The EMBO Journal vol.15 no.14 pp.3547–3554, 1996

Bacterial glycoproteins: a link between glycosylation and proteolytic cleavage of a 19 kDa antigen from *Mycobacterium tuberculosis*

J.L.Herrmann[1,2], P.O'Gaora, A.Gallagher, J.E.R.Thole and D.B.Young

Department of Medical Microbiology, Imperial College School of Medicine at St Mary's, Norfolk Place, London W2 1PG, UK

[1]Present address: Laboratoire de Microbiologie, Hôpital Saint Louis, Paris, France

[2]Corresponding author

Protein glycosylation has an important influence on a broad range of molecular interactions in eukaryotes, but is comparatively rare in bacteria. Several antigens from *Mycobacterium tuberculosis*, the causative agent of human tuberculosis, have been identified as glycoproteins on the basis of lectin binding, or by detailed structural analysis. By production of a set of alkaline phosphatase (PhoA) hybrid proteins in a mycobacterial expression system, the peptide region required for glycosylation of the 19 kDa lipoprotein antigen from *M.tuberculosis* was defined. Mutagenesis of two threonine clusters within this region abolished lectin binding by PhoA hybrids and by the 19 kDa protein itself. Substitution of the threonine residues also resulted in generation of a series of smaller forms of the protein as a result of proteolysis. In a working model to account for these observations, we propose that the role of glycosylation is to regulate cleavage of a proteolytically sensitive linker region close to the acylated N-terminus of the protein.

Keywords: bacteria/glycoproteins/mycobacteria/proteolysis/tuberculosis

Introduction

A wide variety of eukaryotic proteins are modified by post-translational addition of O-linked and N-linked carbohydrate residues, but glycosylation is much less frequent among bacterial proteins. Analysis of the role of protein glycosylation in bacterial systems may provide insights into the evolution of mammalian glycoproteins. Glycoproteins involved in maintaining the structure of some bacterial cell walls ('S-layer glycoproteins') have been characterized in detail, particularly among the archaebacteria, and there is evidence of novel tyrosine-linked glycans in addition to more conventional O-linked and N-linked sugars (Mescher and Strominger, 1976a,b; Messner *et al.*, 1992; Bock *et al.*, 1994). A second group of bacterial glycoproteins includes a range of secreted or cell wall-associated enzymes, often with a function related to hydrolysis of carbohydrate substrates (Langsford *et al.*, 1984; Gerwig *et al.*, 1989, 1992; Kluepfel *et al.*, 1990; Milhoc and Kluepfel, 1990; Ong *et al.*, 1994; Plummer *et al.*, 1995; Reinhold *et al.*, 1995). Proteins involved in interactions with host cell components represent a third group of bacterial glycoproteins, including pilin proteins from pathogenic neisseriae (Parge *et al.*, 1995; Stimson *et al.*, 1995) and *Pseudomonas aeruginosa* (Castric, 1995) and a platelet aggregation-associated protein from *Streptococcus sanguis* (Erickson and Herzberg, 1993).

Three prominent antigens from *Mycobacterium tuberculosis* have been described as glycoproteins. In the case of a 45 kDa secreted antigen, detailed structural analysis has demonstrated covalent attachment of mannose to a threonine residue close to the N-terminus of the mature protein, providing definitive evidence for the existence of O-linked glycoproteins in mycobacteria (Dobos *et al.*, 1995). Evidence for the glycosylation of the other two antigens, 38 kDa and 19 kDa lipoproteins, is based on their ability to bind to the lectin concanavalin A (ConA) in Western blot assays (Espitia and Mancilla, 1989; Garbe *et al.*, 1993). Similar results have been reported in studies of the closely related mycobacterium responsible for bovine tuberculosis (Fifis *et al.*, 1991). Although sequence data are available for all three antigens (Ashbridge *et al.*, 1989; Andersen *et al.*, 1990; Laqueyrerie *et al.*, 1995), a functional role has been ascribed only in the case of the 38 kDa lipoprotein, which is related to the PstS periplasmic binding protein involved in phosphate uptake by *Escherichia coli* (Andersen *et al.*, 1990).

The present study focuses on analysis of the *M.tuberculosis* 19 kDa lipoprotein antigen (Ashbridge *et al.*, 1989; Young and Garbe, 1991). The 19 kDa antigen was originally identified by monoclonal antibodies raised against whole cell extracts, and is a prominent target of humoral and cellular immune responses to *M.tuberculosis* in both mouse and man (for review, see Ivanyi and Thole, 1994). Hybrid proteins based on the 19 kDa gene have been used to induce a strong immune response to heterologous antigens in recombinant BCG vaccines (Stover *et al.*, 1993). We have previously shown that glycosylation of the *M.tuberculosis* 19 kDa antigen, as evidenced by ConA binding, can be reproduced by its expression as a recombinant protein in a rapidly growing mycobacterial host (Garbe *et al.*, 1993). The recombinant protein expressed in conventional *E.coli* systems shows no evidence of glycosylation (Garbe *et al.*, 1993). As an approach to studying the biological function and possible immunological significance of protein glycosylation in mycobacteria, we have exploited the mycobacterial expression system to identify regions of the 19 kDa antigen involved in ConA binding, and to prepare non-glycosylated variants in which key amino acid residues have been substituted by site-directed mutagenesis. Changes in molecular size of the mutant proteins suggest that the function of glycosylation may be to regulate proteolytic cleavage of the 19 kDa antigen.

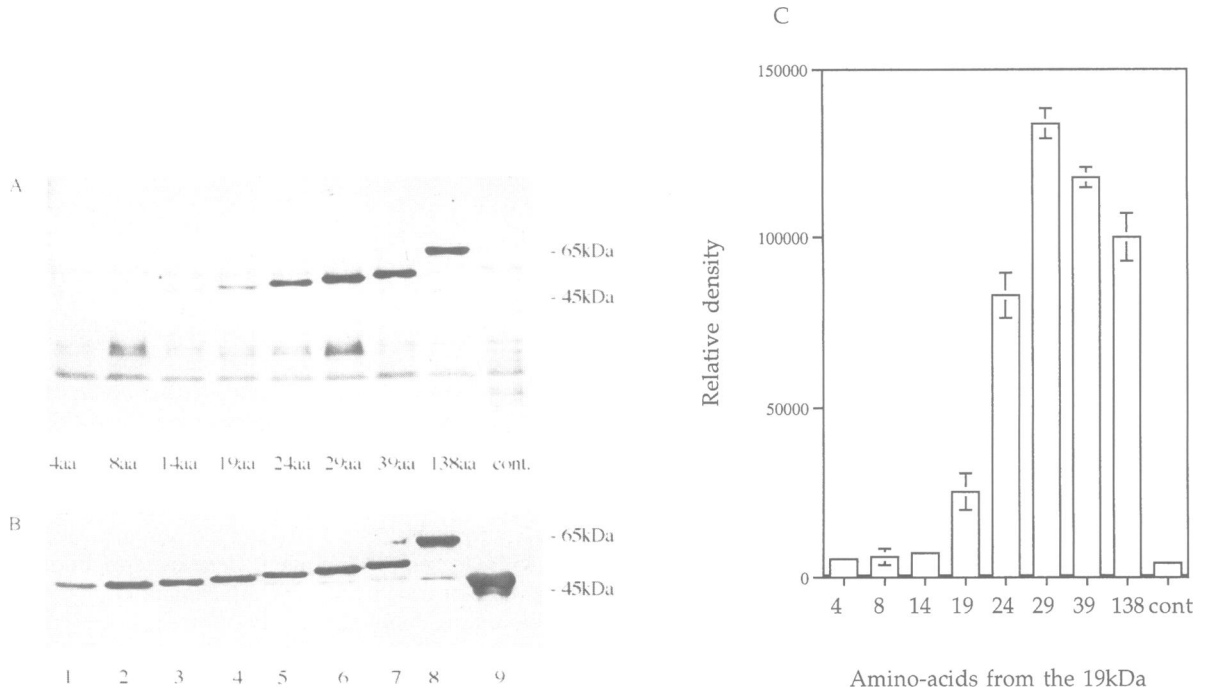

Fig. 1. ConA binding by 19 kDa–PhoA hybrid proteins. Sonicated preparations of *M.smegmatis* expressing PhoA fused to various fragments from the 19 kDa protein were analysed by SDS–PAGE and Western blotting. (**A**) Staining for ConA binding. Lanes 1–8 show hybrid proteins with an increasing number of amino acids from the 19 kDa protein as indicated at the foot of each lane. Amino acid numbering is based on the mature protein, after removal of the signal peptide. Lane 9 contains an extract from *M.smegmatis* transformed with pSMT3 containing the *phoA* gene only. Standard molecular size markers are indicated. In addition to several lower molecular size bands, ConA binds weakly to two bands of 53 and 58 kDa in all *M.smegmatis* extracts. A stronger binding to hybrid proteins of increasing molecular size (54–65 kDa) is clearly seen in lanes 4–8. (**B**) The same preparations stained with a polyclonal antiserum to PhoA. Lane 9 contains purified PhoA protein. All of the hybrid constructs expressed comparable levels of PhoA. (**C**) Analysis of the ConA binding blot by scanning densitometry. Intensity of staining in the area of the blot corresponding to the PhoA hybrid proteins is expressed as the area under the curve deducted from the background (Shea, 1994); bars indicate standard deviation of density values for each sample scanned on different time-exposed gels (n = 3–4). ConA binding by the three shortest hybrid proteins is less than 10% of that seen with the full-length hybrid, and is indistinguishable from the vector control.

Results

Analysis of PhoA hybrid proteins

To identify regions of the 19 kDa protein involved in ConA binding, in-frame hybrid proteins were prepared in which portions of the gene encoding the 19 kDa protein, expressed from its natural promoter, were linked to *E.coli* alkaline phosphatase (PhoA) lacking its own signal sequence. PhoA was used as a hybrid partner in order to allow monitoring of the level of expression of the different constructs, and to verify that protein constructs were being exported across the cell membrane. All of the hybrid proteins described below expressed functionally active alkaline phosphatase in a recombinant mycobacterial expression system (*Mycobacterium smegmatis*) as judged by formation of blue colonies in the presence of appropriate substrate. Glycosylation of PhoA hybrid proteins in mycobacterial protein extracts was monitored by ConA binding activity, after fractionation by denaturing gel electrophoresis (SDS–PAGE) and transfer to nitrocellulose membranes.

Initial experiments with hybrid proteins containing 8, 39 and the full 138 amino acids of the mature 19 kDa lipoprotein indicated that ConA binding was associated with the N-terminal region, and additional hybrid proteins were generated to allow detailed analysis of this part of the protein. Figure 1A shows a nitrocellulose blot of the complete set of hybrid proteins stained with peroxidase-conjugated ConA. The three shortest hybrid proteins (with 4, 8 and 14 amino acids) did not bind ConA; the next hybrid protein, 19 aaPhoA, showed an intermediate level of ConA binding; while the four longest hybrid proteins were all strongly positive. Probing the corresponding blot with antibody to PhoA (Figure 1B) demonstrated a similar amount of hybrid protein in each of the extracts, indicating that differences in ConA binding were not a consequence of differences in expression levels. Densitometric analysis of the ConA blots (Figure 1C) showed that ConA binding by the three 'negative' hybrid proteins was less than 10% of that seen with the hybrid protein containing the full-length 19 kDa gene, and was indistinguishable from the background seen with the *M.smegmatis* vector control. ConA binding was blocked by addition of α-methyl mannoside, or by pretreatment of extracts overnight at room temperature with mild alkali (0.1 M NaOH) (data not shown), suggesting that the interaction is mediated by carbohydrate associated with the hybrid proteins. These experiments demonstrated that ConA binding does not require the whole 19 kDa protein, but can be reproduced by fusion of a short fragment of the 19 kDa gene to PhoA.

Site-directed mutagenesis of PhoA hybrid proteins

The results described above indicate an important role in ConA binding for the peptide region that distinguishes the shortest ConA-positive hybrid protein, 19 aaPhoA, from the longest ConA-negative hybrid protein, 14 aaPhoA. Five threonine residues are present in this region

Table I. Site-directed mutagenesis of 19 aaPhoA

			ConA binding (%)
	1 10	20 30	
	19 aaPhoA lipid - CSSNKSTTGSGE T T T A A G T	-gs-pvl-PhoA	100
	14 aaPhoA lipid - CSSNKSTTGSGE T T	-gs-pvl-PhoA	<10
	V 15 T T V A A G T		70
	V 15,19 T T V A A G V		55
Site-directed mutants in19 aaPhoA	V 13 V T T A A G T		10
	V 14 T V T A A G T		45
	G 16,17 T T T G G G T		<10

ConA binding by 19 aaPhoA was characterized by substitution of amino acid residues. Mutant proteins were analysed by SDS–PAGE and blotting with peroxidase-conjugated ConA as described in Figure 1. Blots were analysed by scanning densitometry, and results shown as percentage binding, relative to unmodified 19 aaPhoA. (Residues indicated in lower case include two amino acids, glycine and serine, introduced by the *Bam*HI linker, and the first three amino acids from PhoA.)

of the 19 kDa protein, at positions 13, 14, 15, 19 and 20. In light of the identification of threonine as attachment site for O-linked carbohydrate in the *M.tuberculosis* 45 kDa antigen (Dobos *et al.*, 1995), we focused particular attention on the effect of mutagenesis of threonine residues in 19 aaPhoA (Table I). Substitution of individual threonine residues by valine did not abolish ConA binding, though a marked reduction in binding was observed in some cases. Substitution of Thr13 in 19 aaPhoA had the most marked effect, reducing ConA binding by 90%. Substitution of alanine residues by glycine at positions 16 and 17 also inhibited ConA binding by 19 aaPhoA.

Similar mutations were subsequently introduced into longer PhoA hybrid proteins containing 39 and the full 138 amino acids from the 19 kDa protein (Figure 2). Substitution of valine for threonine at positions 13 and 14 reduced, but did not eliminate, ConA binding by 39 aaPhoA (Figure 2A, lane 3). In contrast to results with 19 aaPhoA, however, substitution of alanine residues 16 and 17 had no effect on ConA binding by the longer hybrid protein (Figure 2A, lane 4). Substitution of the threonine doublet at positions 19–20 had less effect on its own (Figure 2A, lane 5) but, in combination with the replacement of the 13–15 triplet, caused complete inhibition of ConA binding by 39 aaPhoA (Figure 2A, lane 7). Results obtained with the full-length 138 aaPhoA hybrid protein were similar to those with 39 aaPhoA, indicating a critical role in ConA binding for both the threonine triplet (13–15) and the threonine dimer (19–20) (Figure 2A, lanes 7–10). Again, simultaneous substitution of both threonine clusters generated a ConA-negative hybrid protein. Immunoblot analysis of the hybrid constructs with antibody to PhoA revealed some heterogeneity in the molecular size of site-directed mutants (Figure 2B).

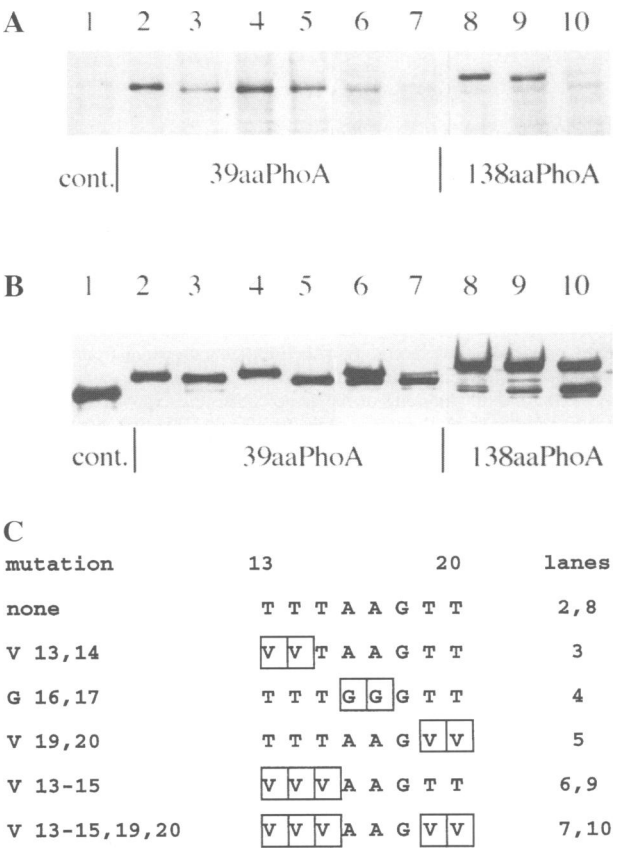

Fig. 2. Effect of site-directed mutagenesis on ConA binding by PhoA hybrids. Amino acids within the 13–20 region of the 19 kDa fragment were substituted in 39 aaPhoA (lanes 2–7) and 138 aaPhoA hybrid proteins (lanes 8–10). Panels (A) and (B) show portions of the gel corresponding to the molecular size region 53 to 84 kDa. (**A**) ConA binding was analysed by SDS–PAGE and blotting as in Figure 1. Lane 1 contains vector-transformed *M.smegmatis* control. Lanes 2 and 8 contain unmodified 39 aaPhoA and 138 aaPhoA respectively. Mutant proteins in the remaining lanes are detailed in (C). For each hybrid protein, simultaneous substitution of both threonine clusters removed ConA binding (lanes 7 and 10). (**B**) A blot equivalent to that in (A) was stained with antibody to PhoA. Lane 1 contained purified PhoA protein. Variations were observed in apparent molecular sizes of hybrid constructs altered by site-directed mutagenesis. (**C**) Amino acid substitutions introduced by site-directed mutagenesis.

Site-directed mutagenesis of the 19 kDa protein

To test whether results obtained with the PhoA hybrid system could be extrapolated to the 19 kDa protein itself, the key threonine triplet and doublet sites were modified in constructs expressing the non-fused 19 kDa coding gene (Figure 3). Immunoblot analysis with monoclonal antibody to the 19 kDa antigen showed marked changes in apparent molecular size of mutant forms of the protein. The original recombinant protein migrates at a molecular size of 22 kDa in our gel system (Figure 3A, lane 2). When the 13–15 threonine triplet was substituted by valine residues (3V2T-19 kDa; Figure 3A, lane 3), two major bands were observed on antibody blots, with apparent molecular sizes of 21 kDa and 17 kDa. Substitution of

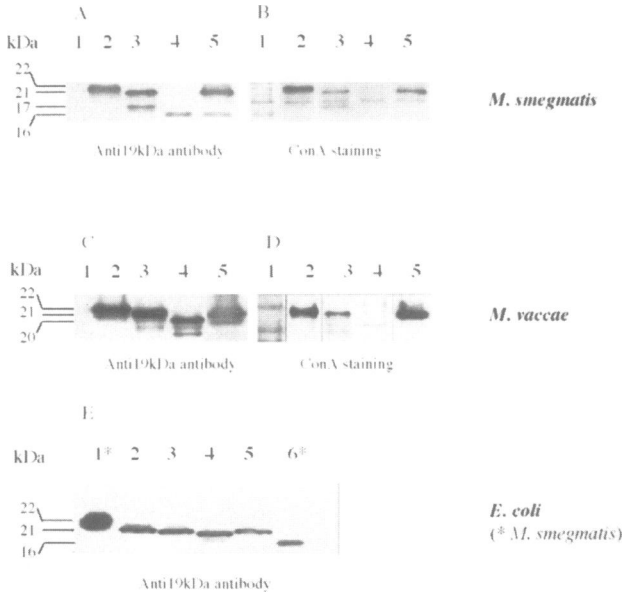

Fig. 3. Site-directed mutagenesis of the 19 kDa antigen. The 19 kDa antigen was mutated by substitution of the threonine triplet at positions 13–15 (3V2T-19 kDa), the threonine doublet at positions 19 and 20 (3T2V-19 kDa), or all five threonines (5V-19 kDa). Portions of the gel corresponding to the molecular size region 15 to 23 kDa are shown. (**A**) A 19 kDa antibody blot of cell extracts from *M.smegmatis* expressing each of the constructs. Lane 1, vector control; 2, unmodified 19 kDa; 3, 3V2T-19 kDa; 4, 5V-19 kDa; 5, 3T2V-19 kDa. (**B**) ConA staining of the same samples. (**C** and **D**) Antibody and ConA staining respectively of cell extracts prepared from constructs expressed in *M.vaccae*. Lane numbers are the same as in (A). (**E**) Antibody staining of the same constructs expressed in *E.coli*. Lanes 1 and 6 show the unmodified 19 kDa and 5V-19 kDa expressed in *M.smegmatis*, respectively. Lanes 2–5 show unmodified 19 kDa, 3V2T-19 kDa, 5V-19 kDa and 3T2V-19 kDa expressed in *E.coli*, respectively.

the 19–20 threonine doublet (3T2V-19 kDa; Figure 3A, lane 5) also generated two bands, at 21 kDa and 16 kDa. When all five threonine residues were substituted (5V-19 kDa; Figure 3A, lane 4), a single antibody-positive band was seen with apparent molecular size of 16 kDa. Analysis of culture filtrate samples showed that 5V-19 kDa was present predominantly in a soluble, as distinct from cell-associated form (data not shown).

Figure 3B illustrates ConA binding activity of the various forms of the 19 kDa protein. The blot has been exposed for a relatively long period to show ConA binding by the lower molecular size band from 3V2T-19 kDa (Figure 3B, lane 3), in contrast to the ConA-negative low molecular size forms of 3T2V-19 kDa and 5V-19 kDa (Figure 3B, lanes 4 and 5). A ConA-positive band seen in all the lanes at a molecular size of ~19 kDa corresponds to an endogenous *M.smegmatis* protein, that is not recognized by antibody to the *M.tuberculosis* antigen. As observed with the PhoA hybrid proteins, modification of the threonine triplet or doublet alone has some effect on ConA binding, while substitution of all five threonine residues reduces binding to background levels.

When the same constructs were expressed in an alternative mycobacterial host, *Mycobacterium vaccae*, a different pattern was seen. Some reduction in molecular size was observed following threonine substitutions, but the 16 kDa and 17 kDa forms seen in *M.smegmatis* were much less prominent (Figure 3C). In *M.vaccae*, the predominant

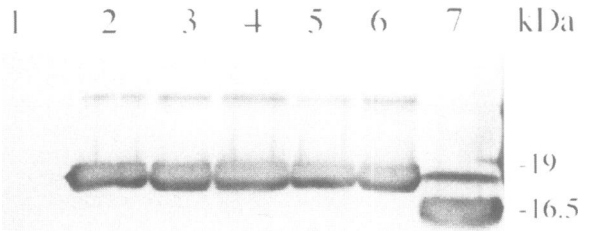

Fig. 4. Proteolytic activity of a soluble extract from *M.smegmatis*. His-19 kDa was purified from *E.coli* and incubated with a sonic extract prepared from *M.smegmatis* as described in Materials and methods. Aliquots were removed at different times and analysed by Western blot with a monoclonal antibody (F29-47) directed against the 19 kDa antigen. Lane 1 shows a control extract without His-19 kDa. Lanes 2–4 show His-19 kDa incubated with boiled extract for 0, 4 and 24 h, respectively. Lanes 5–7 show similar incubations in presence of fresh soluble extract at 0, 4, and 24 h, respectively. After overnight incubation, His-19 kDa is degraded to a stable 16.5 kDa product.

form of 5V-19 kDa had an apparent molecular size of ~20 kDa (Figure 3C, lane 4). As in the case of the *M.smegmatis* recombinant, ConA binding by 5V-19 kDa expressed in *M.vaccae* was less than 10% of that of the unmodified 19 kDa protein (Figure 3D, lane 4), and was not readily distinguishable from bands seen in the vector control (Figure 3D, lane 1).

The 19 kDa antigen is expressed at a slightly lower molecular size in *E.coli* in comparison with *M.smegmatis* (Garbe *et al.*, 1993); threonine substitutions had only a very slight effect on the molecular size of 19 kDa constructs expressed in *E.coli* (Figure 3E). *E.coli*-expressed recombinants showed no detectable binding to ConA.

Proteolytic cleavage of the 19 kDa protein

The low molecular size form of the mutant 19 kDa protein expressed in *M.smegmatis* was further characterized by N-terminal amino acid sequencing. Concentrated culture filtrate from 5V-19 kDa was fractionated by gel electrophoresis and transferred to a PVDF membrane. Sequence analysis of excised protein bands generated three signals. The major component (accounting for 66% of the total protein) had an N-terminal sequence initiating at the valine residue corresponding to amino acid 19 of the mature protein, i.e. VVASPGAASGPKVVIDGKDQ. Minor components were detected with N-terminal sequences initiating at Ala21 (20%), and at Ser22 (14%). Thus, the low molecular size form of the protein corresponds to a partial degradation product of the 19 kDa antigen, lacking the N-terminal acylation region.

To test whether cleavage of the 19 kDa protein was a consequence of the presence of the valine substitutions, a non-glycosylated 19 kDa protein was purified from *E.coli* and incubated with a *M.smegmatis* extract. The *E.coli* recombinant (His-19 kDa) contained the original threonine sequence and was expressed with a poly-histidine tag at the C-terminus to facilitate purification. After overnight incubation, His-19 kDa (observed molecular size 19 kDa) was converted to a low molecular size form (16.5 kDa) (Figure 4). Proteolysis was prevented by heat-inactivation of the *M.smegmatis* extract before addition to the assay.

Discussion

Mapping of glycosylation sites

The PhoA hybrid strategy has been used to study protein secretion in a wide range of bacteria (Manoil *et al.*, 1990), including mycobacteria (Timm *et al.*, 1994), but has not previously been applied to the analysis of glycosylation. The demonstration that the ConA binding activity of the 19 kDa antigen is retained by a short fragment of the gene fused to PhoA, and furthermore, that binding is lost by the alteration of a few specific amino acid residues, provides strong support for the hypothesis that ConA binding is due to glycosylation. Confirmation that the 19 kDa antigen is a glycoprotein will, however, require structural evidence of a covalent sugar–protein linkage, comparable with that reported for the *M.tuberculosis* 45 kDa antigen (Dobos *et al.*, 1995). Preliminary results with the purified 19 kDa protein indicate that the associated carbohydrate is predominantly mannose, and can be released under conditions appropriate for hydrolysis of O-glycosidic linkages (C.Abou-Zeid and D.Young, unpublished data).

The two threonine clusters substituted in the ConA-negative constructs (residues 13–15 and 19–20) may provide sites for O-glycosylation, although it is important to bear in mind that, in addition to alteration of sugar attachment sites, the amino acid substitutions could cause conformational changes that may have a less direct effect on glycosylation. In contrast to N-glycosylation sites, there is no strict consensus motif for O-glycosylation. In addition to threonine and serine clusters involved in attachment of sugars, adjacent proline and alanine residues have important local conformational effects (Gooley *et al.*, 1991; O'Connell *et al.*, 1991, 1992; Wilson *et al.*, 1991). Inhibition of ConA binding by substitution of alanine residues in 19 aaPhoA, but not in longer hybrids or in the 19 kDa protein itself, is most easily explained in terms of local conformational influences, and we cannot discount the possibility that some or all of the threonine substitutions in 5V-19 kDa may have conformational effects rather than acting directly as a site for carbohydrate attachment. Screening of the 19 kDa sequence using an algorithm based on mammalian glycoproteins (Hansen *et al.*, 1995) predicted all five threonine residues as potential sites of O-glycosylation.

It is interesting to compare the amino acid sequence in the ConA binding region with that of similar 19 kDa antigens from other mycobacteria. The 19 kDa gene of *Mycobacterium bovis* is identical to that of *M.tuberculosis* (Collins *et al.*, 1990), and closely related genes have been described from two members of the *Mycobacterium avium-intracellulare* (MAI) complex (Nair *et al.*, 1992; Booth *et al.*, 1993). All three genes encode a typical signal peptide ending with a lipoprotein motif, and share more than 70% sequence identity. In contrast to the bulk of the protein, however, the region between amino acids 7 and 29 is hypervariable, with very little conservation between *M.tuberculosis* and MAI, and considerable diversity even between the two MAI strains (Figure 5). Structural requirements for sequence conservation within the glycosylation-associated N-terminal portion of the 19 kDa protein, clearly differ from those affecting the major C-terminal domain of the protein. The MAI sequences lack both of the

```
                1                                   40
M. tuberculosis CSS-NKSTTGSGETTTAAGTTASP--GAASGPKVVIDGKDQNV
                ||| |||   |  ***  **||  | | | || ||||||||
M. avium        CSSGNKSAPSSSASSSSTSPSASS--GGAAGTKVIIDGKDQNV
                || ||||  | | | ||  |||  |||||||| ||||||||
M. intracellulare CSGGNKSGTSASSSASSSGTSASASTGGAAGTKVTIDGKDQNV
```

Fig. 5. Comparison of mycobacterial 19 kDa sequences. In contrast to the C-terminal domain, the region of the *M.tuberculosis* 19 kDa protein containing threonine clusters implicated in ConA binding (*) is not conserved in corresponding MAI proteins. There is marked diversity in this region even between the closely related MAI isolates.

threonine clusters implicated in ConA binding, although corresponding regions are rich in serine residues. The serine-rich regions resemble mammalian O-glycoprotein sites (Hansen *et al.*, 1995), but the glycosylation status of the MAI proteins has yet to be determined. The modified peptide identified from the *M.tuberculosis* 45 kDa glycoprotein (Dobos *et al.*, 1995) also conforms to the mammalian glycosylation sequence, and a similar threonine-rich O-glycosylation site is present in the N-terminal region of the 38 kDa ConA binding lipoprotein (21-TTPASSPVT-29). A PhoA hybrid protein containing 40 N-terminal amino acids from the 38 kDa antigen is positive in a ConA binding assay (our unpublished data). The PhoA hybrid protein strategy, together with predictions based on mammalian sequences (Hansen *et al.*, 1995), will be useful in screening for further examples of bacterial proteins with analogous glycosylated linker regions.

Glycosylation and proteolysis

In addition to the effect of threonine substitutions on glycosylation, an unexpected effect was found in alterations of SDS–PAGE profiles. The substitutions themselves, replacement of hydroxyl groups by methyl groups, have negligible effect on molecular size, and caused only minor alteration in mobilities of the *E.coli* recombinants. Reductions of 1–2 kDa in apparent molecular size of the *M.vaccae*-expressed mutants are consistent with the loss of carbohydrate residues associated with one or both of the threonine clusters. In *M.smegmatis*, similar 1–2 kDa reductions were seen, as well as a series of lower molecular size forms with reductions of 5–6 kDa. Proteolytic cleavage of the recombinant products, as confirmed by N-terminal sequencing of the 5V-19 kDa, is the most likely origin of the low molecular size forms in *M.smegmatis*. Protease sensitivity of the mutants could be due to the presence of the valine substitutions, or to the absence of glycosylation. The observation that the 'native' 19 kDa sequence expressed in *E.coli* is similarly converted to a low molecular size form during incubation with *M.smegmatis* extracts, suggests that glycosylation is the key factor. Figure 6 illustrates a working model that takes into account the molecular size and ConA binding properties observed for each of the mutants, in which it is proposed that glycosylation protects a proteolytically sensitive linker region of the 19 kDa protein. Protection of a proteolytically sensitive site by the presence of glycan moieties has previously been proposed for the β-1,4-glycanase from *Cellulomonas fimi* (Langsford *et al.*, 1987; Ong *et al.*, 1994). In the case of the 19 kDa protein, proteolytic cleavage would have the effect of releasing the major conserved domain of the protein from its lipid anchor, and would account for the presence of the lower molecular size forms of the recombinant protein in culture filtrate

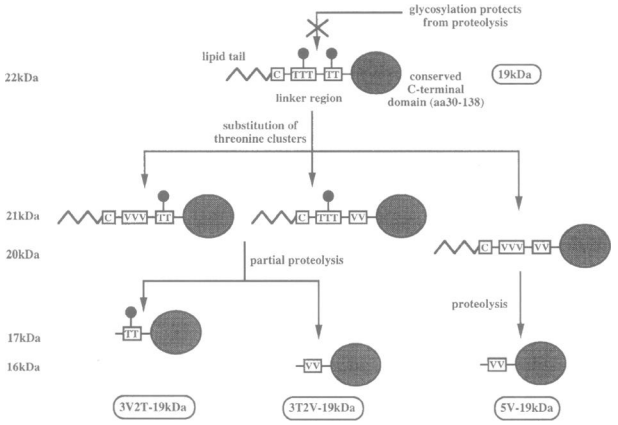

Fig. 6. Working model to account for effect of mutations on apparent molecular size of the recombinant 19 kDa antigen. The *M.tuberculosis* 19 kDa lipoprotein comprises a C-terminal domain that is conserved between different mycobacteria, and a variable linker region implicated in glycosylation. To account for the effect of amino acid substitutions on ConA binding and molecular size, a model is proposed in which glycosylation of one or both of the threonine clusters at positions 13–15 and 19–20 inhibits proteolysis of the linker region. Prevention of glycosylation, by substitution of valine for threonine residues, causes a reduction of 1–2 kDa in molecular size due to loss of sugar moieties, and a larger reduction of 5–6 kDa due to proteolysis. Proteolysis is most efficient after substitution of both threonine clusters, and occurs more readily in *M.smegmatis* than in *M.vaccae*. Glycosylation may be an important mechanism by which bacteria can regulate a balance between cell-associated and cell-free forms of a protein.

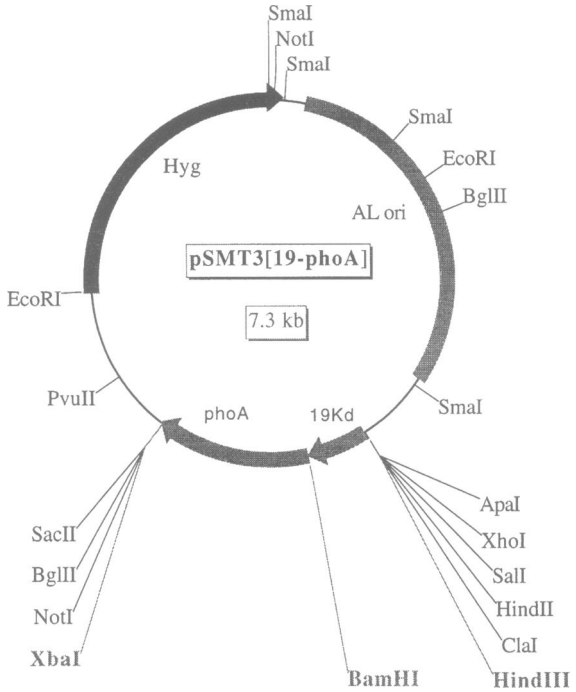

Fig. 7. pSMT3 shuttle plasmid. pSMT3 was generated from p16R1 (Garbe *et al.*, 1994) by insertion of polylinker regions at the original *Eco*RV cloning site. The strategy for construction of the 19 kDa–PhoA hybrid proteins is illustrated.

samples. Different molecular size forms of the native 19 kDa antigen are also evident in immunoblots of *M.tuberculosis* (e.g. Young and Garbe, 1991). An important feature of the working model proposed in Figure 6 is that it makes a series of predictions about structures and proteolytic activities that are now open to further experimental analysis.

The proteolysis model is attractive in providing an explanation for our findings and also in suggesting a possible biological function for glycosylation of bacterial proteins. Attachment and release of cell surface proteins is known to play a critical role in the physiology and virulence of Gram-positive bacteria (Schneewind *et al.*, 1993, 1995; Jenkinson, 1995; Sutcliffe and Russel, 1995). In *Staphylococcus aureus*, for example, secreted and cell-associated (lipoprotein) forms of the β-lactamase enzyme are thought to arise as a result of differential processing by type I and type II signal peptidases (Wiley *et al.*, 1996). The glycosylation/proteolysis system proposed here would allow bacteria to regulate the relative amounts of cell-associated and soluble forms of a protein by changes in activity of glycosylation or proteolytic enzymes. As noted above, a similar N-terminal linker region is present in other mycobacterial lipoproteins, and regulation of proteolysis may prove to be one general function of protein glycosylation in bacteria.

It has been proposed that antigens released into the culture medium play a particularly important role in the immune response to *M.tuberculosis* (Andersen, 1994; Andersen *et al.*, 1995; Horwitz *et al.*, 1995; Orme, 1995), and a mechanism for release of lipoprotein antigens could have important immunological consequences. Recombinant proteins carrying the glycosylated and non-glycosylated linker region from the 19 kDa lipoprotein will be of use in analysing antigen presentation by *M.tuberculosis* in animal and in cell culture models.

Materials and methods

Construction of PhoA hybrid proteins

Fragments of the *M.tuberculosis* 19 kDa antigen gene were amplified by polymerase chain reaction (PCR) using appropriate primers for the 19 kDa gene (Ashbridge *et al.*, 1989) and *phoA* (Haas *et al.*, 1993). PCR was carried out in a DNA thermal cycler (Perkin Elmer Corporation, Bucks, UK) using $Vent_R$ DNA polymerase (New England Biolabs, Hitchin, UK) in reactions consisting of 30 cycles of denaturation at 95°C (1 min), hybridization at 60°C (1 min) and elongation at 72°C (1 min). Amplified fragments were gel-purified (Qiaex, Qiagen Ltd, Dorking, UK), or purified by using Qiagen columns (Qiagen Ltd) and digested with *Bam*HI, *Hin*dIII or *Xba*I, as appropriate, and, after further purification, cloned in the mycobacterial shuttle plasmid pSMT3 (Figure 7) using standard procedures (Sambrook *et al.*, 1989). A 1.4 kbp fragment containing the structural gene for *E.coli phoA* was similarly amplified from pTn*Max3* (kindly provided by Rainer Haas, Max-Planck Institute for Biology, Tübingen, Germany) and cloned downstream of the 19 kDa fragments (Figure 7).

Constructs were initially transformed into *E.coli* DH5α, grown in Luria–Bertani (LB) broth or agar medium with selection in the presence of hygromycin (Boehringer Mannheim, Lewes, UK) at 250 μg/ml, and plasmid preparations obtained using standard methodology. Plasmids were subsequently introduced into *M.smegmatis* $mc^2/1$-2c or *M.vaccae* by electroporation (Garbe *et al.*, 1994). Mycobacteria were grown in LB broth or on agar medium with 0.05% tyloxapol (Sigma, UK). Hygromycin was added as appropriate at a concentration of 50 μg/ml. For visualization of alkaline phosphatase activity of recombinant colonies, mycobacteria were grown on similar plates supplemented with 5-bromo-4-chloro-3-indolyl-phosphate-*p*-toluidine (BCIP; Sigma, Poole, UK) at 40 μg/ml. Positive colonies were dark blue.

For site-directed mutagenesis, forward and reverse PCR primers were designed incorporating appropriate substitutions and with an overlapping region of 18 bp. These were used in combination with the native forward and reverse primers described above initially to amplify two fragments of the mutated 19 kDa gene. After purification, both overlapping 19 kDa amplified fragments were mixed and full mutated insert amplified again using the original forward and reverse primers. These were then digested

by *Hin*dIII and *Bam*HI, purified, and ligated into the *Hin*dIII/*Bam*HI restricted pSMT3::*phoA* (see Figure 7) as previously described.

The nucleotide sequence of the 19 kDa gene fragment and the first 50 nucleotides from *phoA* was determined for all plasmid constructs. Sequencing was performed using a T7 Sequencing kit (Pharmacia Biotech, UK) according to the manufacturer's protocol.

Preparation of samples for analysis

Mycobacterial pellets were obtained by centrifugation (10 000 *g*, 10 min) of 10 ml of late-exponential phase (48 h) cultures in LB medium supplemented with 50 μg/ml of hygromycin and 0.05% tyloxapol (Sigma Corp., UK). Pellets were resuspended in 0.1 ml of phosphate-buffered saline (PBS) and sonicated on ice at 18 μm amplitude (Soniprep 150, MSE, UK) for three periods of 30 s, at 90-s intervals. Total protein content of sonicated preparations was estimated using the Micro BCA protein assay (Pierce, Rockefeller, IL, USA) with bovine serum albumin (BSA) as standard. Culture filtrates were concentrated 100-fold before analysis by centrifugation using Microcon 3 filtration columns (Amicon, Inc., Beverly, MA, USA).

Gel electrophoresis and Western blotting

Sonicated preparations were boiled in the presence of SDS and β-mercaptoethanol and samples (2–4 μg total protein) were analysed by polyacrylamide gel electrophoresis (PAGE) in the presence of SDS, followed by transfer to nitrocellulose membranes (Hybond C, Amersham, UK). For visualization of ConA binding, blots were initially incubated with constant shaking at room temperature for 1 h with 5% BSA in PBS containing 0.1% Tween 20. Blots were then incubated for 1 h with peroxidase-conjugated ConA (Sigma Corp., UK) at a concentration of 0.5 μg/ml in blocking buffer. They were washed three times with 0.05% Tween 20 in PBS, once with PBS alone, and then stained for peroxidase activity using chemoluminescence detection as described by the supplier (Amersham Corp., UK). For visualization of PhoA protein expression, duplicate blots were stained with a rabbit polyclonal antibody raised against PhoA (kindly provided by Juliano Timm, Institut Pasteur, France) at a dilution of 1/5000. Blots were then incubated with peroxidase-conjugated anti-rabbit immunoglobulin (1/5000) (Dako Immunoglobulins, Denmark), washed, and stained for enzyme activity as described for ConA blots. Purified PhoA (Sigma Corp., UK) was included as a positive control. For visualization of expressed 19 kDa site-directed mutants, blots were stained with mouse monoclonal antibody HYT6 (Andersen *et al.*, 1986) or CAM 134.1 (kindly provided by Glyn Hewinson, Central Veterinary Laboratory, New Haw, UK) at a dilution of 1/10000. Blots were then incubated with peroxidase-conjugated anti-mouse immunoglobulin (1/5000) (Bio-Rad Lab., Hercules, CA, USA). Densitometric analysis of autoradiographs was performed with Scan Analysis software (Biosoft, Ferguson, MO, USA) installed on a Macintosh computer (Apple Computer, London, UK) using a Studio Scan II flatbed scanner (Agfa, London, UK). Values were expressed as arbitrary units (relative density) (Shea, 1994), after deduction of background values.

N-Terminal sequencing

Concentrated culture filtrate from 5V-19 kDa expressed in *M.smegmatis* was analysed by Tricine gel electrophoresis (Schagger and von Jagow, 1987), followed by transfer to a PVDF membrane (ABI ProBlot) (Matsudaira, 1987). Protein bands stained with sulforhodamine B (Pappin *et al.*, 1990) were excised and analysed using an ABI476A protein sequencer with standard blot sequencing cycles.

Determination of proteolytic activity in M.smegmatis extracts

M.smegmatis cells (500 mg, wet weight) were resuspended in a buffer containing 50 mM 3-(*N*-morpholin)propanesulfonic acid (MOPS), sonicated as described above and then centrifuged for 10 min at 14 000 r.p.m. (Eppendorf, UK). Freshly prepared supernatants were used in the proteolysis assay. Boiled extract (5 min at 100°C) was used as a negative control. To prepare non-glycosylated 19 kDa protein substrate, the 19 kDa gene was cloned by PCR between a *Nco*I site and a *Bam*HI site in the pQE60 vector (Qiagen Ltd, UK). The resulting lipoprotein (His-19 kDa), tagged with six histidine residues at the C-terminus, was purified by nickel affinity column chromatography (Qiagen Ltd, UK). The purified product was dialysed and adsorbed on a Detoxi-Gel™ column (Pierce, UK) to remove lipopolysaccharide. Samples were freeze-dried and stored at −20°C. His-19 kDa (1 mg/ml) was incubated with *M.smegmatis* extract (1 mg of total protein/ml) at 37°C in a buffer containing 50 mM MOPS and 10 mM $MgCl_2$ in a final volume of 1 ml. Aliquots were removed at 0, 4 and 24 h, and analysed by SDS–PAGE and immunoblot.

Acknowledgements

We are grateful to N.Freeman and W.Blackstock for N-terminal sequencing, and Ken Duncan, Brian Robertson, Koen De Smet, Mark Chambers, Christiane Abou-Zeid and Sarah Ely for helpful discussion and technical advice. J.L.H. thanks P.H.Lagrange for his constant interest in this work. This work was supported by the GlaxoWellcome ActionTB Initiative.

References

Andersen,Å.B., Yuan,Z.L., Haslov,K., Vergmann,B. and Bennesden,J. (1986) Interspecies reactivity of five monoclonal antibodies to *Mycobacterium tuberculosis* as examined by immunoblotting and enzyme-linked immunosorbent assay. *J. Clin. Microbiol.*, **23**, 446–451.

Andersen,Å.B., Ljungqvist,L. and Olsen,M. (1990) Evidence that protein antigen b of *Mycobacterium tuberculosis* is involved in phosphate metabolism. *J. Gen. Microbiol.*, **136**, 477–480.

Andersen,P. (1994) Effective vaccination of mice against *Mycobacterium tuberculosis* with a soluble mixture of secreted mycobacterial proteins. *Infect. Immun.*, **62**, 2536–2544.

Andersen,P., Andersen,Å.B., Sørensen,A.L. and Nagai,S. (1995) Recall of long-lived immunity to *Mycobacterium tuberculosis* infection in mice. *J. Immunol.*, **154**, 3359–3372.

Ashbridge,K.R., Booth,R.J., Watson,J.D. and Lathigra,R.B. (1989) Nucleotide sequence of the 19 kDa antigen from *Mycobacterium tuberculosis*. *Nucleic Acids Res.*, **17**, 1249.

Bock,K., Schuster-Kolbe,J., Altman,E., Allmaier,G., Stahl,B., Christian,R., Sleytr,U.B. and Messner,P. (1994) Primary structure of the O-glycosidically linked glycan chain of the crystalline surface layer glycoprotein of *Thermoanaerobacter thermohydrosulfuricus* L111-69. Galactosyl tyrosine as novel linkage unit. *J. Biol. Chem.*, **269**, 7137–7144.

Booth,R.J., Williams,D.L., Moudgil,K.D., Noonan,L.C., Grandison,P.M., McKee,J.J., Prestidge,R.L. and Watson,J.D. (1993) Homologs of *Mycobacterium leprae* 18-kilodalton and *Mycobacterium tuberculosis* 19-kilodalton antigens in other Mycobacteria. *Infect. Immun.*, **61**, 1509–1515.

Castric,P. (1995) PilO, a gene required for glycosylation of *Pseudomonas aeruginosa* 1244 pilin. *Microbiology*, **141**, 1247–1254.

Collins,M.E., Patki,A., Wall,S., Nolan,A., Goodger,J., Woodward,M.J. and Dale,J.W. (1990) Cloning and characterization of the gene for the 19kDa antigen of *Mycobacterium bovis*. *J. Gen. Microbiol.*, **136**, 1429–1436.

Dobos,K.M., Swiderek,K., Hooi-Khoo,K., Brennan,P.J. and Belisle,J.T. (1995) Evidence for glycosylation sites on the 45-kilodalton glycoprotein of *Mycobacterium tuberculosis*. *Infect. Immun.*, **63**, 2846–2853.

Erickson,P.R. and Herzberg,M.C. (1993) Evidence for covalent linkage of carbohydrate polymers to a glycoprotein from *Streptococcus sanguis*. *J. Biol. Chem.*, **268**, 23780–23783.

Espitia,C. and Mancilla,R. (1989) Identification, isolation and partial characterization of *Mycobacterium tuberculosis* glycoprotein antigens. *Clin. Exp. Immunol.*, **77**, 378–383.

Fifis,T., Costopoulos,C., Radford,A.J., Bacic,A. and Wood,P.R. (1991) Purification and characterization of major antigens from a *Mycobacterium bovis* culture filtrate. *Infect. Immun.*, **59**, 800–807.

Garbe,T., Harris,D., Vordermeir,M., Lathigra,R., Ivanyi,J. and Young,D. (1993) Expression of the *Mycobacterium tuberculosis* 19-kilodalton antigen in *Mycobacterium smegmatis*: immunological analysis and evidence of glycosylation. *Infect. Immun.*, **61**, 260–267.

Garbe,T.R., Barathi,J., Barnini,S., Zhang,Y., Abou-Zeid,C., Tang,D., Mukherjee,R. and Young,D.B. (1994) Transformation of mycobacterial species using hygromycin resistance as a selectable marker. *Microbiology*, **140**, 133–138.

Gerwig,G.J., de Waard,P., Kamerling,J.P., Vliegenthart,J.F.G., Morgenstern,E., Lamed,R. and Bayer,E.A. (1989) Novel O-linked carbohydrate chains in the cellulase complex (cellulosome) of *Clostridium thermocellum*. *J. Biol. Chem.*, **264**, 1027–1035.

Gerwig,G.J., Kamerling,J.P., Vliegenthart,J.F.G., Morag,E., Lamed,R. and Bayer,E.A. (1992) Novel oligosaccharide constituents of the cellulase complex of *Bacteroides cellulosolvens*. *Eur. J. Biochem.*, **205**, 799–808.

Gooley,A.A., Classon,B.J., Marschalek,R. and Williams,K.L. (1991) Glycosylation sites identified by detection of glycosylated amino acids released from Edman degradation: the identification of Xaa-Pro-Xaa-Xaa as a motif for Thr-O-glycosylation. *Biochem. Biophys. Res. Commun.*, **178**, 1194–1201.

Haas,R., Kahrs,A.F., Facius,D., Allmeir,H., Schmitt,R. and Meyer,T.F. (1993) Tn*Max* – a versatile mini-transposon for the analysis of cloned genes and shuttle mutagenesis. *Gene*, **130**, 23–31.

Hansen,J.E., Lund,O., Engelbrecht,J., Bohr,H., Nielsen,J.O., Hansen, J.E.S. and Brunak,S. (1995) Prediction of O-glycosylation of mammalian proteins: specificity patterns of UDP-GalNAc polypeptide *N*-acetylgalactosaminetransferase. *Biochem. J.*, **308**, 801–813.

Horwitz,M.A., Lee,B.W.E., Dillon,B.J. and Harth,G. (1995) Protective immunity against tuberculosis induced by vaccination with major extracellular proteins of *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA*, **92**, 1530–1534.

Ivanyi,J. and Thole,J.E.R. (1994) Specificity and function of T-and B-cell recognition in tuberculosis. In Bloom,B.R. (ed.), *Tuberculosis: Pathogenesis, Protection, and Control.* American Society for Microbiology, Washington, DC, pp. 437–458.

Jenkinson,H.F. (1995) Anchorage and release of Gram-positive bacterial cell surface polypeptides. *Trends Microbiol.*, **3**, 333–335.

Kluepfel,D., Vats-Mehta,S., Aumont,F., Shareck,F. and Morsoli,R. (1990) Purification and characterization of a new xylanase (xylanase B) produced by *Streptomyces lividans* 66. *Biochem. J.*, **267**, 45–50.

Langsford,M.L., Gilkes,N.R., Wakarchuk,W.W., Kilburn,D.G., Miller, R.C.,Jr and Warren,R.A.J. (1984) The cellulase system of *Cellulomonas fimi*. *J. Gen. Microbiol.*, **130**, 1367–1376.

Langsford,M.L., Gilkes,N.R., Singh,N.R., Moser,B., Miller,R.C.,Jr, Warren,R.A.J. and Kilburn,D.G. (1987) Glycosylation of bacterial cellulases prevents proteolytic cleavage between functional domains. *FEBS Lett.*, **225**, 163–167.

Laqueyrerie,A., Militzer,P., Romain,F., Eiglmeier,K., Cole,S. and Marchal,G. (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.

Manoil,C., Mekalanos,J.J. and Beckwith,J. (1990) Alkaline phosphatase hybrids: sensors of subcellular location. *J. Bacteriol.*, **172**, 515–518.

Matsudaira,P. (1987) Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem.*, **262**, 10035–10038.

Mescher,M.F. and Strominger,J.L. (1976a) Purification and characterization of a prokaryotic glycoprotein from the cell envelope of *Halobacterium salinarium*. *J. Biol. Chem.*, **251**, 2005–2014.

Mescher,M.F. and Strominger,J.L. (1976b) Structural (shape-maintaining) role of the cell surface glycoprotein of *Halobacterium salinarium*. *Proc. Natl Acad. Sci. USA*, **73**, 2687–2691.

Messner,P., Christian,R., Kolbe,J., Schulz,G. and Sleytr,U.B. (1992) Analysis of a novel linkage unit of O-linked carbohydrates from the crystalline surface layer glycoprotein of *Clostridium thermohydrosulfuricum* S102–70. *J. Bacteriol.*, **174**, 2236–2240.

Milhoc,A. and Kluepfel,D. (1990) Purification and characterization of a β-glucosidase from *Streptomyces lividans* 66. *Can. J. Microbiol.*, **36**, 53–56.

Nair,J., Rouse,A. and Morris,S.L. (1992) Nucleotide sequence analysis and serologic characterization of the *Mycobacterium intracellulare* homologue of the *Mycobacterium tuberculosis* 19kDa antigen. *Mol. Microbiol.*, **6**, 1431–1439.

O'Connell,B.C., Tabak,L.A. and Ramasubbu,N. (1991) The influence of flanking sequences on O-glycosylation. *Biochem. Biophys. Res. Commun.*, **180**, 1024–1030.

O'Connell,B.C., Hagen,F.K. and Tabak,L.A. (1992) The influence of flanking sequence on the O-glycosylation of threonine *in vitro*. *J. Biol. Chem.*, **267**, 25010–25018.

Ong,E., Kilburn,D.G., Miller,R.C.,Jr and Warren,R.A.J. (1994) *Streptomyces lividans* glycosylates the linker region of β-1,4-glycanase from *Cellulomonas fimi*. *J. Bacteriol.*, **176**, 999–1008.

Orme,I.M. (1995) Prospects for new vaccines against tuberculosis. *Trends Microbiol.*, **3**, 401–404.

Pappin,D.J., Coull,J.M. and Koster,H. (1990) Solid-phase sequence analysis of proteins electroblotted or spotted onto polyvinylidene difluoride membranes. *Anal. Biochem.*, **187**, 1–9.

Parge,H.E., Forest,K.T., Hickey,M.J., Christensen,D.A., Getzoff,E.D. and Tainer,J.A. (1995) Structure of the fibre-forming protein pilin at 2.6Å resolution. *Nature*, **378**, 32–38.

Plummer,T.H.,Jr, Tarantino,A.L. and Hauer,C.R. (1995) Novel, specific O-glycosylation of secreted *Flavobacterium meningosepticum* proteins. *J. Biol. Chem.*, **270**, 13192–13196.

Reinhold,B.B., Hauer,C.R., Plummer,T.H. and Reinhold,V.N. (1995) Detailed structure analysis of a novel specific O-linked glycan from the prokaryote *Flavobacterium meningosepticum*. *J. Biol. Chem.*, **270**, 13197–13203.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schagger,H. and von Jagow,G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.*, **166**, 368–379.

Schneewind,O., Mihaylova-Petkov,D. and Model,P. (1993) Cell wall sorting signals in surface proteins of Gram-positive bacteria. *EMBO J.*, **12**, 4803–4811.

Schneewind,O., Fowler,A. and Faull,K.F. (1995) Structure of the cell wall anchor of surface proteins in *Staphylococcus aureus*. *Science*, **268**, 103–106.

Shea,T.B. (1994) An inexpensive densitometric analysis system using a Macintosh™ computer and desktop scanner. *Biotechniques*, **16**, 1126–1128.

Stimson,E. *et al.* (1995) Meningococcal pilin: a glycoprotein substituted with digalactosyl 2,4-diacetamido-2,4,6-trideoxyhexose. *Mol. Microbiol.*, **17**, 1201–1214.

Stover,C.K., Bansal,G.P., Hanson,M.S., Burlein,J.E., Palaszynski,S.R., Young,J.F., Koenig,S., Young,D.B., Sadziene,A. and Barbour,A.G. (1993) Protective immunity elicited by recombinant bacille Calmette-Guerin (BCG) expressing outer surface protein A (OspA) lipoprotein: a candidate Lyme disease vaccine. *J. Exp. Med.*, **178**, 197–209.

Sutcliffe,I.C. and Russel,R.R.B. (1995) Lipoproteins of Gram-positive bacteria. *J. Bacteriol.*, **177**, 1123–1128.

Timm,J., Perilli,M.G., Duez,C., Trias,J., Orefici,G., Fattorini,L., Amisocante,G., Oratore,A., Oris,B., Frere,J.M., Pugsley,A.P. and Gicquel,B. (1994) Transcription and expression analysis, using *lacZ* and *phoA* gene hybrids, of *Mycobacterium fortuitum* β-lactamase genes cloned from a natural isolate and a high-level β-lactamase producer. *Mol. Microbiol.*, **12**, 491–504.

Wiley,W.N., Daefler,S. and Schneewind,O. (1996) Cell wall sorting of lipoproteins in *Staphylococcus aureus*. *J. Bacteriol.*, **178**, 441–446.

Wilson,I.B.H., Gavel,Y. and von Heijne,G. (1991) Amino acids distribution around O-linked glycosylation sites. *Biochem. J.*, **275**, 529–534.

Young,D.B. and Garbe,T.R. (1991) Lipoprotein antigens of *Mycobacterium tuberculosis*. *Res. Microbiol.*, **142**, 55–65.

Received on February 15, 1996; revised on March 20, 1996