INFECTION AND IMMUNITY, Sept. 2001, p. 5777–5785 Vol. 69, No. 9
0019-9567/01/$04.00+0 DOI: 10.1128/IAI.69.9.5777–5785.2001

Identification and Characterization of Mycobacterial Proteins Differentially Expressed under Standing and Shaking Culture Conditions, Including Rv2623 from a Novel Class of Putative ATP-Binding Proteins

MATTHEW A. FLORCZYK,[1,2] LEE ANN MCCUE,[2] ROBERT F. STACK,[2] CHARLES R. HAUER,[1,2] AND KATHLEEN A. MCDONOUGH[1,2]*

Department of Biomedical Sciences, University at Albany, Albany, New York 12222,[1] and Wadsworth Center, New York State Department of Health, Albany, New York 12202-2002[2]

Received 15 February 2001/Returned for modification 2 April 2001/Accepted 4 June 2001

The environmental signals that affect gene regulation in *Mycobacterium tuberculosis* remain largely unknown despite their importance to tuberculosis pathogenesis. Other work has shown that several promoters, including *acr* (also known as *hspX*) (α-crystallin homolog), are upregulated in shallow standing cultures compared with constantly shaking cultures. Each of these promoters is also induced to a similar extent within macrophages. The present study used two-dimensional gel electrophoresis and mass spectrometry to further characterize differences in mycobacterial protein expression during growth under standing and shaking culture conditions. Metabolic labeling of *M. bovis* BCG showed that at least 45 proteins were differentially expressed under standing and shaking culture conditions. Rv2623, CysA2-CysA3, Gap, and Acr were identified from each of four spots or gel bands that were specifically increased in bacteria from standing cultures. An additional standing-induced spot contained two comigrating proteins, GlcB and KatG. The greatest induction was observed with Rv2623, a 32-kDa protein of unknown function that was strongly expressed under standing conditions and absent in shaking cultures. Analysis using PROBE, a multiple sequence alignment and database mining tool, classified *M. tuberculosis* Rv2623 as a member of a novel class of ATP-binding proteins that may be involved in *M. tuberculosis*'s response to environmental signals. These studies demonstrate the power of combined proteomic and computational approaches and demonstrate that subtle differences in bacterial culture conditions may have important implications for the study of gene expression in mycobacteria.

Worldwide, tuberculosis causes nearly 3 million deaths per year, the highest mortality rate of any single disease caused by an infectious organism (28). Approximately one-third of the world's population is thought to be infected with *Mycobacterium tuberculosis*, which is a facultative intracellular pathogen that resides within the macrophages of its host. For a great majority of these infected individuals, the tubercle bacilli are contained within granulomas by the immune system. The interior of the granuloma is characterized by an acidic pH, the presence of toxic fatty acids, and the low availability of oxygen (7). The survival of *M. tuberculosis* within macrophages and granulomas is likely to depend upon its ability to mount an effective genetic response to these hostile environments.

Several in vitro model systems have been developed using two-dimensional (2-D) gel electrophoresis to examine the protein level response of *M. tuberculosis* to environmental stress and intracellular residence within macrophages (13, 17, 32, 38, 40). Additional studies have shown that expression of the 16-kDa *M. tuberculosis* α-crystallin (Acr) protein (encoded by the *acr* [also known as *hspX*] gene) is strongly induced during stationary-phase growth (11, 41) and during growth in unagitated, submerged cultures in which oxygen is believed to be limiting (5, 34, 35). We have recently shown, using a green fluorescence protein reporter promoter fusion assay, that the *M. tuberculosis acr* promoter is also upregulated in *M. bovis* BCG grown in shallow, standing cultures, compared to cultures constantly agitated on a rocking platform (A. Purkayastha, L. A. McCue, and K. McDonough, submitted for publication).

A growing family of *acr*-coregulated genes (ACG) is being identified and includes genes whose promoters are upregulated under shallow standing growth conditions and during intracellular growth within macrophages (Purkayastha et al., submitted for publication). The present study was done to compare overall mycobacterial protein expression profiles in response to growth under standing and shaking culture conditions and to identify additional ACG candidate genes. We found that standing and shaking culture conditions significantly affect protein expression in mycobacteria. Sequence-based analyses showed that Rv2623, one of the proteins whose expression was most affected by these differential culture conditions, is a member of a novel class of putative ATP-binding proteins that may be involved in environmental gene regulation. These findings indicate that shaking and standing growth conditions significantly impact mycobacterial protein expression and may affect the subsequent outcome of experiments using these bacteria.

* Corresponding author. Mailing address: Wadsworth Center, New York State Department of Health, P.O. Box 22002, Albany, NY 12202-2002. Phone: (518) 486-4253. Fax: (518) 474-3181. E-mail: Kathleen.McDonough@wadsworth.org.

MATERIALS AND METHODS

Bacterial growth and labeling conditions. *M. bovis* BCG (Pasteur strain; Trudeau Institute) was grown in mycomedia (liquid 7H9 medium [Difco] supplemented with 0.5% [vol/vol] glycerol, 10% [vol/vol] oleic acid-albumin-dex-

trose-catalase [Difco] and 0.05% [vol/vol] Tween 80), as previously described (18). Standing cultures were grown undisturbed in 10 ml of mycomedia (approximately 2 mm deep) in 75-cm^2 flat-bottom tissue culture flasks (catalog no. 353083; Falcon) with the caps tightly sealed, lying flat at 37°C. Shaking cultures were grown on a gently rocking platform (Model 55 Rocking Platform; Reliable Scientific, Inc.) at 24 cycles per minute.

Bacteria were labeled with [^{35}S]-L-methionine and [^{35}S]-L-cysteine (Pro-mix; 100 μCi/ml; Amersham) for a 24-h period prior to harvesting the bacteria for 2-D gel electrophoresis. Radioactive label was carefully added to the standing cultures, without mixing, to minimize any disruption of the bacterial sediment at the bottom of the tissue culture flask.

Sample preparation and 2-D gel electrophoresis of *M. bovis* BCG proteins. Bacteria were harvested by centrifugation and washed three times with ice-cold DPBS (Dulbecco's phosphate-buffered saline [10 mM sodium phosphate; 126 mM NaCl, pH 7.2]) plus 0.2% (wt/vol) EDTA (disodium EDTA dihydrate) containing a protease inhibitor cocktail (catalog no. P8340; Sigma) (DPBS-I). Cells were then resuspended in Tris-sodium dodecyl sulfate (Tris-SDS) buffer (0.3% [wt/vol] SDS and 50 mM Tris-HCl, pH 8.0) and lysed by several rounds of sonication and freeze-thawing as described previously (17).

Radiolabeled *M. bovis* BCG proteins were separated by 2-D SDS-polyacrylamide gel electrophoresis (PAGE) as described previously, with the following modifications (17). Isoelectric focusing (IEF) was performed using IEF tube gels (1.5 mm [inner diameter] [i.d.] by 16 cm [length]) with a final concentration of 2% (vol/vol) each of Bio-Lyte 4–6, 5–7, and 6–8 ampholytes (Bio-Rad) for 18 h at a constant voltage of 667 V. Proteins were separated in the second dimension on 1.5-mm-thick, 16-cm (length) SDS–10% PAGE gels. Approximately 5×10^6 dpm of radiolabeled bacterial protein was loaded onto each gel.

Concentrations of unlabeled mycobacterial proteins were estimated using the NanoOrange Protein Quantitation Kit (Molecular Probes). Approximately 500 μg of total protein was loaded onto each IEF tube gel (3 mm [i.d.] by 15 cm [length]) with a final concentration of 4% each of Bio-Lyte 4–6, 5–7, and 6–8 ampholytes (Bio-Rad). Protein samples were focused as described above and separated in the second dimension on 2-mm-thick, 16-cm (length), SDS–10% PAGE gels. Gels were stained for 1 h in 0.05% (wt/vol) Coomassie R-250 and destained in 5% (vol/vol) methanol–7% (vol/vol) acetic acid overnight. Coomassie-stained 2-D gels were analyzed and compared using the ImageMaster computer software (Amersham Pharmacia Biotech) and ZERO-Dscan (version 1.0; Scanalytics, Billerica, Mass).

In-gel proteolytic digestion of proteins. Protein spots of interest were isolated from Coomassie blue-stained 2-D PAGE gels, destained, and partially dehydrated with 0.1 M Tris (pH 8.0)–50% (vol/vol) acetonitrile for 30 min at 37°C, and this was followed by 5 min in a sonicating water bath to ensure that all visible Coomassie stain was removed from the gel pieces. Gel pieces were then dried in a Speed-Vac at ambient temperature under vacuum for approximately 30 min and rehydrated in 0.1 M Tris (pH 8.0)–0.05% (wt/vol) *n*-octylglucoside (Sigma) containing modified trypsin (0.1 μg/μl; Boehringer Mannheim) or endoproteinase Glu-C (protease V8) (Roche Molecular Biochemicals). Proteins were digested for 18 to 24 h at 37°C.

After digestion, samples were mixed with an equal volume of 0.1 M Tris (pH 8.0)–50% (vol/vol) acetonitrile, and dithiothreitol was added to a final concentration of 1 mM before incubating at 50°C for 20 min. Free sulfhydryl groups of the digested peptide fragments were alkylated by adding Iodo-Acetamide (Sigma) to a final concentration of 3 to 4 mM and incubating at 37°C for 40 min. Peptide fragments were recovered by extracting the gel pieces twice with 60% (vol/vol) acetonitrile–0.1% (vol/vol) trifluoroacetic acid and shaking for 40 min at room temperature. Peptide extracts were combined and concentrated 10-fold in a Speed-Vac vacuum centrifuge prior to mass spectrometry (MS) analysis.

MS analysis. The peptides extracted from gel slices were desalted using a C_{18} ZipTip (Millipore, Bedford, Mass.) and eluted onto a matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) sample plate with 60% (vol/vol) acetonitrile, using 1-μl wash volumes. After partial air drying of the wash volumes on the sample plate, 0.5 μl of MALDI matrix solution containing alpha-4-hydroxy-cinnamic acid (10 mg/ml; Sigma) was added to each sample spot.

MALDI-TOF MS analysis was performed with a Bruker REFLEX instrument (Bruker Daltonics, Inc., Billerica, Mass.) equipped with delayed ion extraction. Monoisotopic mass values were identified from 100 laser-shot, reflectron spectra obtained at an accelerating potential of 25,000 kV. Accurate mass values were obtained by recalibrating the instrument using modified trypsin self-digestion product ions as internal standards.

Ion trap mass spectrometry was performed with a Finnigan LCQ DECA (Thermo Finnigan, San Jose, Calif.) equipped with a New Objective PicoView nano-ESI source (New Objective, Inc., Cambridge, Mass.). Peptide digests were loaded onto a Michrom peptide CapTrap (Michrom BioResources, Inc., Auburn, Calif.) plumbed to a 75-μm-i.d. New Objective PicoFrit liquid chromatography (LC) column packed with 5 cm of Aquasil C_{18}. Peptide digests were desalted on the CapTrap and then eluted from the reverse-phase column with a gradient of 2 to 70% (vol/vol) acetonitrile–0.1% (vol/vol) formic acid versus 0.1% (vol/vol) formic acid in water over 70 min. The LC gradient was generated by an ABI 140B dual-syringe pump (ABI/Kratos, Foster City, Calif.) and split to a flow of less than 500 nanoliters per min with an Acurate flow splitter (LC Packings, San Francisco, Calif.). Data-dependent LC-MS, zoom scans, and LC-MS-MS of eluting peptides were acquired in repetitive 3-s cycles by the triple-play software of the LCQ system.

Database searches. Genomic databases were searched using MS-FIT (http://prospector.ucsf.edu). The MALDI-TOF MS peptide mass fingerprint data were used to search the National Center for Biotechnology Information (NCBI) nonredundant database using a species-limited search filter, which restricted the search to *Mycobacterium* species. The search parameters were protein molecular mass ranging from 1,000 to 100,000 Da, trypsin digests (one missed cleavage), peptide mass tolerance of ±50 ppm, monoisotopic peptide masses, and cysteines modified by carbamidomethylation. Positive protein identification required matched peptides to cover a minimum of 15% of the protein sequence if the predicted molecular weight and pI of the identified protein were consistent with those of the observed protein spot isolated from the gel. A second search was routinely performed using only unassigned peptide masses to identify additional proteins that may be present in the isolated protein spot (12). MS/MS spectra data sets were used to query the complete *M. tuberculosis* H37Rv genome (NCBI accession no. NC_000962) with the SEQUEST protein identification software (15).

Western analysis. Approximately 50 μg of total mycobacterial protein was separated on an SDS–12.5% PAGE gel for each bacterial growth condition. After electroblotting onto polyvinylidene difluoride membranes (Millipore), the blots were blocked with buffer containing 5% (wt/vol) milk and 0.05% (vol/vol) Tween 20 (Sigma) in Tris-buffered saline (50 mM Tris, 0.5 M NaCl, pH 7.6). Rabbit antisera generated against *M. tuberculosis* isocitrate lyase (ICL) and malate synthase (both kindly provided by D. Russell) and catalase-peroxidase (KatG) (a generous gift of S. Cole) were diluted 1:10,000 in blocking buffer. The ICL antibody was generated against the whole recombinant protein from *M. tuberculosis* (10). The malate synthase antibody was generated against amino acids 424 to 439 (QNTMKIGIMDEERRTT) of the *M. tuberculosis* H37Rv malate synthase protein. Primary antibodies were detected using goat anti-rabbit immunoglobulin G antibodies conjugated to horseradish peroxidase (Jackson Labs) at a 1:5,000 dilution. Chemiluminescence detection was performed using the ECL detection kit (Amersham). Western blots were analyzed using Zero-Dscan (version 1.0; Scanalytics).

Sequence analysis. Pairwise local sequence alignments were performed with BLAST (1). Database mining and multiple sequence alignment of the Rv2623 family was performed with PROBE (16, 23). Given a single protein query sequence, PROBE performs a transitive BLAST search to identify a set of related sequences in a protein database and then proceeds to align only functionally constrained regions of the proteins, creating a "superlocal" multiple sequence alignment model consisting of individual motifs. The alignment model is then used to search the database for additional sequences. PROBE iterates between refining the alignment model and database searching until no additional sequences are recruited. PROBE calculates maximum a posteriori (MAP) scores for the model as well as for each motif. MAP scores are a measure of how well the alignment model describes the data, relative to unaligned random background (MAP scores represent the log of the probability ratio). A MAP score above zero indicates that the motif is more likely to belong to the alignment model than to unaligned random background.

PROBE was run with the default settings using the Rv2623 protein sequence as the query and the NCBI nonredundant protein database (29 November 2000 version) to identify the ATP-binding protein family. Sequences near the cutoff for family membership (E-value ≤ 0.01) were further subjected to a jackknife test for statistical significance. Jackknife tests were performed by removing the sequence to be tested, as well as any pairwise similar sequences, from the family sequence set. Pairwise similar sequences were those with BLAST scores of >31 bits. The sequences remaining in the family were then realigned, and the nonredundant database was searched using PROBE. Sequences recruited back into the family by this procedure were considered true family members.

RESULTS AND DISCUSSION

[^{35}S]-L-methionine-cysteine labeling of mid-logarithmic-phase shaking and standing cultures. The expression of *M.*

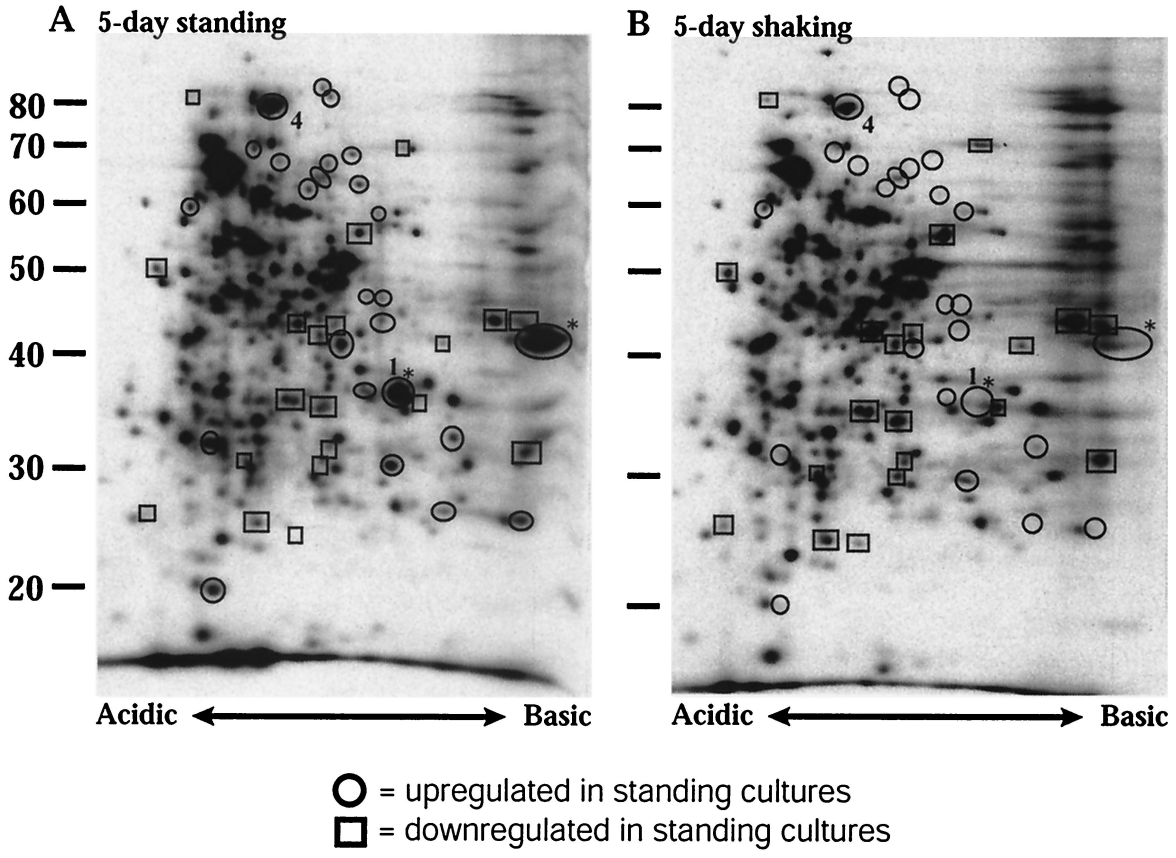

FIG. 1. 2-D gel analysis of metabolically labeled 5-day *M. bovis* BCG cultures growing under standing (A) or shaking (B) growth conditions. Approximately 5×10^6 dpm of radiolabeled bacterial protein was loaded onto each gel. Proteins were first separated by IEF (pH 4 to 8, left to right) and then by SDS-PAGE (10% polyacrylamide) in the second dimension. Open circles indicate proteins that were more abundant under the standing growth condition, and open squares indicate proteins that were less abundant. Asterisks mark the positions of two highly induced mycobacterial proteins under standing growth conditions. Spots 1 and 4 indicate the positions of two mycobacterial proteins that were identified by 2-D GEMS analysis in Fig. 2. Approximate molecular masses are in kilodaltons.

tuberculosis ACG family promoters is upregulated similarly within macrophages and under shallow standing, undisturbed growth conditions, compared to constantly agitated culture conditions (Purkayastha et al., submitted for publication). We reasoned that the ACG family could belong to a larger set of genes that are differentially expressed under standing and shaking growth conditions.

Mid-logarithmic-phase *M. bovis* BCG cultures pulse-labeled with [^{35}S]-L-methionine-cysteine were examined by 2-D gel electrophoresis to determine the extent of differential protein synthesis under standing and shaking growth conditions (Fig. 1). Levels of at least 25 proteins increased, and those of 20 proteins decreased, in the standing culture relative to the shaking culture. The increases in expression of two of these proteins were particularly notable (Fig. 1A).

Identification of differentially regulated proteins in standing and shaking cultures. Steady-state protein levels were examined by comparing total protein lysates of mid-logarithmic-phase shaking and standing cultures on Coomassie blue-stained 2-D protein gels. Comparison of these gels showed 15 protein spots that were upregulated under standing and shaking culture conditions and 2 protein spots that were downregulated. Differences in the protein expression profiles of radiolabeled (Fig. 1) and Coomassie-stained (Fig. 2) gels are largely due to differences in the sensitivity and nature of these two protein detection methods. Coomassie-stained gels detect accumulated, steady-state proteins and have a detection limit of 0.3 to 1 μg per gel spot. In contrast, only de novo-synthesized proteins appear on [^{35}S]-L-methionine-cysteine-labeled protein gels and are detected at a limit of several picograms of protein, depending on amino acid composition (2).

Several well-resolved differentially expressed protein spots were excised from the Coomassie-stained 2-D standing culture gels (Fig. 2A), including two (spots 1 and 4) whose positions were consistent with the positions of two of the differentially regulated spots on the radiolabeled gels (Fig. 1). 2D gel electrophoresis and MS (2D-GEMS) analysis was used to assign putative identities to a total of four of these protein spots (Fig. 2A) that were upregulated under standing and shaking cultures in at least two independent experiments.

(i) Rv2623. Protein gel spot number 1 (Fig. 2A) was identified as Rv2623, a putative *M. tuberculosis* open reading frame of unknown function. No corresponding protein spot was observed on the Coomassie-stained shaking gel. The equivalent spot was also dramatically increased under standing conditions in radiolabeled 2-D gels (spot 1; Fig. 1A). Eight peptides from the tryptic peptide mixture matched the predicted protein sequence of Rv2623 and provided 25% coverage of this 297-

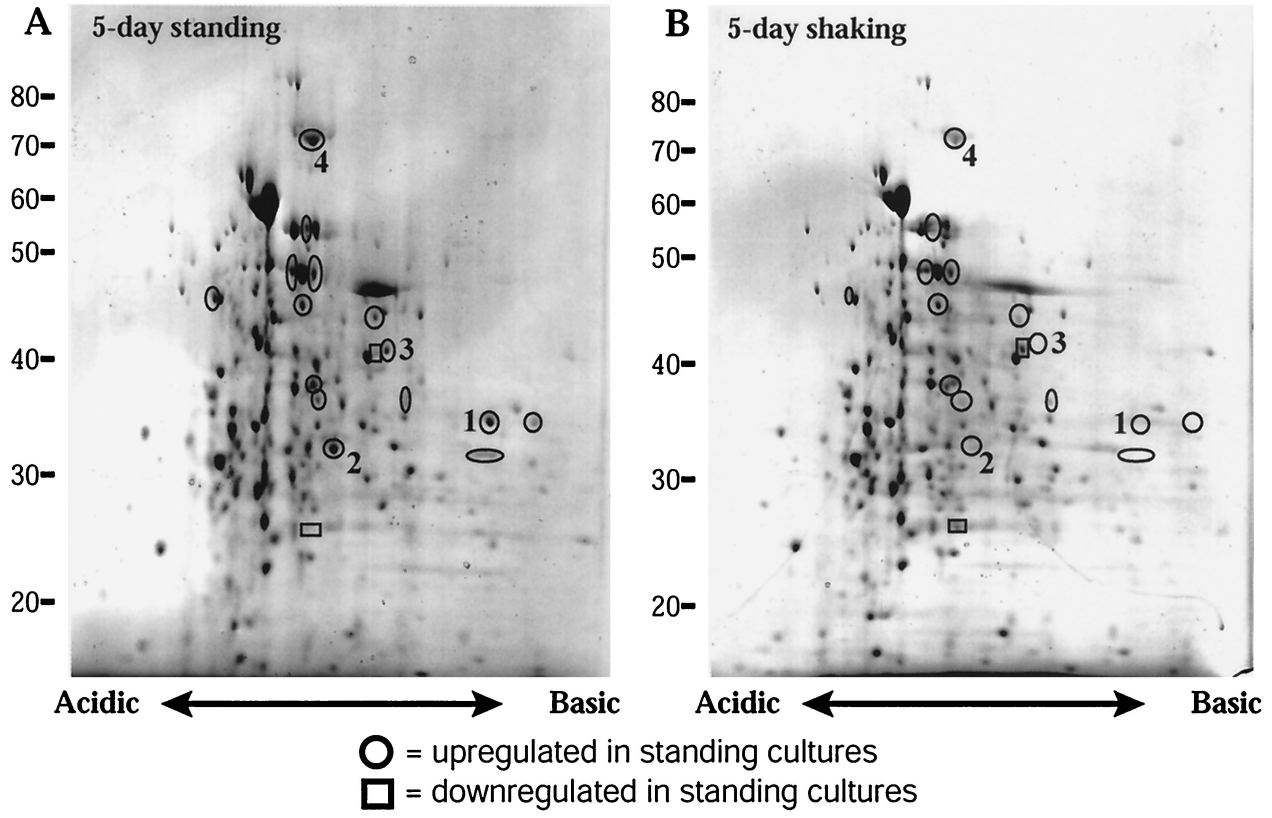

FIG. 2. 2-D gel analysis of steady-state levels of *M. bovis* BCG proteins growing in culture for 5 days under standing (A) or shaking (B) growth conditions. Approximately 500 μg of total protein was loaded onto each gel. Proteins were separated as described for Fig. 1 and stained with Coomassie blue dye. Open circles indicate proteins that were more abundant under standing growth conditions; open squares indicate the position of proteins that were less abundant. Large numbers in boldface type indicate protein spots identified by 2D-GEMS and correspond to the protein spot numbers in Table 1.

amino-acid hypothetical *M. tuberculosis* protein (Table 1). Digestion with endoproteinase Glu-C resulted in eight peptide fragments that matched Rv2623 and provided 31% coverage of the predicted protein sequence. Rv2623 has a predicted molecular mass of 31.65 kDa and an isoelectric point of 5.46, which are consistent with the position of protein spot 1 on the 2-D gel (Fig. 2A).

BLAST analysis revealed that Rv2623 shares significant pairwise homology with four other hypothetical *M. tuberculosis* proteins (Rv2005c [52% identity over 295 amino acids], Rv1996 [40% identity over 315 amino acids], Rv2026c [49% identity over 295 amino acids], and Rv2624c [27% identity over 284 amino acids]), as well as a number of hypothetical *Streptomyces coelicolor* proteins, all of unknown function. To gain additional structural and functional insight into the potential role of Rv2623 in *M. tuberculosis* biology, we employed a Bayesian Gibbs sampling method, PROBE, to construct a multiple-alignment model of a family of related proteins (23). PROBE identifies and aligns only those regions of protein sequence patterns that are functionally constrained and therefore essential to the protein family as a whole. This results in a superlocal alignment model that may consist of several colinear, conserved motifs and describes the protein domain common to that family.

A total of 153 proteins were identified in the Rv2623 family, and the alignment model consisted of three statistically significant motifs, presented as sequence logos in Fig. 3. This Rv2623 family is similar to the Pfam (3) Usp (universal stress protein) family (PF00582). A majority of the proteins identified in this family were hypothetical proteins of unknown function, but two have been defined with respect to function or structure. These include MJ0577, an ATP-binding protein from the hyperthermophile, *Methanococcus jannaschii*, and UspA (universal stress protein A) from *Escherichia coli.* The relationship between Rv2623 and the better-characterized proteins in this family may provide further insight into the biological role of this protein.

The three-dimensional crystal structure of MJ0577 has been shown to contain a bound ATP, although the ATP-binding fold is significantly different from that of other ATP-binding proteins (42). The three motifs of the PROBE family model encompass the ATP-binding domain of MJ0577 (Fig. 3), suggesting that this model describes a novel ATP-binding domain family. Biochemical experiments showed that MJ0577 is capable of hydrolyzing ATP only in the presence of *M. jannaschii* crude cell extract, suggesting that it functions biochemically as an ATPase or an ATP-binding molecular switch for a cellular process (42).

UspA is an autophosphorylating serine-threonine phosphoprotein that is phosphorylated under conditions that induce growth arrest (8). Its expression is also greatly increased during growth arrest or as a result of diverse stress conditions that

TABLE 1. Summary of differentially expressed *M. bovis* BCG proteins identified by 2D-GEMS from Coomassie-stained 2-D gels under standing and shaking growth conditions

Spot no.	Protein name (NCBI accession no.)	MW[a]	pI	Functional category	Mass of peptide MH^+ (Da)	Change in mass	Start position–end position	Tryptic peptide sequence
1	Rv2623 (gi|2104288)	31,652.3	5.46	Unknown	515.2	−0.07	124–127	(R)WPGR(L)
					859.5	0.06	224–230	(R)LAGWQER(Y)
					924.5	−0.02	70–77	(R)HLIDDALK(V)
					926.5	0.08	63–69	(R)WQQDHGR(H)
					933.5	−0.02	231–238	(R)YPNVAITR(V)
					1,101.7	0.04	128–138	(R)LLGSVSSGLLR(H)
					1,273.7	0.02	253–264	(R)SEEAQLVVVGSR(G)
					1,936.0	0.004	267–286	(R)GGYAGMLVGSVGETVAQLAR(T)
2	cysA2 (gi|3122949)	31,014.7	5.14	Intermediary metabolism and respiration	729.4	−0.03	159–164	(K)NLIDVR(S)
					1,132.6	0.01	239–247	(R)SSHTWFVLR(E)
					1,184.5	−0.06	58–67	(R)DFVDAQQFSK(L)
					1,445.8	−0.00	146–158	(R)AFRDEVLAAINVK(N)
					1,587.8	−0.04	235–247	(R)IGERSSHTWFVLR(E)
					1,743.7	−0.11	22–36	(K)VVFVEVDEDTSAYDR(D)
					2,549.2	−0.05	22–44	(K)VVFVEVDEDTSAYDRDHIAGAIK(L)
2	gap (gi|3122120)	35,956.1	5.19	Intermediary metabolism and respiration	599.3	0.01	16–19	(R)NFYR(A)
					807.5	0.06	52–58	(K)FDSILGR(L)
					807.5	0.06	233–239	(K)LDGYALR(V)
					819.5	0.05	5–12	(R)VGINGFGR(I)
					1,134.6	−0.01	169–178	(K)VLDDEFGIVK(G)
					1,174.9	0.12	314–327	(K)VVSWYDNEWGYSNR(L)
					1,871.1	0.09	240–257	(R)VPIPTGSVTDLTVDLSTR(A)
4a	glcB (gi|2497795)	80,403.4	5.03	Intermediary metabolism and respiration	752.4	−0.01	334–339	(R)SLMFVR(N)
					772.4	−0.04	313–318	(R)VLNRDR(N)
					1,199.6	−0.006	586–595	(K)ELAWAPDEIR(E)
					1,280.7	−0.0	453–463	(R)VVFINTGFLDR(T)
					1,300.7	−0.04	639–649	(R)ISSQLLANWLR(H)
					1,321.6	−0.05	72–82	(R)VIEPIDMDAYR(Q)
					1,509.7	−0.04	622–634	(K)VPDIHDVALMEDR(A)
					1,535.8	0.04	319–333	(R)NYTAPGGGQFTLPGR(S)
					1,619.9	0.06	297–312	(K)GDLAAAVDKDGTAFLR(V)
					1,723.0	0.05	42–57	(K)VVADLTPQNQALLNAR(D)
					2,574.2	−0.04	485–508	(K)SQPWILAYEDHNVDAGLAAGFSGR(A)
4b	Catalase-peroxidase (gi | 1150710)	80,476.1	5.09	Virulence, detoxification, adaptation	888.4	−0.04	180–187	(K)TFGFGFGR(V)
					1,335.7	0.01	694–705	(R)VDLVFGSNSELR(A)
					1,437.7	−0.06	558–571	(K)AAGHNITVPFTPGR(T)
					1,541.8	0.07	357–373	(K)DGAGAGTIPDPFGGPGR(S)
					1,774.8	0.01	28–42	(K)YPVEGGGNQDWWPNR(L)
					1,938.0	0.06	499–515	(R)LQPQVGWEVNDPDGDLR(K)

[a] Molecular weight.

cause perturbations in unrestricted balanced growth (24). It has been suggested that UspA is a kinase that plays a role in the posttranslational modification of other proteins during growth arrest (8, 24). UspA may also be distantly related to eukaryotic MADS box transcription factors (21).

Nine additional *M. tuberculosis* proteins were included in the Rv2623 PROBE family: Rv1636, Rv2026c, Rv1996, Rv2005c, Rv2624c, Rv2028c, Rv3134c, Rv2319c, and Rv1028c. The large number of these putative ATP-binding proteins may indicate that this type of ATPase or ATP-binding molecular switch is functionally important for *M. tuberculosis*. In addition, Rv2623 and several other *M. tuberculosis* proteins (Rv2026c, Rv1996, Rv2005c, Rv2624c, Rv2028c, and Rv2319c) contain two copies of the domain described by the PROBE model. These proteins may belong to a unique subclass of proteins within this family that are capable of binding two ATP molecules per protein monomer.

Recently it has been shown that the *M. tuberculosis* Rv2623 protein is expressed at high levels during infection of the macrophage cell line THP1 (19). This is consistent with preliminary promoter analysis experiments in our laboratory indicating an increase in transcription from the Rv2623 promoter when *M. bovis* BCG cells are grown within J774.16 macrophages (M. A. Florczyk and K. A. McDonough, unpublished data). The possibility that Rv2623 is a member of the ACG superfamily that plays a role in macrophage infection is presently under investigation.

(ii) CysA2-CysA3. Protein gel spot 2 in Fig. 2A was identified as a putative thiosulfate sulfurtransferase, which is a rhodanese-like protein (4). The amount of protein in this spot was approximately sixfold greater under standing culture conditions than under shaking culture conditions. The peptide mixture contained 7 peptides that matched CysA2 (Rv0815c) and CysA3 (Rv3117), which are 100% identical, providing 23%

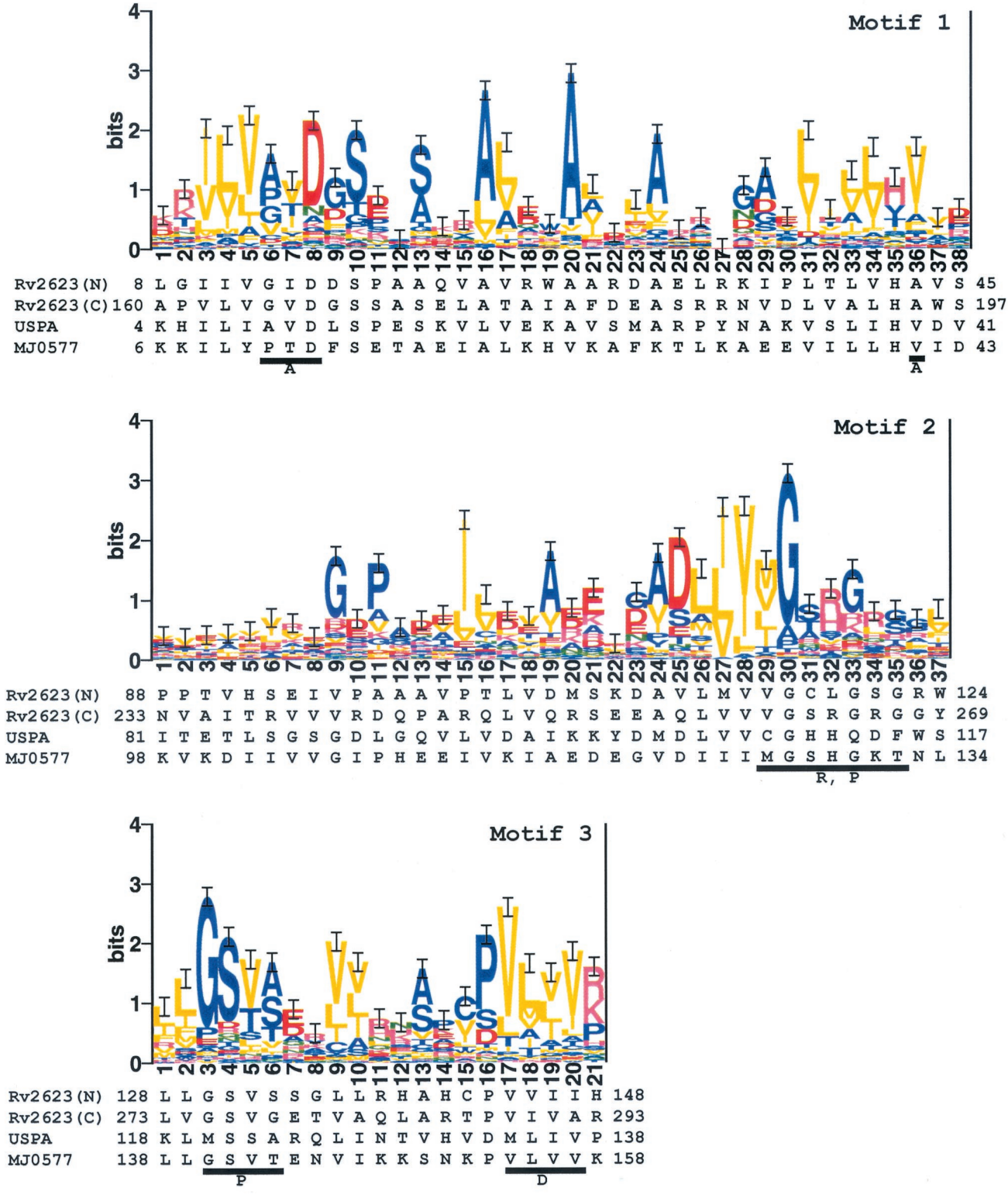

FIG. 3. Sequence logos (30) of the Rv2623 protein family motifs representing the multiple sequence alignment output of PROBE. Below each motif is the corresponding protein sequence for that motif in Rv2623, UspA, and MJ0577. Rv2623 contains two complete sets of these motifs, with N corresponding to the copy in the amino portion of the protein and C denoting the copy closer to the carboxy terminus. The MAP scores for the motifs are 1162.49 (motif 1), 980.12 (motif 2), and 832.65 (motif 3). The underlined sequences indicate amino acids involved in ATP binding in MJ0577 (42). Abbreviations: A, adenine; R, ribose; P, phosphate; D, dimer interface.

coverage of these 277 amino acid proteins (Table 1). These proteins have a predicted molecular mass of 31 kDa and a predicted isoelectric point of 5.14, consistent with the gel position of the isolated spot. The *M. tuberculosis* genome encodes a total of four putative thiosulfate transferases (CysA2, CysA3, SseA [Rv3283], and SseB Rv2291]), suggesting an important role for these proteins in mycobacteria. BLAST analysis showed that *M. tuberculosis* CysA2-CysA3 shares 50% identity with *M. tuberculosis* SseA over 271 amino acids and 26% identity with *M. tuberculosis* SseB over 253 amino acids.

Rhodanese-like proteins have been found in a variety of organisms, including bacteria, plants, and mammals (9). Rhodanese reacts with sulfur-containing anions such as thiosulfate and transfers them to thiophilic acceptor molecules (27). An interesting aspect of this enzyme in the context of our study is its possible role in the assembly of iron-sulfur clusters (4Fe-4S). In *E. coli*, iron-sulfur clusters have been identified as biosensors of both oxygen and iron concentrations (33). Conditions such as low oxygen tension within the standing cultures may create the need to increase the intracellular pools of labile sulfide and iron-sulfur clusters in the bacterium. This may also be important in the response of *M. tuberculosis* to low-oxygen environments within macrophages and granulomas.

(iii) Gap. Protein gel spot 3 (Fig. 2A) was identified as Gap (Rv1436), glyceraldehyde 3-phosphate dehydrogenase (GAPDH) (4). The corresponding spot was not detected under shaking culture conditions. The peptide mixture contained seven peptides that matched the Gap protein sequence (Table 1) and provided 20% coverage of this 339-amino-acid protein.

GAPDH is a key enzyme in intermediary metabolism and operates at the sixth step in glycolysis, during which glucose is converted to pyruvate with the subsequent net generation of two ATP molecules (20). In yeast, pyruvate can subsequently be anaerobically metabolized to ethanol by the reduction of acetaldehyde by alcohol dehydrogenase and NADH. This reaction regenerates the pool of NAD^+, allowing glycolysis to occur under anaerobic conditions. A similar situation may occur in *M. tuberculosis*. Using the Kyoto Encyclopedia of Genes and Genomes (http://www.genome.ad.jp/kegg/), four open reading frames in *M. tuberculosis* (*adhA* [Rv1862)], *adhB* [Rv0761c], *adhC* [Rv3045], and *adh*E2 [Rv2259]) were identified as alcohol dehydrogenase-coding regions that may play a role in the production of NAD^+ under anaerobic conditions to facilitate anaerobic glycolysis (26). The upregulation of GAPDH in standing BCG cultures may be necessary to satisfy the energy demands of these bacteria during adjustment to growth under reduced oxygen tension in standing cultures.

(iv) GlcB-KatG. Protein gel spot 4 was found to contain two proteins. The amount of protein in this spot was approximately 3- to 14-fold greater under standing culture conditions than under shaking culture conditions in different experiments. Increased de novo synthesis of a protein at this gel position was also observed in radiolabeling experiments (spot 4; Fig. 1A). These proteins were identified as an *M. tuberculosis* malate synthase (Rv1837c, GlcB) and an *M. tuberculosis* catalase-peroxidase protein (Rv1908c, KatG) (4). The predicted molecular mass (80.4 kDa) and pI (4.9) for *M. tuberculosis* GlcB are essentially identical to those for KatG, making it difficult to resolve these two proteins by 2-D PAGE. Eleven tryptic peptide fragments (Table 1) matched the *M. tuberculosis* GlcB protein and resulted in 18% coverage of this 741-amino-acid protein. KatG was identified as a possible second component with 11% coverage on a second-pass search after the subtraction of all tryptic peptides matching the predicted *glcB* gene product.

Further analysis by MS-MS unambiguously confirmed the presence of both GlcB and KatG proteins in this spot, and no other significant matches were identified. Data from a nano-ESI LC-MS-MS ion trap analysis of tryptic digest peptides from spot 4 were used to query the complete *M. tuberculosis* H37Rv genome (NCBI accession no. NC_000962) with the SEQUEST protein identification computer program. Strict peptide sequence correlation parameters (cross-correlation > 2, initial sequence ranking of 1 or 2, and visual inspection of sequence assignments) were used, and MS-MS spectra that correspond to 55% of the amino acid count of GlcB and 34% of that of KatG were observed.

Malate synthase is the second bypass enzyme of the glyoxylate cycle, an alternative to the tricarboxylic acid cycle that allows isocitrate to bypass the remainder of the tricarboxylic acid cycle. Glyoxylate metabolism has previously been implicated in the adaptation of *M. tuberculosis* to survival under anaerobic growth conditions (36). Wayne and Lin proposed that a shift to the glyoxylate synthesis pathway increases the supply of NAD^+ and the subsequent production of ATP, allowing the bacteria to respond appropriately to growth under low oxygen tension (36).

KatG is encoded by the *M. tuberculosis katG* gene. Mutations in *katG* result in isoniazid-resistant organisms (22). Mycobacterial catalases have also been implicated as virulence factors, because they provide a means to protect the bacteria from a variety of reactive oxygen intermediates produced by activated host macrophages (29, 31, 39). KatG has been shown to play an important role in the growth and survival of *M. tuberculosis* in both murine and guinea pig models of tuberculosis infection (14).

Western analysis was used to evaluate the relative contributions of GlcB and KatG to the quantitative changes observed in spot number 4 (Fig. 2). Three distinct protein bands were detected in bacteria from shaking cultures using an anti-GlcB antibody, but only the two higher-molecular-mass bands were present in standing conditions. The intensity of the 72-kDa band, which is consistent with the position of protein spot 4 (Fig. 2A), was similar under both conditions (Fig. 4B). Western analysis of these same cultures using an anti-KatG antibody showed no differences in KatG expression (Fig. 4C), with four protein bands of similar intensity detected under each growth condition. ICL, the first bypass enzyme of the glyoxylate cycle, appeared as one band that was unaffected by growth under shaking and standing conditions at 5 days (Fig. 4D), although levels in shaking cultures were reduced in 7-day cultures relative to standing cultures (not shown).

The differences in the GlcB banding patterns may represent differential processing or degradation of this protein in shaking versus standing growth conditions. The apparent increase in the amount of protein in gel spot number 4 may represent a posttranslational modification to GlcB and/or KatG that results in a shift in isoelectric point of the protein during growth under shaking conditions. A change in isoelectric point could alter migration of these proteins on a two-dimensional gel, without significantly affecting the mobility of the protein on a one-dimensional SDS-PAGE gel. This is consistent with a

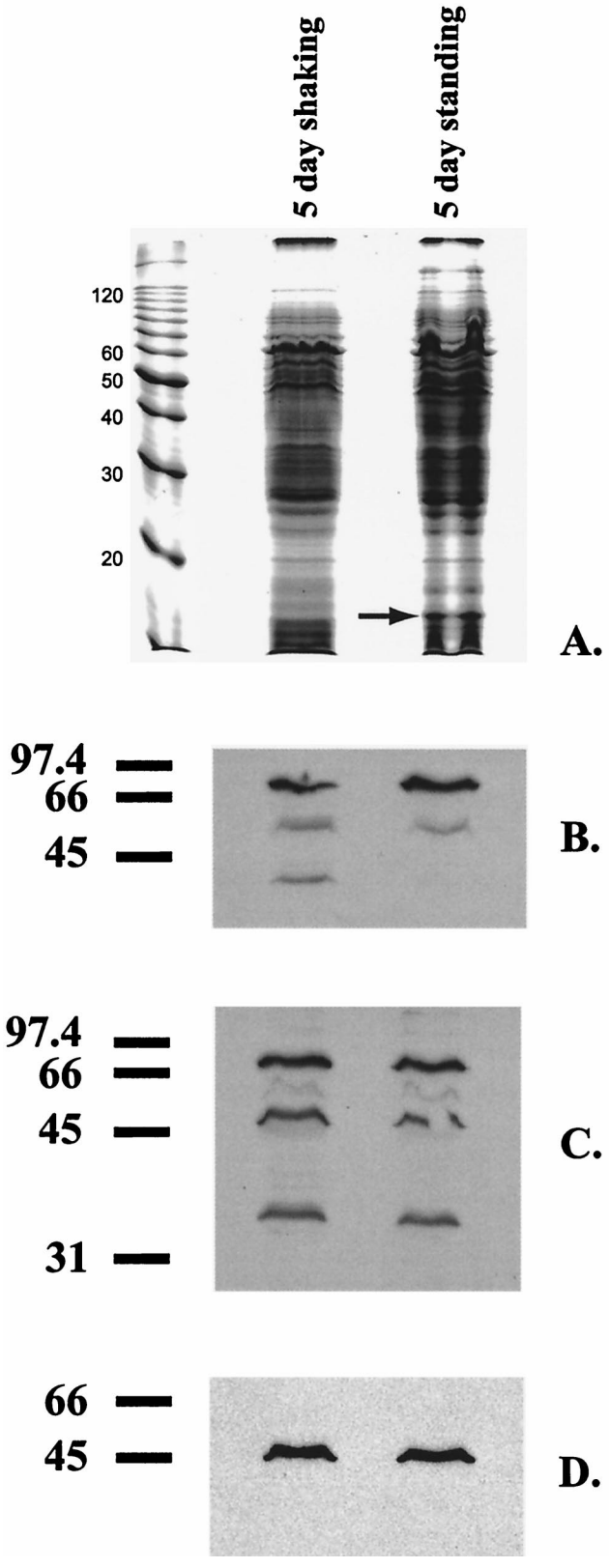

FIG. 4. Western analysis of 5-day *M. bovis* BCG cultures grown under shaking and standing growth conditions for the expression of ICL, malate synthase, and catalase-peroxidase (KatG) in *M. bovis* BCG. (A) Coomassie blue-stained SDS-PAGE gel of *M. bovis* BCG proteins. (A) Arrows indicate the position of the protein band identified as the α-crystallin (Acr) protein. Immunoblots were probed with anti-malate synthase (B), anti-catalase-peroxidase (C), or anti-ICL polyclonal antibodies (D). Molecular mass markers are in kilodaltons.

change in shape and position, rather than intensity, of this spot in one of three 2D gel experiments. A modification of this kind could affect the protein's sensitivity to further processing and explain the differences in the GlcB protein profile observed under shaking and standing growth conditions (Fig. 4B). Additional studies are needed to resolve this issue.

Previous studies showed that the specific activity of ICL, but not that of malate synthase, increased fourfold when the bacteria were grown submerged in deep liquid cultures without agitation (36). Variation in the results between this and the previous study is probably due to differences in the standing culture models (shallow versus deep medium depth) or bacterial growth stage (day 5 logarithmic versus day 19 to 28 stationary phase).

(v) Acr (HspX). In Fig. 4A, we observed a low-molecular-weight protein that was significantly induced in standing mycobacterial cultures. Tryptic peptides from this 16-kDa protein were analyzed by MS-MS. MS-MS spectra of the most abundant protein in each band corresponded to 72% of the amino acid count of the *M. tuberculosis* HspX (Acr) protein. This finding is consistent with the work of Purkayastha et al. who have shown that the *acr* promoter is upregulated under shallow standing growth conditions (Purkayastha et al., submitted for publication).

Relevance of standing and shaking culture conditions as a model system. The aim of this study was to begin to provide a more detailed understanding of the differences between mycobacteria grown under shallow standing and shaking culture conditions through the identification of differentially expressed proteins. These studies showed that a number of proteins are differentially regulated in response to growth under shallow standing culture conditions compared to constantly shaking cultures.

Several of the proteins identified in this study have functions that are consistent with growth under reduced oxygen conditions. The putative function of CysA2-CysA3 is consistent with a role in oxygen sensing, and preliminary experiments suggest that Rv2623 is transcriptionally induced in response to growth under low-oxygen as well as standing culture conditions (Florczyk and McDonough, unpublished data.). Additionally, the *M. tuberculosis* Acr protein, which is highly expressed under hypoxic conditions (11, 41), was induced under standing culture conditions in this study (Fig. 4).

However, the contribution of limiting-oxygen tension to differential protein expression in our model system requires further investigation. The headspace air-to-liquid ratio in our shallow standing cultures is 20- to 60-fold greater and the depth of the medium in our cultures is 20- to 35-fold less than what is typically used in the low-oxygen dormancy model developed by Wayne and colleagues (6, 36, 37). Nonetheless, it is possible that localized oxygen depletion occurs within the bacterial sediment of our cultures, despite an increased availability of oxygen in our system. Additional factors that may be contributing to the differential protein expression of mycobacteria growing under the shallow standing culture condition include, but are not limited to, changes in pH, availability of nutrients, metabolic waste buildup, quorum sensing, and bacterial cell-cell interactions.

This work demonstrates the power of combined proteomic and computational approaches to explore questions about differential protein expression. These studies also show the extent to which subtle differences in culture conditions can impact mycobacterial protein expression, which may affect the outcome of subsequent experiments. These findings should be of particular interest to researchers involved in studying the genetic expression of bacterial genes and proteins.

ACKNOWLEDGMENTS

We thank David Russell and Kerstin Hoener zu Bentrup for providing antibodies to *M. tuberculosis* ICL and malate synthase proteins and Stuart Cole and Brigitte Saint-Joanis for antibodies directed against *M. tuberculosis* catalase peroxidase. We also thank C. Wilcox, R. Jovell, M. Ryan, A. Purkayastha, and C. E. Lawrence for useful discussions and critical reading of the manuscript. We acknowledge support from the Wadsworth Center's Computational Molecular Biology and Statistics Core and Mass Spectrometry Facilities.

This work was supported in part by National Institutes of Health (NIH) grant AI4565801 (K.A.M.) and the Potts Memorial Foundation (K.A.M.). L.A.M. was supported by NIH grant HG01257 (to C. E. Lawrence) and Department of Energy grant 96ER62266 (to C. E. Lawrence).

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Ausubel, F., R. Brent, R. Kingston, D. Moore, J. Seidman, J. Smith, and K. Struhl (ed.).** 1992. Analysis of proteins, vol. 2. Greene Publishing Associates and Wiley Interscience, New York, N.Y.
3. **Bateman, A., E. Birney, R. Durbin, S. R. Eddy, K. L. Howe, and E. L. Sonnhammer.** 2000. The Pfam protein families database. Nucleic Acids Res. **28:**263–266.
4. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. G. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M.-A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead, and B. G. Barrell.** 1998. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature **393:**537–544.
5. **Cunningham, A. F., and C. L. Spreadbury.** 1998. Mycobacterial stationary phase induced by low oxygen tension: cell wall thickening and localization of the 16-kilodalton α-crystallin homolog. J. Bacteriol. **180:**801–808.
6. **Dick, T., B. H. Lee, and B. Murugasu-Oei.** 1998. Oxygen depletion induced dormancy in *Mycobacterium smegmatis*. FEMS Microbiol. Lett. **163:**159–164.
7. **Fenton, M. J., and M. W. Vermeulen.** 1996. Immunopathology of tuberculosis: roles of macrophages and monocytes. Infect. Immun. **64:**683–690.
8. **Freestone, P., T. Nystrom, M. Trinei, and V. Norris.** 1997. The universal stress protein, UspA, of *Escherichia coli* is phosphorylated in response to stasis. J. Mol. Biol. **274:**318–324.
9. **Gliubich, F., M. Gazerro, G. Zanotti, S. Delbono, G. Bombieri, and R. Berni.** 1996. Active site structural features for chemically modified forms of rhodanese. J. Biol. Chem. **271:**21054–21061.
10. **Höner Zu Bentrup, K., A. Miczak, D. L. Swenson, and D. G. Russell.** 1999. Characterization of activity and expression of isocitrate lyase in *Mycobacterium avium* and *Mycobacterium tuberculosis*. J. Bacteriol. **181:**7161–7167.
11. **Hu, Y., and A. R. Coates.** 1999. Transcription of the stationary-phase-associated *hspX* gene of *Mycobacterium tuberculosis* is inversely related to synthesis of the 16-kilodalton protein. J. Bacteriol. **181:**1380–1387.
12. **Jensen, O. N., P. Mortensen, O. Vorm, and M. Mann.** 1997. Automation of matrix-assisted laser desorption/ionization mass spectrometry using fuzzy logic feedback control. Anal. Chem. **69:**1706–1714.
13. **Lee, B. Y., and M. A. Horwitz.** 1995. Identification of macrophage and stress-induced proteins of *Mycobacterium tuberculosis*. J. Clin. Investig. **96:**245–249.
14. **Li, Z., C. Kelley, F. Collins, D. Rouse, and S. Morris.** 1998. Expression of *katG* in *Mycobacterium tuberculosis* is associated with its growth and persistence in mice and guinea pigs. J. Infect. Dis. **177:**1030–1035.
15. **Link, A. J., J. Eng, D. M. Schieltz, E. Carmack, G. J. Mize, D. R. Morris, B. M. Garvik, and J. R. Yates III.** 1999. Direct analysis of protein complexes using mass spectrometry. Nat. Biotechnol. **17:**676–682.
16. **Liu, J. S., and C. E. Lawrence.** 1999. Bayesian inference on biopolymer models. Bioinformatics **15:**38–52.
17. **McDonough, K. A., M. A. Florczyk, and Y. Kress.** 2000. Intracellular passage within macrophages ;ffects the trafficking of virulent tubercle bacilli upon reinfection of other macrophages in a serum-dependent manner. Tuber. Lung Dis. **80:**259–271.
18. **McDonough, K. A., Y. Kress, and B. R. Bloom.** 1993. The interaction of *Mycobacterium tuberculosis* with macrophages: a study of phagolysosome fusion. Infect. Agents Dis. **2:**232–235.
19. **Monahan, I., J. Betts, D. Banerjee, and P. Butcher.** 2001. Differential expression of mycobacterial proteins following phagocytosis by macrophages. Microbiology **147:**459–471.
20. **Moran, L., K. Scrimgeour, H. Horton, R. Ochs, and J. Rawn.** 1994. Biochemistry, 2nd ed. Neil Patterson Publishers, Prentice Hall, Englewood Cliffs, N.J.
21. **Mushegian, A. R., and E. V. Koonin.** 1996. Sequence analysis of eukaryotic developmental proteins: ancient and novel domains. Genetics **144:**817–828.
22. **Musser, J. M.** 1995. Antimicrobial agent resistance in mycobacteria: molecular genetic insights. Clin. Microbiol. Rev. **8:**496–514.
23. **Neuwald, A. F., J. S. Liu, D. J. Lipman, and C. E. Lawrence.** 1997. Extracting protein alignment models from the sequence database. Nucleic Acids Res. **25:**1665–1677.
24. **Nystrom, T., and F. C. Neidhardt.** 1992. Cloning, mapping and nucleotide sequencing of a gene encoding a universal stress protein in *Escherichia coli*. Mol. Microbiol. **6:**3187–3198.
25. **Nystrom, T., and F. C. Neidhardt.** 1996. Effects of overproducing the universal stress protein, UspA, in *Escherichia coli* K-12. J. Bacteriol. **178:**927–930.
26. **Ogata, H., S. Goto, K. Sato, W. Fujibuchi, H. Bono, and M. Kanehisa.** 1999. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. **27:**29–34.
27. **Ploegman, J. H., G. Drent, K. H. Kalk, W. G. Hol, R. L. Heinrikson, P. Keim, L. Weng, and J. Russell.** 1978. The covalent and tertiary structure of bovine liver rhodanese. Nature **273:**124–129.
28. **Raviglione, M. C., D. E. Snider, Jr., and A. Kochi.** 1995. Global epidemiology of tuberculosis. Morbidity and mortality of a worldwide epidemic. JAMA **273:**220–226.
29. **Rouse, D. A., J. A. DeVito, Z. Li, H. Byer, and S. L. Morris.** 1996. Site-directed mutagenesis of the *katG* gene of *Mycobacterium tuberculosis*: effects on catalase-peroxidase activities and isoniazid resistance. Mol. Microbiol. **22:**583–592.
30. **Schneider, T. D., and R. M. Stephens.** 1990. Sequence logos: a new way to display consensus sequences. Nucleic Acids Res. **18:**6097–6100.
31. **Sherman, D. R., K. Mdluli, M. J. Hickey, T. M. Arain, S. L. Morris, C. E. Barry III, and C. K. Stover.** 1996. Compensatory *ahpC* gene expression in isoniazid-resistant *Mycobacterium tuberculosis*. Science **272:**1641–1643.
32. **Sturgill-Koszycki, S., P. L. Haddix, and D. G. Russell.** 1997. The interaction between Mycobacterium and the macrophage analyzed by two-dimensional polyacrylamide gel electrophoresis. Electrophoresis **18:**2558–2565.
33. **Unden, G., S. Becker, J. Bongaerts, G. Holighaus, J. Schirawski, and S. Six.** 1995. O2-sensing and O2-dependent gene regulation in facultatively anaerobic bacteria. Arch. Microbiol. **164:**81–90.
34. **Wayne, L. G., and G. A. Diaz.** 1967. Autolysis and secondary growth of *Mycobacterium tuberculosis* in submerged culture. J. Bacteriol. **93:**1374–1381.
35. **Wayne, L. G., and L. G. Hayes.** 1996. An in vitro model for sequential study of shiftdown of *Mycobacterium tuberculosis* through two stages of nonreplicating persistence. Infect. Immun. **64:**2062–2069.
36. **Wayne, L. G., and K. Y. Lin.** 1982. Glyoxylate metabolism and adaptation of *Mycobacterium tuberculosis* to survival under anaerobic conditions. Infect. Immun. **37:**1042–1049.
37. **Wayne, L. G., and H. A. Sramek.** 1994. Metronidazole is bactericidal to dormant cells of *Mycobacterium tuberculosis*. Antimicrob. Agents Chemother. **38:**2054–2058.
38. **Weldingh, K., I. Rosenkrands, S. Jacobsen, P. B. Rasmussen, M. J. Elhay, and P. Andersen.** 1998. Two-dimensional electrophoresis for analysis of *Mycobacterium tuberculosis* culture filtrate and purification and characterization of six novel proteins. Infect. Immun. **66:**3492–3500.
39. **Wilson, T. M., G. W. de Lisle, and D. M. Collins.** 1995. Effect of *inhA* and *katG* on isoniazid resistance and virulence of *Mycobacterium bovis*. Mol. Microbiol. **15:**1009–1015.
40. **Young, D. B., and T. R. Garbe.** 1991. Heat shock proteins and antigens of *Mycobacterium tuberculosis*. Infect. Immun. **59:**3086–3093.
41. **Yuan, Y., D. D. Crane, R. M. Simpson, Y. Q. Zhu, M. J. Hickey, D. R. Sherman, and C. E. Barry III.** 1998. The 16-kDa alpha-crystallin (Acr) protein of *Mycobacterium tuberculosis* is required for growth in macrophages. Proc. Natl. Acad. Sci. USA **95:**9578–9583.
42. **Zarembinski, T. I., L. W. Hung, H. J. Mueller-Dieckmann, K. K. Kim, H. Yokota, R. Kim, and S. H. Kim.** 1998. Structure-based assignment of the biochemical function of a hypothetical protein: a test case of structural genomics. Proc. Natl. Acad. Sci. USA **95:**15189–15193.

Editor: B. B. Finlay