

NOTES

Identification of Secreted Proteins of *Mycobacterium tuberculosis* by a Bioinformatic Approach

MANUEL GOMEZ, SADIE JOHNSON, AND MARIA LAURA GENNARO*

Public Health Research Institute, New York, New York

Received 3 September 1999/Returned for modification 18 October 1999/Accepted 13 December 1999

Proteins secreted by *Mycobacterium tuberculosis* are usually targets of immune responses in the infected host. Here we describe a search for secreted proteins that combined the use of bioinformatics and *phoA'* fusion technology. The 3,924 proteins deduced from the *M. tuberculosis* genome were analyzed with several computer programs. We identified 52 proteins carrying an NH₂-terminal secretory signal peptide but lacking additional membrane-anchoring moieties. Of these 52 proteins—the TM1 subgroup—only 7 had been previously reported to be secreted proteins. Our predictions were confirmed in 9 of 10 TM1 genes that were fused to *Escherichia coli phoA'*, a marker of subcellular localization. These findings demonstrate that the systematic computer search described in this work identified secreted proteins of *M. tuberculosis* with high efficiency and 90% accuracy.

Proteins released by *Mycobacterium tuberculosis* to the extracellular environment have been the focus of much of the research directed at identifying antigens that induce protective immunity or those that elicit immune responses of diagnostic value (reviewed in references 2, 11, and 37). Two different experimental approaches have been used. One is analysis of the protein composition of *M. tuberculosis* culture filtrates. The *M. tuberculosis* culture filtrate, which contains as many as 200 proteins (29), has been investigated by means of protein purification, by immunological methods, and by screening of expression libraries of *M. tuberculosis* DNA with anti-culture filtrate sera (for examples, see references 1, 4, 5, 19, 25, 33, and 34). A second, genetic approach involves the screening of libraries of fusions of *M. tuberculosis* genes to reporter genes encoding enzymes that become active upon translocation across the cell membrane (7, 18). As a result of the combined efforts of several laboratories, more than 30 secreted proteins of *M. tuberculosis* have been characterized (examples are provided in references 3, 6, 9, 10, 12, 15, 17, 18, 20, 22, 26, 30, 32, and 34). Nevertheless, much of the immunological activity of the culture filtrate of *M. tuberculosis* remains unaccounted for.

Analysis of the NH₂-terminal sequences of proteins purified from the culture filtrate vis-à-vis corresponding deduced amino acid sequences (34, 37) indicates that many proteins of *M. tuberculosis* are secreted via the general export pathway (GEP). The GEP mediates protein translocation across the cytoplasmic membrane by means of an NH₂-terminal secretory signal peptide (24, 28). Following translocation, cleavage of the signal peptide by a signal peptidase releases the mature protein, providing that there are no additional membrane-spanning segments or membrane-anchoring moieties (Fig. 1). The conserved features of signal peptides (length and amino acid composition plus the signal peptidase cleavage site [31]) make them amenable to identification by sequence analysis. Thus, we

set out to identify proteins of *M. tuberculosis* secreted via the GEP by using a bioinformatic approach that takes advantage of the recently released nucleotide sequence of the *M. tuberculosis* genome (8).

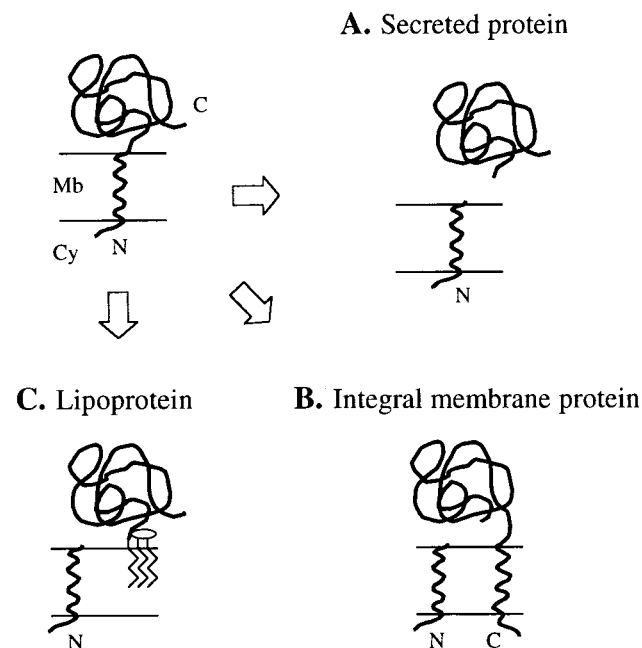


FIG. 1. Fates of proteins having an NH₂-terminal signal peptide. The NH₂-terminal signal peptide directs translocation of a protein across the cell membrane (top left panel). Processing of the preprotein by a type I signal peptidase releases the mature protein (A), providing that there are no additional transmembrane domains (B). Translocation of LPPs across the membrane is followed by modification with glycerol and fatty acids of the lipoprotein and processing by a type II signal peptidase. The lipid moiety anchors the LPP to the membrane (C). Mb, cell membrane; Cy, cytoplasm; N, NH₂-terminal end; C, COOH-terminal end.

* Corresponding author. Mailing address: Public Health Research Institute, 455 First Ave., New York, NY 10016. Phone: (212) 578-0844. Fax: (212) 578-0804. E-mail: gennaro@phri.nyu.edu.

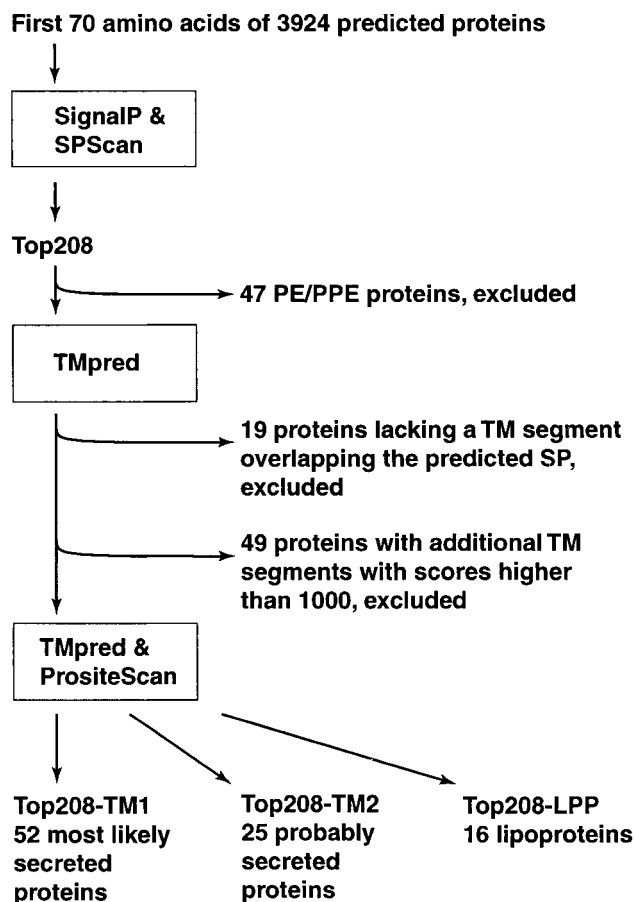


FIG. 2. Schematic representation of the strategy utilized to predict *M. tuberculosis* secreted proteins. TM, transmembrane; SP, signal peptide; PE/PPE, families of *M. tuberculosis* proteins containing multiple copies of sequences rich in small-side-chain amino acid residues (8).

Prediction of secreted proteins of *M. tuberculosis*. The computer strategy used to identify proteins having secretory signal peptides but lacking additional membrane attachment domains is presented in Fig. 2. The amino acid sequences of 3,924 proteins deduced from the nucleotide sequence of the *M. tuberculosis* genome were downloaded from the Sanger Center database (http://www.sanger.ac.uk/Projects/M_tuberculosis). Segments containing the NH₂-terminal 70 amino acid residues of each polypeptide were analyzed for secretory signal peptides with two computer programs, SignalP (<http://www.cbs.dtu.dk/services/SignalP>) and SPScan (13). The two programs assigned scores to potential signal peptides for all of the proteins except two (Rv1572c and Rv3599c) that were too short for SPScan analysis. Scores ranged from 0.039 to 0.888 for SignalP predictions and from -3.8 to +14.8 for SPScan predictions (Fig. 3). Cutoff values for the computer predictions were chosen on the basis of scores assigned to nine known secreted proteins of *M. tuberculosis* that contain a signal peptide (Ag85A, Ag85B, Ag85C, MPT32, MPT51, MPT53, MPT63, MPT64, and MPB70) (34). For these proteins, SignalP scores ranged from 0.425 to 0.758 and SPScan scores ranged from 8 to 11.2. We chose score cutoffs of 0.4 for SignalP and 8.0 for SPScan. Both values are more restrictive than the default cutoff values (0.34 for SignalP and 3.5 for SPScan). Two hundred eight proteins (5% of the proteome) scored above the cutoff with both programs (boxed inset in Fig. 3). This set of proteins is henceforth referred to as the Top208 group.

Of the proteins in the Top208 group, 47 had been previously annotated as members of the PE and PPE families (8). Proteins in these families contain multiple copies of motifs rich in small-side-chain amino acid residues, such as alanine and glycine. Since a biased amino acid composition and a repetitive primary sequence could lead to unreliable results in sequence analyses, the 47 PE and PPE proteins in the Top208 group were not further analyzed.

Proteins that cross the membrane via the GEP are released to the external environment after cleavage of the signal peptide only if they lack membrane-anchoring sequences (Fig. 1). To eliminate from our study those proteins that contain membrane-spanning segments, the remaining 161 proteins in the Top208 group were analyzed with the program TMpred (http://www.ch.embnet.org/software/TMPRED_form.html) (Fig. 2). The presence of an NH₂-terminal transmembrane segment (i.e., the putative signal peptide) was confirmed by TMpred only for 142 proteins. Of these, 49 were putative integral membrane proteins, for they contained transmembrane segments within the mature protein that scored >1,000 (the default cutoff score for TMpred is 500) (Fig. 2). The remaining 93 proteins were analyzed for membrane lipoprotein (LPP) lipid attachment sites with the program PrositeScan (http://www.isrec.isb-sib.ch/software/PSTSCAN_form.html) (LPP motif, PROSITE database entry PS00013). We defined three subgroups based on TMpred and PrositeScan predictions (Fig. 2). The first subgroup, Top208-TM1, comprised 52 proteins that were

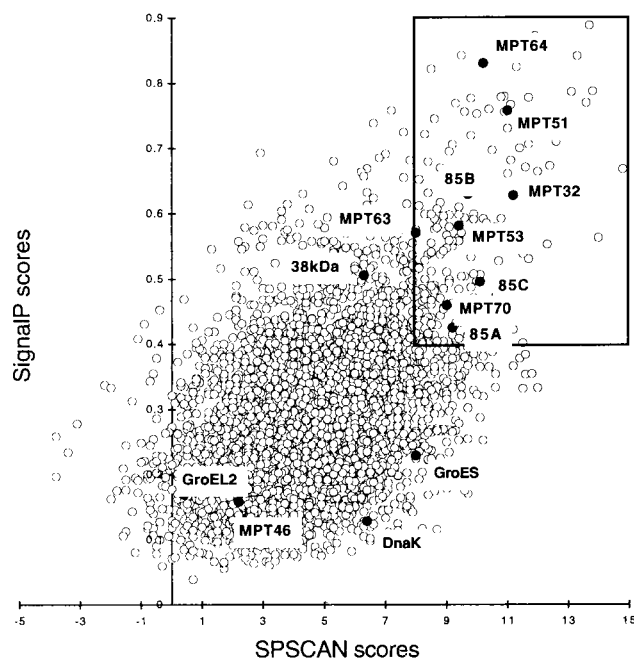


FIG. 3. Correlation between SignalP and SPScan scores. Each circle represents one of the 3,924 deduced proteins of *M. tuberculosis*. The position of the circle indicates the scores assigned by SignalP and SPScan to the signal peptide. For this study, the two computer programs were set in the mode for gram-positive bacteria. SPScan was used in the adjusted mode, which penalizes signal peptides longer than 45 amino acid residues (i.e., the maximal length for gram-positive bacteria). The boxed area contains the Top208 group of proteins that scored ≥ 0.4 with SignalP and ≥ 8 with SPScan. The solid circles within the boxed area represent nine known secreted proteins (Ag85A, Ag85B, Ag85C, MPT32, MPT51, MPT53, MPT63, MPT64, and MPB70) (34) that were used to set cutoff values for SignalP and SPScan scores. The solid circles outside of the boxed area represent four known proteins (MPT46, GroES, GroEL2, and DnaK) containing no signal peptide (34). Also outside of the boxed area is located the 38-kDa antigen, a secreted protein which exhibits an LPP type of signal peptide (14).

TABLE 1. Top208-Tm1 proteins^a

Gene name	Set	SPScan score	SignalP score	SPScan-predicted signal peptide sequence	Relevant notes regarding function or identity
rv0315*	A	14.8	0.667	MLMPEMDRRRMMAGFGALAAALPAPTAWA^DP	Glucanase (SCDB)
rv3668c*	A	13.8	0.787	LQTAHRRFAAAFAAVLLAVVCLPANTAAA^DD	Possible protease (SCDB)
rv0477	B	13.7	0.888	MKALVAVSAVAVVALLGVSSAQA^DP	
rv0398c	A	12.6	0.71	MGVIARVVVGAACGLSLAVLAAAAPTAGA^EP	
rv3096	A	12.4	0.621	VHRRRTALKLPLLLAAGTVLGGQAPRAAA^EE	Glycosylhydrolase family 5 signature (PS00659)
rv3354*	A	12.4	0.672	MNLRRHQTLTLRLLAASAGILSAAFAAPAQA^NP	
rv1291c	B	12	0.663	MFTRRFAASMVGTTLTAATLGLAALGFAGTASA^SS	
rv0559c	B	11.9	0.873	MKGTKLAVVVGMTVAAVSLAAPAQA^DD	
rv1891	A	11.7	0.53	MIRELVTTAAITGAAIGGAPVAGA^DP	
rv3333c	A	11.5	0.538	MFTGIASHAGALGAALVVLIGAAIILHDGPA^DP	
rv0040c	C	11.4	0.7	MIQIARTWRVVFAGGMATGFIGVVLVTAGKASA^DP	MTC28 (19)
rv1269c*	A	11.4	0.73	MTTMITLRRRFVAVAVGATAAAATVTLAPAPANA^AD	
rv2253	B	11.3	0.824	MSGHRKKAMLALAAASLAAATLAPNAVAAA^AE	
rv1860	C	11.2	0.627	MHQVDPNLTTRKGRLAALAAAMASASLVTVAVPATANA^DP	MPT32 (17)
rv0455c	B	11	0.661	MSRSLSSILRGAFAFLVLGIAAATFPQSAAA^DS	
rv1268c*	A	10.9	0.779	MTTSKIATAFKTATFALAAGAVALGLASPADA^AA	Cys proteases histidine active site (PS00639)
rv1174c	C	10.6	0.579	MRLSLTALSAGVAVAMSLTVGAGVASA^DP	Mtb8.4SA-5K (9, 12)
rv2376c	C	10.5	0.438	MAGGPVVYQMPVVFAGPLPLDPASA^PD	MTB12 (32)
rv2450c*	B	10.5	0.697	LKNARTTLIAAAIAGTLVTTSPAGIANA^DD	Putative pheromone (23)
rv1271c	A	10.2	0.592	MLSPLSPRIIAAFTTAVGAAAIGLAVATAGTAGA^NT	
rv1980c	C	10.2	0.83	VRIKIFMLVTAIVVLLCCSGVATA^AP	MPT64 (36)
rv0617	A	10	0.458	VTVLLDANVLIALVVAEHHV^DA	
rv0674	A	10	0.635	MPAMTARSVVLVLLGAHPAWA^TA	
rv1906c*	A	10	0.753	MRLKPAPSPAAAFAVAGLILAGWAGSVGLAGA^DP	
rv1006	A	9.9	0.432	MVLRSRKSTLGVVCLALVLLGGPLNGCSSSA^SH	
rv2389c	A	9.5	0.842	MFVALLGLSTISSKA^DD	Putative pheromone (23)
rv2878c	C	9.4	0.581	MSLRLVSPIKAFADGIVAVAIIVLMFGLANTPRAVA^AD	MPT53, DsbE (34)
rv3036c	B	9.4	0.598	MRYLIATAVLLVAVVVLVGVWPAAGA^PP	Similar to MPT64 (SCDB)
rv1566c*	B	9.2	0.55	MKRSMKSGSFAIGLAMMLAPMVAAPGLAAA^DP	Similar to <i>Listeria</i> invasion-associated p60 (SCDB)
rv1974	A	9.2	0.557	VQRQSLMPQQTLAGVVFVGLLCCGVVTA^AV	
rv1419	A	9.1	0.695	MGELRLVGGVLRVLLVVGVAVFDVAVLNAGAASA^DG	
rv3106	A	9.1	0.494	MRPYIIAIVGSGPSAFFA^AA	FprA, NADPH ferredoxin reductase? (SCDB)
rv1813c	A	8.9	0.584	MITNLRRTTAMAAAGLGAALGLGILLVPTVDAHL^AN	
rv3013	A	8.9	0.543	VRSYLLRIELADRPGLSGLSVALGSGVGA^DI	
rv0199	A	8.8	0.537	MFSTYGIASTLLGVLSVAAVVVGAMIWSAHR^DD	
rv0867c	B	8.8	0.616	MSGRRKPTTNSVSVAKIAFTGAVLGGGGIAMA^AQ	Putative pheromone (23)
rv2576c	A	8.8	0.669	MTSVRTVPSAVALVTFAGAALSGVIPAIAARA^DP	
rv3170	A	8.8	0.444	VTNPPWTVDVVVVVGAGFAGLAAA^RE	Probable monoamine oxidase (SCDB)
rv3572	A	8.7	0.578	MTRLIPGCTLVGLMLTLLPAPTSAA^GS	
rv0592	B	8.6	0.413	MSTIFDIRSLRPLKLSAKVVVVGGLVVVLAVVAAA^AG	Part of <i>mce-2</i> operon (SCDB)
rv1352	A	8.6	0.592	MARTLALRASAGLVAGMAMAAITLAPGARA^ET	
rv1242	A	8.5	0.555	VIIIPDINLLLYAVITGFPQHRRHAH^WW	
rv0309	A	8.4	0.537	MSRLLALLCAAVCTGCVAVVLPVSLAVVNPWFA^NS	Heavy-metal-associated domain (PS01047)
rv0360c	A	8.4	0.412	VTKRITIPMTSMGDLLEPEPILLPGDSDAEALLANESPSIVAA^AH	
rv1804c	A	8.4	0.631	MRVVSTLLSIPLMIGLAVPAHA^GP	
rv2223c*	B	8.4	0.679	MAAMWRRRPLSSALLSFLGGLLPLAAPPPLAGA^TE	Lipase, serine active site (SCDB)
rv1488	A	8.3	0.721	VQGAVAGLVFLAVLVIFAIIVVAKSVALIPQAEA^AV	Hemopexin domain signature (PS00024)
rv2290*	B	8.3	0.424	LTDPRHTVRIAVGATALGVSAALGATLPACSAHS^GP	LppO, similar to 19-kDa antigen (SCDB)
rv3207c	A	8.3	0.469	VSTYGWYALPVLMLVTVVVVYQTVTGTSTPRPAAA^QT	Zn protease (SCDB)
rv0203	B	8.2	0.485	MKTGTATRRRLAVLIALALPGAVALLAEPSATG^AS	
rv0603	A	8.2	0.466	MNRIVQFGVSAVAAAIGIGAGSGIAAA^FD	
rv1926c*	C	8	0.57	MKLTMMIKTAVAVVAMAAIATFAAPVALA^AY	MPT63 (20)

^a The 52 proteins of *M. tuberculosis* in the Top208-TM1 subgroup are sorted by decreasing SPScan scores. Genes analyzed by *phoA'* fusion technology are marked with asterisks. Set A includes 32 proteins that had no annotations in the Sanger Center database (SCDB) referring to protein topology or subcellular localization, set B includes 13 proteins that were annotated as "probably secreted, or exported, or having a putative signal sequence, or a hydrophobic amino acid stretch at the NH₂-terminus," and set C includes seven proteins known to be secreted by a signal peptide-dependent mechanism (see text). The position of the predicted cleavage site is indicated by a caret. Relevant notes regarding function or identity are based on previously published reports, Sanger Center database annotations, or Blast and PrositeScan searches.

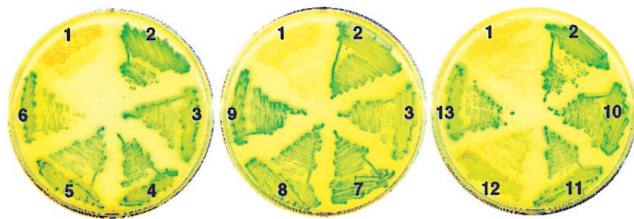


FIG. 4. Alkaline phosphatase activity of 10 *M. tuberculosis* TM1 proteins fused to PhoA. *phoA'* fusions were constructed in *E. coli* plasmid pUCCMPHOA (16), a pUC18 derivative that carries *phoA'* downstream of codon 18 of *lacZ*. *phoA'* lacks the nucleotide sequences coding for the NH₂-terminal signal peptide and the first 11 amino acid residues of the mature PhoA protein. Hybrid gene expression in pUCCMPHOA utilizes the *lac* promoter and the *lacZ* ribosome binding site and start codon. Full-length *M. tuberculosis* coding sequences (*Mtcds*) were synthesized from chromosomal DNA by PCR, cloned into plasmid pCR2.1 (Invitrogen), and subsequently transferred into pUCCMPHOA using appropriate restriction endonuclease sites to obtain in-frame *lacZ-Mtcds-phoA'* fusions. Recombinant plasmids were characterized by nucleotide sequence analysis of the DNA junctions. *E. coli* clones bearing vector pUCCMPHOA and 12 different *phoA* fusions were tested for alkaline phosphatase activity on Luria-Bertani agar plates containing BCIP at 40 $\mu\text{g ml}^{-1}$. Plate sectors: 1, pUCCMPHOA (negative control); 2, *B. subtilis* membrane protein Comp (positive control) (27); 3, secreted protein MPT63 of *M. tuberculosis* (positive control—also one of the TM1 proteins) (20); 4, Rv0315; 5, Rv1268c; 6, Rv1269c; 7, Rv1566c; 8, Rv1906c; 9, Rv2223c; 10, Rv2290; 11, Rv2450c; 12, Rv3354; 13, Rv3668c.

classified as “most likely secreted.” These proteins contained one transmembrane domain corresponding to the signal peptide, no additional membrane-spanning segments having TMpred scores above 500, and no LPP motifs. The second subgroup, Top208-TM2, consisted of 25 proteins that were classified as “possibly secreted” because the mature proteins contained segments having TMpred scores between 500 and 1,000. The remaining 16 proteins (the Top208-LPP subgroup) were predicted to be LPPs, for they contained a properly positioned LPP motif.

The 52 proteins comprising the Top208-TM1 subgroup are listed in Table 1. Thirty-two of the TM1 proteins were previously unrecognized as secreted proteins (set A in Table 1, column 2), while 13 TM1 proteins had annotations in the Sanger Center database that suggest the presence of NH₂-terminal signal peptides (set B in Table 1, column 2). Only seven TM1 proteins were previously known as secreted proteins (Mtb8.4, MTB12, MTC28, MPT32, MPT53, MPT63, and MPT64) (9, 12, 17, 19, 20, 32, 34, 36) (set C in Table 1, column 2). Sixteen of the 52 TM1 proteins were assigned a probable function on the basis of the presence of functional motifs or similarities to proteins with known functions (Table 1, column 6).

Analysis of TM1 proteins by *phoA'* fusion technology. A method to study protein topology with respect to the cell membrane is to analyze the enzymatic activity of a hybrid consisting of the test protein and the reporter *Escherichia coli* PhoA protein devoid of its original signal peptide (21). To test the prediction that the proteins in the TM1 subgroup contained a single transmembrane domain, the signal peptide (Fig. 1), fusions were constructed between *phoA'* and full-length *M. tuberculosis* TM1 coding sequences. To introduce a minimum of bias in these experiments, we chose 10 proteins (marked with asterisks in Table 1, column 1) represented a broad range of SPScan scores (8.3 to 14.8) and SignalP scores (0.469 to 0.787) and were equally divided between those having (set B) and those not having (set A) database annotations with reference to topology or localization.

Recombinant plasmids bearing fusions of *phoA'* with *M.*

tuberculosis TM1 genes were constructed in *E. coli*. Clones carrying the fusions were evaluated for alkaline phosphatase activity by plating on Luria-Bertani agar containing the chromogenic substrate 5-bromo-4-chloro-3-indolylphosphate (BCIP). On indicator plates, alkaline phosphatase activity is seen as blue-green colonies. The colony color phenotype of one representative clone for each fusion is shown in Fig. 4. Nine of 10 fusions exhibited alkaline phosphatase activity. The only *phoA'* fusion that yielded a weak, albeit detectable, blue-green colony color was that with Rv3354 (clone 12 in Fig. 4). All of the hybrid proteins, except that encoded by the Rv3354:*phoA'* fusion, were detected by Western blot analysis of whole-cell lysates with anti-PhoA antibodies (data not shown). These results suggest that the hybrid protein encoded by the Rv3354:*phoA'* fusion is either synthesized at low levels or rapidly degraded. The finding that 9 of 10 mycobacterial protein hybrids directed export of the PhoA moiety provides strong evidence that proteins in the Top208-TM1 subgroup are bona fide secreted proteins.

In summary, 52 proteins (the Top208-TM1 subgroup) were identified by computer-based analyses as most likely secreted. The computer predictions were confirmed in 90% of the TM1 proteins tested by *E. coli phoA'* gene fusion methods. Only 7 of the 52 TM1 proteins had been previously reported to be secreted proteins. Thus, the bioinformatic method used in the present paper is highly efficient and accurate. The high accuracy is explained by the use of stringent selection criteria (i.e., high cutoff scores for SignalP, SPScan, and TMPred).

The computer-based approach that we have used to identify secreted proteins differs in several important ways from genetic screening methods employing random fusions to *phoA'* or *bla*, markers of extracellular localization. First, methods that employ random fusions do not allow distinction between secretory signal peptides and transmembrane domains of membrane-bound proteins (reference 35 and references therein). Second, expression of alkaline phosphatase activity by clones generated by random insertion of *phoA*-coupled transposons critically depends on the strength of promoters located upstream of the transposition site (21). The same limitation applies to the screening of random fusion libraries in promoterless plasmids. Third, gene fusion library screening methods often yield multiple hits in the same genes, thus making the process of gene identification labor intensive (7). A disadvantage of the computer-based approach is that it is limited to only those proteins secreted via the GEP. While many secreted antigens of *M. tuberculosis* fall into this category, the presence in culture filtrates of proteins lacking secretory signal peptides (e.g., ESAT-6, SodA, GlnA, and KatG [15, 29, 30, 38]) suggests the existence of other, still undefined, mechanisms of protein secretion that operate in *M. tuberculosis*.

The identification of novel secreted proteins of *M. tuberculosis* opens the way to studies on their subcellular localization in *M. tuberculosis* and to the immunological characterization of these proteins to define their potential for immunological diagnosis of tuberculosis or vaccine design.

We thank David Dubnau for providing plasmid pUCCMPHOA and Carol Lusty and Karl Drlica for critical reading of the manuscript.

This work was supported by NIH grant AI-36989 (M.L.G.).

REFERENCES

1. Amara, R. R., and V. Satchidanandam. 1996. Analysis of a genomic DNA expression library of *Mycobacterium tuberculosis* using tuberculosis patient sera: evidence for modulation of host immune response. *Infect. Immun.* 64:3765–3771.
2. Andersen, A. B., and P. Brennan. 1994. Proteins and antigens of *Mycobacterium tuberculosis*, p. 307–327. *In* B. R. Bloom (ed.), *Tuberculosis: patho-*

- genesis, protection, and control. American Society for Microbiology, Washington, D.C.
3. 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.
 4. Andersen, P., D. Askgaard, A. Gottschau, J. Bennedsen, S. Nagai, and I. Heron. 1992. Identification of immunodominant antigens during infection with *Mycobacterium tuberculosis*. Scand. J. Immunol. **36**:823–831.
 5. Andersen, P., D. Askgaard, L. Ljungqvist, J. Bennedsen, and I. Heron. 1991. Proteins released from *Mycobacterium tuberculosis* during growth. Infect. Immun. **59**:1905–1910.
 6. Ashbridge, K. R., R. J. Booth, J. D. Watson, and R. B. Lathigra. 1989. Nucleotide sequence of the 19 kDa antigen gene from *Mycobacterium tuberculosis*. Nucleic Acids Res. **17**:1249.
 7. Chubb, A. J., Z. L. Woodman, F. M. da Silva Tatley, H. J. Hoffmann, R. R. Scholle, and M. R. Ehlers. 1998. Identification of *Mycobacterium tuberculosis* signal sequences that direct the export of a leaderless beta-lactamase gene product in *Escherichia coli*. Microbiology **144**:1619–1629.
 8. Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry 3rd, F. Tekaiia, 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, B. G. Barrell, et al. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393**:537–544.
 9. Coler, R. N., Y. A. Skeiky, T. Vedvick, T. Bement, P. Owendale, A. Campos-Neto, M. R. Alderson, and S. G. Reed. 1998. Molecular cloning and immunologic reactivity of a novel low molecular mass antigen of *Mycobacterium tuberculosis*. J. Immunol. **161**:2356–2364.
 10. Content, J., A. de la Cuvelierie, L. De Wit, V. Vincent-Levy-Frebault, J. Ooms, and J. De Bruyn. 1991. The genes coding for the antigen 85 complexes of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG are members of a gene family: cloning, sequence determination, and genomic organization of the gene coding for antigen 85-C of *M. tuberculosis*. Infect. Immun. **59**:3205–3212.
 11. Cooper, A. M., and J. L. Flynn. 1995. The protective immune response to *Mycobacterium tuberculosis*. Curr. Opin. Immunol. **7**:512–516.
 12. Freer, G., W. Florio, B. Dalla Casa, D. Bottai, G. Batoni, G. Maisetta, S. Senesi, and M. Campa. 1998. Identification and molecular cloning of a novel secretion antigen from *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG. Res. Microbiol. **149**:265–275.
 13. Genetics Computer Group. 1997. Wisconsin Package Version 9.1. Genetics Computer Group, Madison, Wis.
 14. Harboe, M., and H. G. Wiker. 1992. The 38-kDa protein of *Mycobacterium tuberculosis*: a review. J. Infect. Dis. **166**:874–884.
 15. Harth, G., D. L. Clemens, and M. A. Horwitz. 1994. Glutamine synthetase of *Mycobacterium tuberculosis*: extracellular release and characterization of its enzymatic activity. Proc. Natl. Acad. Sci. USA **91**:9342–9346.
 16. Inamine, G. S., and D. Dubnau. 1995. ComEA, a *Bacillus subtilis* integral membrane protein required for genetic transformation, is needed for both DNA binding and transport. J. Bacteriol. **177**:3045–3051.
 17. Laqueyrie, A., P. Miltzer, F. Romain, K. Eiglmeier, S. Cole, and G. Marchal. 1995. Cloning, sequencing, and expression of the *apa* gene coding for the *Mycobacterium tuberculosis* 45/47-kilodalton secreted antigen complex. Infect. Immun. **63**:4003–4010.
 18. Lim, E. M., J. Rauzier, J. Timm, G. Torrea, A. Murray, B. Gicquel, and D. Portnoi. 1995. Identification of *Mycobacterium tuberculosis* DNA sequences encoding exported proteins by using *phoA* gene fusions. J. Bacteriol. **177**:59–65.
 19. Manca, C., K. Lyashchenko, R. Colangeli, and M. L. Gennaro. 1997. MTC28, a novel 28-kilodalton proline-rich secreted antigen specific for the *Mycobacterium tuberculosis* complex. Infect. Immun. **65**:4951–4957.
 20. Manca, C., K. Lyashchenko, H. G. Wiker, D. Usai, R. Colangeli, and M. L. Gennaro. 1997. Molecular cloning, purification, and serological characterization of MPT63, a novel antigen secreted by *Mycobacterium tuberculosis*. Infect. Immun. **65**:16–23.
 21. Manoil, C. 1991. Analysis of membrane protein topology using alkaline phosphatase and beta-galactosidase gene fusions. Methods Cell. Biol. **34**:61–75.
 22. Matsumoto, S., T. Matsuo, N. Ohara, H. Hotokezaka, M. Naito, J. Minami, and T. Yamada. 1995. Cloning and sequencing of a unique antigen MPT70 from *Mycobacterium tuberculosis* H₃₇Rv and expression in BCG using *E. coli*-mycobacteria shuttle vector. Scand. J. Immunol. **41**:281–287.
 23. Mukamolova, G. V., A. S. Kaprelyants, D. I. Young, M. Young, and D. B. Kell. 1998. A bacterial cytokine. Proc. Natl. Acad. Sci. USA **95**:8916–8921.
 24. Murphy, C. K., and J. Beckwith. 1987. Export of proteins to the cell envelope in *Escherichia coli*, p. 967–978. In F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umberger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 1. American Society for Microbiology, Washington, D.C.
 25. Nagai, S., H. G. Wiker, M. Harboe, and M. Kinomoto. 1991. Isolation and partial characterization of major protein antigens in the culture fluid of *Mycobacterium tuberculosis*. Infect. Immun. **59**:372–382.
 26. Ohara, N., H. Kitaura, H. Hotokezaka, T. Nishiyama, N. Wada, S. Matsumoto, T. Matsuo, M. Naito, and T. Yamada. 1995. Characterization of the gene encoding the MPB51, one of the major secreted protein antigens of *Mycobacterium bovis* BCG, and identification of the secreted protein closely related to the fibronectin binding 85 complex. Scand. J. Immunol. **41**:433–442.
 27. Piazza, F., P. Tortosa, and D. Dubnau. 1999. Mutational analysis and membrane topology of CompP, a quorum-sensing histidine kinase of *Bacillus subtilis* controlling competence development. J. Bacteriol. **181**:4540–4548.
 28. Pugsley, A. P. 1993. The complete general secretory pathway in gram-negative bacteria. Microbiol. Rev. **57**:50–108.
 29. Sonnenberg, M. G., and J. T. Belisle. 1997. Definition of *Mycobacterium tuberculosis* culture filtrate proteins by two-dimensional polyacrylamide gel electrophoresis, N-terminal amino acid sequencing, and electrospray mass spectrometry. Infect. Immun. **65**:4515–4524.
 30. Sorensen, A. L., S. Nagai, G. Houen, P. Andersen, and A. B. Andersen. 1995. Purification and characterization of a low-molecular-mass T-cell antigen secreted by *Mycobacterium tuberculosis*. Infect. Immun. **63**:1710–1717.
 31. Watson, M. E. E. 1984. Compilation of published signal sequences. Nucleic Acids Res. **12**:5145–5164.
 32. Webb, J. R., T. S. Vedvick, M. R. Alderson, J. A. Guderian, S. S. Jen, P. J. Owendale, S. M. Johnson, S. G. Reed, and Y. A. Skeiky. 1998. Molecular cloning, expression, and immunogenicity of MTB12, a novel low-molecular-weight antigen secreted by *Mycobacterium tuberculosis*. Infect. Immun. **66**:4208–4214.
 33. 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.
 34. Wiker, H. G., S. L. Michell, R. G. Hewinson, E. Spierings, S. Nagai, and M. Harboe. 1999. Cloning, expression and significance of MPT53 for identification of secreted proteins of *Mycobacterium tuberculosis*. Microb. Pathog. **26**:207–219.
 35. Worley, M. J., I. Stojiljkovic, and F. Heffron. 1998. The identification of exported proteins with gene fusions to invasins. Mol. Microbiol. **29**:1471–1480.
 36. Yamaguchi, R., K. Matsuo, A. Yamazaki, C. Abe, S. Nagai, K. Terasaka, and T. Yamada. 1989. Cloning and characterization of the gene for immunogenic protein MPB64 of *Mycobacterium bovis* BCG. Infect. Immun. **57**:283–288.
 37. Young, D. B., S. H. Kaufmann, P. W. Hermans, and J. E. Thole. 1992. Mycobacterial protein antigens: a compilation. Mol. Microbiol. **6**:133–145.
 38. Zhang, Y., R. Lathigra, T. Garbe, D. Catty, and D. Young. 1991. Genetic analysis of superoxide dismutase, the 23 kilodalton antigen of *Mycobacterium tuberculosis*. Mol. Microbiol. **5**:381–391.