-3 and 2C1-5 react with pab (38 kDa), the 37-38-kDa protein, and Ag88 (40 kDa), respectively, from the CF of *M. bovis* BCG (2, 22).

We compared the recognition of the three recombinant proteins and CF antigens by Western blot analysis using these three different MAbs. As shown in Fig. 2B, rPstS-2 is recognized by MAb 2A1-2, rPstS-3 is recognized by MAb 2F8-3, and rPstS-1 (38) is recognized by HBT12. These results demonstrate that the *pstS-1*, *-2*, and *-3* genes exist and are expressed as CF antigens in *M. bovis* BCG. The MAbs are specific since no cross-reactivity was detected between them.

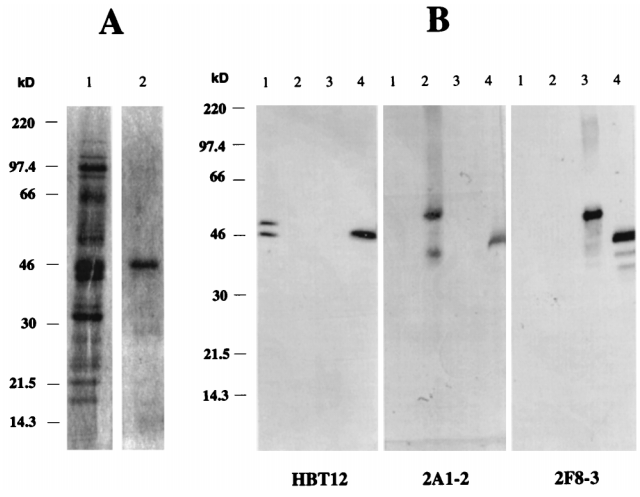

FIG. 2. (A) Coomassie blue-stained SDS-PAGE gel showing the purified recombinant PstS-2 protein. Lane 1 represents the flow-through and lane 2 represents the elution at pH 6.5. (B) Immunoblotting analysis with MAbs HBT12, 2A1-2, and 2F8-3 of recombinant proteins PstS-1 (lane 1), PstS-2 (lane 2), and PstS-3 (lane 3) and CF of *M. bovis* BCG (lane 4).

Are PstS-1, PstS-2, and PstS-3 secreted and/or anchored proteins? As described above the three PstS proteins are predicted to be lipoproteins which may bind the bacterial membrane. However their presence in CF suggests that they are soluble and probably secreted. To test whether they are attached to and exposed on the *M. bovis* BCG membrane we analyzed 10-day cultures of *M. bovis* BCG populations (OD_{600} of 0.7) by using flow cytometry. We used the five specific MAbs described above. As shown in Fig. 3 MAbs HYT28, 2A1-2, and 2F8-3 each revealed a difference of about 1 $\log_{10}$ in fluorescence intensity compared to negative controls. Similar results were obtained with HBT12 (for PstS-1) and 2C1-5 (for PstS-3) (data not shown). None of the five MAbs showed any emission of fluorescence when *E. coli* bacterial populations at the same OD were used (data not shown). Labeling of the mycobacteria with anti-human CD3 MAbs or anti-human CD4 MAbs (data not shown) was also negative. Altogether, the above results demonstrate that the three PstS subunits are located on the mycobacterial surface, possibly by cell wall or membrane anchoring.

PstS proteins and phosphate metabolism. The effect of phosphate starvation on PstS-1, PstS-2, and PstS-3 expression within whole *M. bovis* BCG cells cultivated for 24 h was analyzed by immunoblot assay (Fig. 4). The 32-kDa protein (Ag85) (50) was used as an internal control in this experiment. The accumulation of the three PstS proteins is enhanced under phosphate starvation while the production of Ag85 is not influenced.

The PstS proteins constitute a family of phosphate receptors for bacterial ABC-type transporters. A homology search for PstS-2 and PstS-3 with the Blast program (20) detected a dozen additional PstS-type proteins in eight distinct bacterial organisms. A dendrogram of these proteins shows that three clusters may be distinguished: a PstS-1 group, including three species of mycobacteria (*M. tuberculosis*, *M. leprae*, and *M. intracellulare*), a second group containing PstS-2 and PstS-3 of

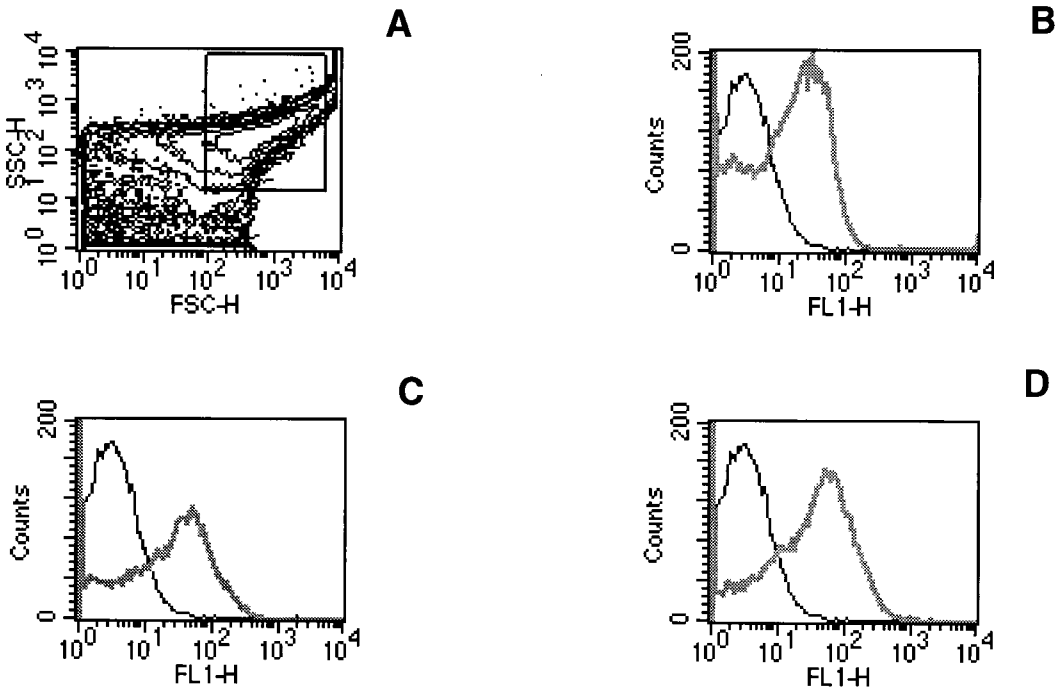

FIG. 3. Flow cytometric analysis of bacterial populations. (A) Dot plots of the gated mycobacterial population used for analysis. The three MAbs used were HYT28 (B) for PstS-1, 2A1-2 (C) for PstS-2, and 2F8-3 (D) for PstS-3. Fine-line tracings represent bacterial populations without MAbs (negative control), and the heavy lines represent the same specimens stained with the indicated MAbs.

M. tuberculosis and PstS-4 of *M. leprae* (Fig. 5). The third group is represented by the external phosphate receptor of gram-negative bacteria and cyanobacteria, this group being distant from the gram-positive bacteria represented by the genus *Bacillus*. It is interesting that according to this diagram the putative evolutionary distance between the PstS-1 cluster and the PstS-2–PstS-3–PstS-4 cluster is nearly as important as their respective distances from the gram-negative cluster, suggesting a very ancient duplication (divergence) within the mycobacteria.

Recently, the X-ray diffraction structure of the liganded form of PstS of *E. coli* has been determined (30). It has indicated that Thr-10, Phe-11, Ser-38, Asp-56, Arg-135, Ser-139, Gly-140, and Thr-141 are directly involved in hydrogen bonding with the phosphate. The alignment between this PstS from *E. coli* and its three *M. tuberculosis* homologs (Fig. 6) showed that amino acids identical (six locations) or equivalent (two locations) to these eight key residues (see details below) are also present in the aligned mycobacterial amino acid sequences. A major difference is apparent in the N-terminal region of the *M. tuberculosis* proteins, which is about 20 residues longer than that of the *E. coli* periplasmic PstS. If the mycobacterial PstS proteins are lipoproteins attached to the cell wall or membrane, this stretch may provide them the structural freedom needed for their function.

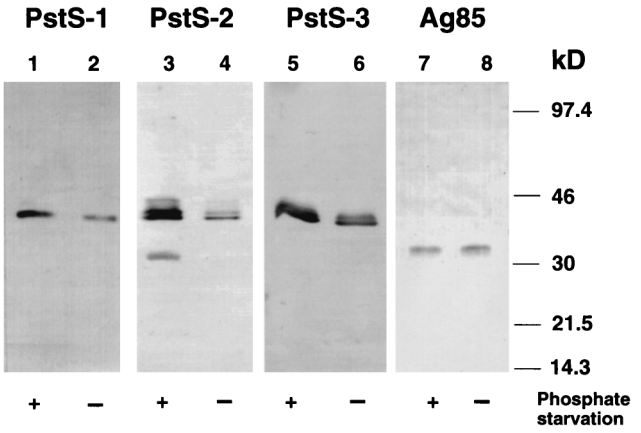

FIG. 4. Western blot analysis of the total proteins from *M. bovis* BCG submitted (+) or not (−) to phosphate starvation. The MAbs used were HBT12 for PstS-1, 2A1-2 for PstS-2, 2C1-5 for PstS-3, and 32TD15 for Ag85.

Molecular modeling of the PstS-1, PstS-2, and PstS-3 structures. Using the coordinates of the structure of the *E. coli* PstS we constructed three-dimensional (3-D) models of its three *M. tuberculosis* homologs (see Materials and Methods) (Fig. 7A). Comparison of our 3-D models with the structure of the *E. coli* PstS has revealed they have high tertiary structure similarity indicating good accommodation of the three *M. tuberculosis* sequences into the *E. coli* PstS structure. All four proteins consist of two distinct globular domains bisected by a groove wherein the ligand binds. The movement of these domains around a predicted hinge was suggested as the mechanism by which the orthophosphate ligand is buried within the protein. Structural differences between the models of *M. tuberculosis* proteins and the X-ray structure of *E. coli* PstS are found in loops (Fig. 7A) where according to the alignment deletions and insertions of amino acid residues had to be accommodated. One of these loops is located at the site of the predicted hinge region. This may affect the constants of equilibrium between the free and bound forms of the receptor. Close inspection of the ligand binding sites of the *M. tuberculosis* proteins has revealed that the amino acid residues predicted to interact directly with the ligand

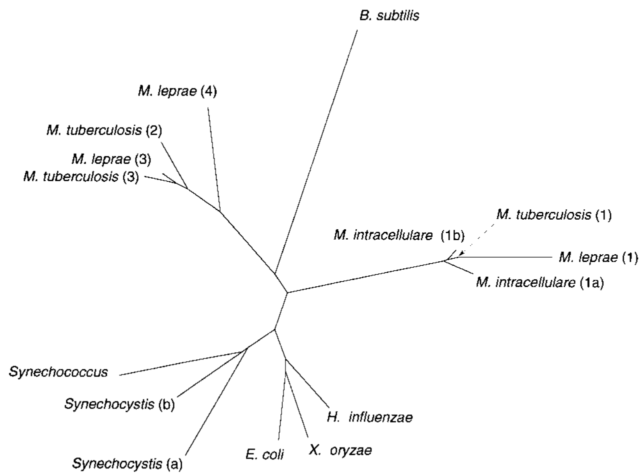

FIG. 5. Dendrogram representing the distances between aligned sequences of proteins related to *E. coli* PstS. The PstS-like protein sequences analyzed are as follows: *E. coli* (43); *X. oryzae* (*Xanthomonas oryzae* [21]); *H. influenzae* (*Haemophilus influenzae* [19]); *Synechocystis* sp. (a) and (b) (*Synechocystis* sp. PCC6803 [25], accession number D90909); *Synechococcus* sp. (*Synechococcus* sp. WH7803 [37]); *B. subtilis* (*Bacillus subtilis*, ORF 108, accession number D58414); *M. tuberculosis* (1), (2), and (3) (*M. tuberculosis* PstS-1 [2], PstS-2 [this study], and PstS-3 [9], respectively); *M. leprae* (1), (3), and (4) (*M. leprae* PstS-1 [C2_156, cosmid B1308], PstS-3 [phoS1, cosmid 650], and PstS-4 [phoS, cosmid B2266]); and *M. intracellulare* (1a) and (1b) (*M. intracellulare* PstS-1a and PstS-1b, accession number X95538). The dendrogram was constructed with the help of the Genetics Computer Group (GCG) software (14). The sequences were aligned with the pileup program. A Kimura distance matrix (27) was computed with the Distances GCG program. The tree was constructed with the GrowTree GCG program by the neighbor-joining method (44) and plotted with the PHYLIP (18) program Drawtree.

are as suggested by our alignment (Fig. 6). Phe-11 of the *E. coli* protein is replaced by Leu-37, Ala-34, and Ala-36 in the mature PstS-1, -2, and -3, respectively, whereas Thr-141 is replaced by Asp-168 in PstS-1 alone. Phe-11 of PstS of *E. coli* was not conserved, probably since its involvement in the hydrogen bonding with the phosphate is ensured by the amino-terminal end of the residue and not by the aromatic side chain (30).

DISCUSSION

In this paper we describe the cloning of the *pstS-2* gene. Together with the *pstS-1* and *pstS-3* genes this is the third *M. tuberculosis* gene that encodes a protein closely related to *E. coli* PstS, the high-affinity soluble ligand binding protein component of an ABC periplasmic phosphate permease (43, 47, 49). Protein sequence alignments and molecular modeling of the mycobacterial proteins, with the crystal structure of the *E. coli* PstS (30) as a template, have demonstrated that the predicted 3-D structures of all of the three proteins, compared with the *E. coli* PstS structure, are very much alike and that the amino acid residues included in the phosphate binding site are conserved (Fig. 7A). Since PstS-1 is able to bind phosphate with an affinity similar to that of PstS of *E. coli* (12) and since the structural similarity between the different mycobacterial PstS proteins is high, it is very likely that all three are phosphate binding proteins. Moreover, the observed induction of these three PstS proteins under phosphate starvation also suggests that they are involved in phosphate transport (3, 17, 49).

By using five MAbs reacting with specific *M. bovis* BCG CF antigens, we have shown by immunoblot assay that the three pstS genes are expressed. Flow cytometry analysis using the same MAbs has demonstrated that the three PstS proteins are accessible on the surface of intact living mycobacteria where they are probably anchored. These data are in agreement with our previous observation that MAb HBT12 (directed against PstS-1) neutralizes phosphate uptake by *M. bovis* BCG (10).

We have recently cloned and sequenced a few other *M. tuberculosis* genes corresponding to two PstA, two PstC, and one PstB components of putative phosphate permeases (8–10). The three *pstS* genes are positioned in different orientations and are probably clustered in three potential adjacent operons controlled by different promoters (Fig. 1).

Altogether, the present data support the presumption that at least one, and possibly several, related and functional phosphate permeases exist in mycobacteria. It was previously shown that a single ABC prokaryotic permease uses different ligand binding subunits specific for distinct substrates (28). However, the existence of several ligand binding subunits for a single ligand has not been described to date among known ABC transporters (45) and may represent an interesting and subtle

```
         1                                                                                                   100
PstS-1   MKIRLHTLLA VLTAAPLLLA AAGCG.SKPP SGSPETGAGA GTVATTPASS PVTLAETGST LLYPLFNLWG PAFHERYPNV TITAQGTGSG AGIAQAAAGT
PstS-2   MKFARSGAAV SLLAAGTLVL TA.CGGGTNS SSSGA..GGT SGSV.HCGGK K.ELHSSGST AQENAMEQFV YAYVRSCPGY TLDYNANGSG AGVTQFLNNE
PstS-3   MKLNRFGAAV GVLAAGALVL SA.CGNDDNV TGGGATTGQA SAKV.DCGGK K.TLKASGST AQANAMTRFV NVFEQACPGQ TLNYTANGSG AGISEFNGNQ
  PstS   MKV..MRTTV ATVVAATLSM SAFSVFAEAS .......... .......... ...LTGAGAT FPAPVYAKWA DTYQKETGN. KVNYQGIGSS GGVKQIIANT

         101                                                                                                 200
PstS-1   VNIGASDAYL SEGDMAAHKG .......LMN IALAISAQQV NYNLPGV.SE HLKLNGKVLA AMYQGTIKTW DDPQIAALNP GVNLPGTAVV PLHRSDGSGD
PstS-2   TDFAGSDVPL NPSTGQPDRS AERCGSPAWD LPTVFGPIAI TYNIKGVST. .LNLDGPTTA KIFNGTITVW NDPQIQALNS GTDLPPTPIS VIFRSDKSGT
PstS-3   TDFGGSDVPL ..SKDEPS.G AARCGSPAWN LPVVFGPIAV TYNLNSVSS. .LNLDGPTLA KIFNGSITQW NNPAIQALNR DFTLPGERIH VVFRSDESGT
  PstS   VDFGASDAPL SDEKL.AQEG .......LFQ FPTVIGGVVL AVNIPGLKSG ELVLDGKTLG DIYLGKIKKW DDEAIAKLNP GLKLPSQNIA VVRRADGSGT

         201                                                                                                 300
PstS-1   TFLFTQYLSK QDPEGWGKSP GFGTTVDFPA VPGALGENGN GGMVTGCAET PGCVAYIGIS FLDQASQRGL GEAQLG.NSS GNFLLPDAQS IQ...AAAAG
PstS-2   SDNFQKYLDG ASNGAWGKGA SETFNGG... ..VGVGASGN NGTSALLQTT DGSITYNEWS F...AVGKQL NMAQIITSAG PDPVAITTES VGKTIAGAKI
PstS-3   TDNFQRYLQA ASNGAWGKGA GKSFQGG... ..VGEGARGN DGTSAAAKNT PGSITYNEWS F...AQAHDL TMANIVTSAG GDPVAITIDS VGQTIAGATI
  PstS   SFVFTSYLAK VN.EEWKNNV GTGSTVKWPI ...GLGGKGN DGIAAFVQRL PGAIGYVEYA Y...AKQNNL AYTKLI.SAD GKPVSPTEEN FA...NAAKG

         301                                                                                         396
PstS-1   FA.SKTPANQ AISMIDGPAP DGYPIINYEY AIVNNRQKDA ATAQTLQAFL HWAITDGNKA SFLDQVHFQP LPPAVVKLSD ALIATISS.. ......
PstS-2   MGQGNDLVLD TSSFYRPTQP GSYPIVLATY EIVCSKYPDA TTGTAVRAFM QAAIGPGQEG ..LDQYGSIP LPKSFQAQLA AAVNAIS... ......
PstS-3   SGVGNDLVLD TDSFYRPKRP GSYPIVLATY EIVCSKYPDS QVGTAVKAFL QSTIGAGQSG ..LGDNGYIP IPDEFKSRLS TAVNAIA... ......
  PstS   ADWSKTKAQD ...LTNQKGE DAWPITSTTF ILIHKDQKKP EQGTEVLKFF DWAYKTGAKQ A..NDLDYAS LPDSVVEQVR AAWKTNIKDS SGKPLY
```

FIG. 6. Alignment of the amino acid sequences of *E. coli* PstS and *M. tuberculosis* PstS-1, PstS-2, and PstS-3. The eight amino acid residues homologous to those involved in phosphate binding of *E. coli* PstS are in black boxes. The potential mycobacterial lipoprotein consensus sequence cleavage site is indicated by an arrow.

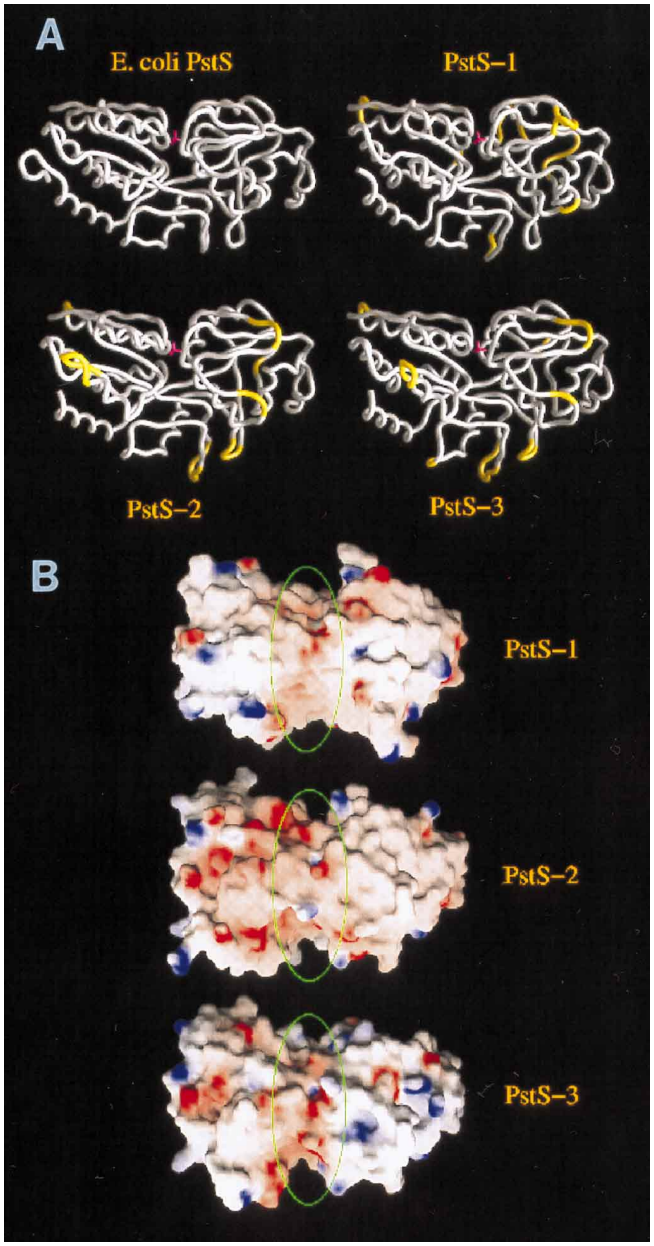

FIG. 7. (A) Schematic backbone representation of the X-ray structure of the *E. coli* PstS and the molecular models of *M. tuberculosis* PstS-1, PstS-2, and PstS-3. The bound phosphate is colored in red. Regions in the *M. tuberculosis* PstS-1, PstS-2, and PstS-3 models which differ in structure compared with the *E. coli* PstS are depicted in yellow. These differences occur only in loop regions where deletions or insertions of amino acid residues had to be accommodated. (B) Graphical representation of the electrostatic potential map of the predicted PstA-PstC binding interface of *M. tuberculosis* PstS-1, PstS-2, and PstS-3 (top view). Colors represent the negative (red) and positive (blue) electrostatic potential at the surface of the molecule. The surface of the phosphate binding cleft is highlighted.

biochemical adaptation of a microorganism capable of growth and survival under very different conditions.

Although the proteins share a high degree of similarity, their predicted isoelectric points are different (4.68, 5.38, and 4.83 for PstS-1, PstS-2, and PstS-3, respectively). Inspection of the predicted PstA-PstC binding interface of the 3-D models of PstS-1, -2, and -3 suggests that the electrostatic potential at the surface of the molecules is different (Fig. 7B). Thus, it is conceivable that the three different mycobacterial PstS proteins interact differently with the different PstA and PstC subunits. In addition, their expression may vary in different environments (extracellular or intracellular or within the phagolysosomes or the cytoplasm) depending on the pH or phosphate concentrations (see above) (13, 41).

It is likely that the three *pstS* genes have originated as the result of several gene duplications. These duplications are probably more recent for *pstS-2* and *pstS-3* since their amino acid sequences are highly similar and relatively distant from PstS-1. We have found that *pstS-1* (Fig. 1) and *pstS-2* (not shown) homologs exist in the genome of *M. leprae* with a similar organization to that described here for *M. tuberculosis*, whereas a *pstS-3 M. leprae* homolog is located at a distant site, within a permease gene cluster similar to the one found in *M. tuberculosis* (9). In the *M. leprae* genome, an additional fourth *pstS*-like gene was found, which we designated *pstS-4*. It is not clear at the moment whether another such gene also exists in *M. tuberculosis*. A somewhat similar diversity also exists in *M. intracellulare* (46) (Fig. 5). This may indicate that in all mycobacteria this diversification of phosphate permeases has occurred.

Detailed knowledge of these phosphate transport systems could open new insights into their physiology, regulation in various conditions, importance for the survival of mycobacteria, and pathogenesis. At least one of the MAbs used here, 2A1-2, was able to distinguish *M. tuberculosis* from other mycobacterial species (15). Therefore, the accessibility of the PstS antigens to these MAbs on the surface of intact mycobacteria may open new possibilities for rapid and specific diagnosis. Finally, molecular modeling may prove useful for the development of specific inhibitors of these transport systems such as mutated binding proteins, peptides, site-directed MAbs, or even synthetic compounds.

ACKNOWLEDGMENTS

We thank R. A. Young for the *M. tuberculosis* λgt11 library, P. Vandenbussche for advice, S. Cole and D. R. Smith for unpublished *M. leprae* sequences, Å. Andersen for the gift of HBT12 and HYT28, and M. Singh for rPstS-1. Competent biocomputing support was provided by the members of the Belgian EMBnet Node.

This work was supported by grant 3.4543.95 from FRSM (Fund for Medical Scientific Research, Belgium), Les Amis de l'Institut Pasteur de Bruxelles, ASBL, and IRSIA (Belgium).

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
2. **Andersen, A. B., and E. B. Hansen.** 1989. Structure and mapping of antigenic domains of protein antigen b, a 38,000-molecular-weight protein of *Mycobacterium tuberculosis*. Infect. Immun. **57:**2481–2488.
3. **Andersen, A. B., L. Ljungqvist, and M. Olsen.** 1990. Evidence that protein antigen b of *Mycobacterium tuberculosis* is involved in phosphate metabolism. J. Gen. Microbiol. **136:**477–480.
4. **Bairoch, A.** 1991. PROSITE: a dictionary of sites and patterns in proteins. Nucleic Acids Res. **25:**2241–2245.
5. **Bloom, B. R. C., and C. J. L. Murray.** 1992. Tuberculosis: commentary on a reemergent killer. Science **257:**1055–1064.
6. **Borodovsky, M., and J. McIninch.** 1993. GENMARK: parallel gene recognition for both DNA strands. Comput. Chem. **17:**123–133.
7. **Borremans, M., L. De Wit, G. Volckaert, J. Ooms, J. De Bruyn, K. Huygen, J. P. Van Vooren, M. Stelandre, R. Verhofstadt, and J. Content.** 1989. Cloning, sequence determination, and expression of a 32-kilodalton-protein gene of *Mycobacterium tuberculosis*. Infect. Immun. **57:**3123–3130.
8. **Braibant, M., L. De Wit, P. Peirs, M. Kalai, J. Ooms, K. Huygen, and J. Content.** 1994. Structure of the *Mycobacterium tuberculosis* antigen 88: a protein related to the *Escherichia coli* PstA periplasmic phosphate permease subunit. Infect. Immun. **62:**849–854.
9. **Braibant, M., P. Lefèvre, L. De Wit, J. Ooms, P. Peirs, K. Huygen, R. Wattiez, and J. Content.** 1996. Identification of a second *Mycobacterium tuberculosis* gene cluster encoding proteins of an ABC phosphate transporter. FEBS Lett. **394:**206–212.

10. **Braibant, M., P. Lefèvre, L. De Wit, P. Peirs, J. Ooms, K. Huygen, A. B. Andersen, and J. Content.** 1996. A *Mycobacterium tuberculosis* gene cluster encoding proteins of a phosphate transporter homologous to the *Escherichia coli* Pst system. Gene **176:**171–176.
11. **Brett, S., and J. Ivanyi.** 1990. Genetic influences on the immune repertoire following tuberculosis infection in mice. Immunology **71:**113–119.
12. **Chang, Z., A. Choudhary, R. Lathigra, and F. A. Quiocho.** 1994. The immunodominant 38-kDa lipoprotein of *M. tuberculosis* is a phosphate binding protein. J. Biol. Chem. **269:**1956–1958.
13. **Chicurel, M., E. Garcia, and F. Goodsaid.** 1988. Modulation of macrophage lysosomal pH by *Mycobacterium tuberculosis*-derived proteins. Infect. Immun. **56:**479–483.
14. **Devereux, J., J. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
15. **Drowart, A., C. L. Cambiaso, K. Huygen, E. Serruys, J.-C. Yernault, and J.-P. Van Vooren.** 1993. Detection of mycobacterial antigens present in short-term culture media using particle counting immunoassay. Am. Rev. Respir. Dis. **147:**1401–1406.
16. **Engers, H. D., et al.** 1986. Results of a World Health Organization-sponsored workshop to characterize antigens recognized by mycobacterium-specific monoclonal antibodies. Infect. Immun. **51:**718–720.
17. **Espitia, C., M. Elinos, R. Hernandez-Pando, and R. Mancilla.** 1992. Phosphate starvation enhances expression of the immunodominant 38-kilodalton protein antigen of *Mycobacterium tuberculosis*: demonstration by immunogold electron microscopy. Infect. Immun. **60:**2998–3001.
18. **Felsenstein, J.** 1989. PHYLIP—phylogeny inference package (version 3.2). Cladistics **5:**164–166.
19. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J.-F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenney, G. Sutton, W. FitzHugh, C. A. Fields, J. D. Gocayne, J. D. Scott, R. Shirley, L.-I. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrmann, N. S. M. Geoghagen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Fraser, H. O. Smith, and J. C. Venter.** 1995. Whole-genome random sequencing and assembly of Haemophilus. Science **269:**496–512.
20. **Gish, W., and D. J. States.** 1993. Identification of protein coding regions by database similarity search. Nat. Genet. **3:**266–272.
21. **Hopkins, C. M., F. F. White, L. A. Heaton, J. A. Guikema, and J. E. Leach.** 1995. A homolog of an *Escherichia coli* phosphate-binding protein gene from *Xanthomonas oryzae* pv. oryzae. DNA Sequence **5:**299–305.
22. **Huygen, K., A. Drowart, M. Harboe, R. ten-Berg, J. Cogniaux, and J. P. Van-Vooren.** 1993. Influence of genes from the major histocompatibility complex on the antibody repertoire against culture filtrate antigens in mice infected with live *Mycobacterium bovis* BCG. Infect. Immun. **61:**2687–2693.
23. **Huygen, K., L. Ljungqvist, R. ten-Berg, and J. P. Van-Vooren.** 1990. Repertoires of antibodies to culture filtrate antigens in different mouse strains infected with *Mycobacterium bovis* BCG. Infect. Immun. **58:**2192–2197.
24. **Huynh, T. V., R. A. Young, and R. W. Davis (ed.).** 1985. Constructing and screening cDNA libraries in λgt10 and λgt11, vol. I. Oxford University Press, Oxford, United Kingdom.
25. **Kaneko, T., A. Tanaka, S. Sato, H. Kotani, T. Sazuka, N. Miyajima, M. Sugiura, and S. Tabata.** 1995. Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome. DNA Res. **2:**153–166.
26. **Kaufmann, S. H. E., and J. D. A. van Embden.** 1993. Tuberculosis: a neglected disease strikes back. Trends Microbiol. **1:**2–5.
27. **Kimura, M.** 1983. Neutral theory of molecular evolution. Cambridge University Press, Cambridge, United Kingdom.
28. **Kustu, S. G., and G. F.-L. Ames.** 1973. The *hisP* protein, a known histidine transport component in *Salmonella typhimurium*, is also an arginine transport component. J. Bacteriol. **116:**107–113.
29. **Ljungqvist, L. A., A. Worsaae, and I. Heron.** 1988. Antibody responses against *Mycobacterium tuberculosis* in 11 strains of inbred mice: novel monoclonal antibody specificities generated by fusions, using spleens from BALB.B10 and CBA/J mice. Infect. Immun. **56:**1994–1998.
30. **Luecke, K., and F. A. Quiocho.** 1990. High specificity of a phosphate transport protein determined by hydrogen bonds. Nature **347:**402–406.
31. **Magota, K., N. Otsuji, T. Miki, T. Horiuchi, S. Tsunasawa, J. Kondo, F. Sakiyama, M. Amemura, T. Morita, H. Shinagawa, and A. Nakata.** 1984. Nucleotide sequence of the *phoS* gene, the structural gene for the phosphate-binding protein of *Escherichia coli*. J. Bacteriol. **157:**909–917.
32. **Marck, C.** 1988. 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. Nucleic Acids Res. **16:**1829–1836.
33. **Nicholls, A., K. A. Sharp, and B. Honig.** 1991. Protein folding and association: insights from the the interfacial and thermodynamic properties of hydrocarbons. Proteins **11:**281–296.

33a. **Peirs, P., L. DeWitt, M. Braibant, K. Huygen, and J. Content.** A serine-threonine protein kinase from Mycobacterium tuberculosis. Eur. J. Biochem., in press.

34. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
36. **Sauton, B.** 1912. Sur la nutrition minérale du bacille tuberculeux. C. R. Acad. Sciences **155:**860–861.
37. **Scanlan, D. J., N. H. Mann, and N. G. Carr.** 1993. The response of the picoplanktonic marine cyanobacterium *Synechococcus* species WH7803 to phosphate starvation involves a protein homologous to the periplasmic phosphate-binding protein of *Escherichia coli*. Mol. Microbiol. **10:**181–191.
38. **Singh, M., A. B. Andersen, J. E. McCarthy, M. Rohde, H. Schutte, E. Sanders, and K. N. Timmis.** 1992. The *Mycobacterium tuberculosis* 38-kDa antigen: overproduction in *Escherichia coli*, purification and characterization. Gene **117:**53–60.
39. **Smith, T. F., and M. S. Waterman.** 1981. Identification of common molecular subsequences. J. Mol. Biol. **147:**195–197.
40. **Snider, D., K. B. Bart, B. Bloom, T. Tokunaga, J. R. David, and A. Pio.** 1988. Summary, conclusion and recommendations from the International Workshop on "Research towards global control and prevention of tuberculosis: with an emphasis on vaccine development." J. Infect. Dis. **158:**248–253.
41. **Sturgill-Koszycki, S., P. Schlesinger, P. Chakraborty, P. Haddix, H. Collins, A. Fok, R. Allen, S. Gluck, J. Heuser, and D. Russell.** 1994. Lack of acidification in *Mycobacterium* phagosomes produced by exclusion of the vesicular proton-ATPase. Science **263:**678–681.
42. **Surin, B. P., D. A. Jans, A. L. Fimmel, D. C. Shaw, G. B. Cox, and H. Rosenberg.** 1984. Structural gene for the phosphate-repressible phosphate-binding protein of *Escherichia coli* has its own promoter: complete nucleotide sequence of the *phoS* gene. J. Bacteriol. **157:**772–778.
43. **Surin, B. P., H. Rosenberg, and G. B. Cox.** 1985. Phosphate-specific transport system of *Escherichia coli*: nucleotide sequence and gene-polypeptide relationships. J. Bacteriol. **161:**189–198.
44. **Swofford, D. L., and G. J. Olsen.** 1990. Phylogeny reconstruction, p. 411–501. *In* D. M. Hillis and C. Moritz (ed.), Molecular systematics. Sinauer Associates, Sunderland, Mass.
45. **Tam, R., and M. H. Saier.** 1993. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. Microbiol. Rev. **57:**320–346.
46. **Thangaraj, H. S., T. J. Bull, M. K. Desmet, D. A. Hill, C. Rouse, C. Moreno, and J. Ivanyi.** 1996. Duplication of genes encoding the immunodominant 38 kDa antigen in *Mycobacterium intracellulare*. FEMS Microbiol. Lett. **144:** 235–240.
47. **Torriani, A.** 1990. From cell membrane to nucleotides: the phosphate regulon in *Escherichia coli*. Bioessays **12:**371–376.
48. **Vriend, G.** 1990. WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. **8:**52–56.
49. **Wanner, B. L.** 1993. Gene regulation by phosphate in enteric bacteria. J. Cell. Biochem. **51:**47–54.
50. **Wiker, H. G., and M. Harboe.** 1992. The antigen 85 complex: a major secretion product of *Mycobacterium tuberculosis*. Microbiol. Rev. **56:**648–661.
51. **Young, D. B., S. H. Kaufmann, P. W. Hermans, and J. E. Thole.** 1992. Mycobacterial protein antigens: a compilation. Mol. Microbiol. **6:**133–145.
52. **Young, R. A., B. R. Bloom, C. M. Grosskinsky, J. Ivanyi, D. Thomas, and R. W. Davis.** 1985. Dissection of *Mycobacterium tuberculosis* antigens using recombinant DNA. Proc. Natl. Acad. Sci. USA **82:**2583–2587.