THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 35, Issue of August 27, pp. 36771–36777, 2004

Printed in U.S.A.

Mycobacterium tuberculosis Antigen 85A and 85C Structures Confirm Binding Orientation and Conserved Substrate Specificity*

Received for publication, January 25, 2004, and in revised form, June 8, 2004
Published, JBC Papers in Press, June 10, 2004, DOI 10.1074/jbc.M400811200

Donald R. Ronning‡§, Varalakshmi Vissa¶, Gurdyal S. Besra‖, John T. Belisle¶, and James C. Sacchettini‡ ‡‡**

From the ‡Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas 77843-2128, the ‖School of Biosciences, The University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom, and the ¶Department of Microbiology, Colorado State University, Fort Collins, Colorado 80523

The maintenance of the highly hydrophobic cell wall is central to the survival of *Mycobacterium tuberculosis* within its host environment. The antigen 85 proteins (85A, 85B, and 85C) of *M. tuberculosis* help maintain the integrity of the cell wall 1) by catalyzing the transfer of mycolic acids to the cell wall arabinogalactan and 2) through the synthesis of trehalose dimycolate (cord factor). Additionally, these secreted proteins allow for rapid invasion of alveolar macrophages via direct interactions between the host immune system and the invading bacillus. Here we describe two crystal structures: the structure of antigen 85C co-crystallized with octylthioglucoside as substrate, resolved to 2.0 Å, and the crystal structure of antigen 85A, which was solved at a resolution of 2.7 Å. The structure of 85C with the substrate analog identifies residues directly involved in substrate binding. Elucidation of the antigen 85A structure, the last of the three antigen 85 homologs to be solved, shows that the active sites of the three antigen 85 proteins are virtually identical, indicating that these share the same substrate. However, in contrast to the high level of conservation within the substrate-binding site and the active site, surface residues disparate from the active site are quite variable, indicating that three antigen 85 enzymes are needed to evade the host immune system.

The antigen 85 proteins (85A, 85B, and 85C) from *Mycobacteria* are homologous proteins that share very high sequence identity (68–79%, in *Mycobacterium tuberculosis*). These secreted enzymes were shown to possess a mycolyltransferase activity that is pivotal for the biosynthesis of the mycobacterial cell wall and for the survival of mycobacteria (1). Sequence analysis identified a carboxylesterase consensus sequence (G*X*-S*X*G), and mutation of the central serine to an alanine residue resulted in a loss of enzymatic function (1).

Subsequent x-ray crystallographic studies identified the fold of the enzyme resulting in placement of the antigen 85 proteins within the group of α/β hydrolases (2). Inclusion of the antigen 85 proteins into this family confirmed the importance of the consensus sequence, because numerous mechanistic and structural studies of α/β hydrolases have implicated a conserved pentapeptide containing serine as the initiating nucleophile in a hydrolytic attack on the enzyme substrate. Extension of this from a hydrolytic to a transfer mechanism, in the case of antigen 85, is rather straightforward (2). Whereas a hydrolytic mechanism would utilize a water molecule to hydrolyze the bond between the enzyme and the acyl chain, the transfer mechanism of the antigen 85 enzymes likely uses a hydroxyl group of a terminal arabinose on the mycobacterial cell wall.

Targeted inhibition of the antigen 85 proteins *in vivo* in *Mycobacterium aurum* using a substrate analog, ADT,[1] showed a decrease in the level of production of trehalose monomycolate, trehalose dimycolate, and cell wall linked mycolic acids (1). In a different study, an antigen 85C knockout strain of *M. tuberculosis* showed a 40% decrease in the amount of cell wall linked mycolic acid, whereas the ratio between the three types of mycolic acid (α-, methoxy-, and ketomycolates) bound covalently to the *M. tuberculosis* cell wall is the same as that found in the parent (H37Rv) strain (3). This study also showed that knockouts of 85A and 85B resulted in no noticeable effect on the amount of cell wall attached mycolic acids, indicating that antigens 85A and 85B have minimal roles in mycobacterial survival or that 85C could compensate for the loss of 85A and 85B. In contrast, a subsequent study confirmed the importance of the antigen 85A and 85B proteins by showing that they are indeed directly involved in the covalent attachment of mycolic acids to the mycobacterial cell wall (4).

Furthermore, a study utilizing antisense sequences to the antigen 85 genes (*fbpA*, *B*, and *C2*) demonstrated that inhibition of protein production by these three genes inhibits bacterial growth significantly better than inhibiting only antigen 85C production (5). Although very important in determining the nature of the chemistry performed by the antigen 85 proteins, as well as the relative importance, these studies did not specifically identify the enzyme's substrate.

Environmental factors also modulate the expression level of

* This work was supported in part by the Robert A. Welch Foundation and by Grant GM62410 from the National Institutes of Health. Use of the Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source was supported by the Office of Energy Research, United States Department of Energy, under Contract W-31-109-ENG-38. Use of the BioCARS Sector 14 was supported by the National Institutes of Health, National Center for Research Resources, under Grant RR07707. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (codes 1SFR and 1VA5) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§ Present address: Laboratory of Molecular Biology, NIDDK, National Institutes of Health, Bethesda, MD 20892.

** Currently a Lister Jenner Research Fellow and supported by the Medical Research Council.

‡‡ To whom correspondence should be addressed: Dept. of Biochemistry and Biophysics, Texas A&M University, 2128 TAMU, College Station, TX 77843-2128. Tel.: 979-862-7636; Fax: 979-862-7638; E-mail: sacchett@tamu.edu.

[1] The abbreviations used are: ADT, 6-azido-6-deoxy-α,α'-trehalose; OSG, octylthioglucoside; MR, molecular replacement.

the different antigen 85 proteins (6). Using reverse transcriptase-PCR assays, mycobacteria grown in Sauton media were shown to produce mRNA from only the *fbpB* and *fbpC2* genes, which encode for the antigen 85B and 85C proteins, respectively. Mycobacteria grown in resting macrophages showed transcription of the mRNA of the *fbpA* gene, which codes for antigen 85A. Gene expression studies of mycobacteria grown in activated macrophages showed production of only the *fbpC2* gene. Thus differential expression of the antigen 85 proteins may be a very important mechanism utilized by *M. tuberculosis* to evade the human immune system or to persist in a chronic infection. Another study looked at bacterial viability after knocking out various antigen 85 genes (7). In the antigen 85B knockout, no apparent growth defects were observed in either synthetic media or "macrophage-like" cell lines, whereas an 85A knockout grew in rich media, but not in minimal media or macrophage-like cell lines. These data strengthen the hypothesis that differential expression of the antigen 85 genes may be a mechanism utilized by mycobacteria to confuse and evade the host immune system.

M. tuberculosis also evades the immune system through interactions with human fibronectin (8). Although antigen 85 proteins possess a fibronectin-binding activity, the identity of the domain, or domains, to which antigen 85 binds is currently not clear. Different binding studies have implicated either the heparin and cellular-binding domains (9), or the collagen-binding domain of fibronectin (10) in this activity. Regardless of which fibronectin domain interacts with the antigen 85 proteins, it seems clear that this is an important interaction. The interaction between the antigen 85 proteins and fibronectin stimulates monocyte-derived macrophages to phagocytose *Mycobacterium leprae* (11). The addition of anti-antigen 85 antibodies to this assay decreased the amount of bacterial uptake. Similar strategies are utilized by other pathogens to ensure cellular attachment and promote the initiation of infection. Surface proteins from both *Streptococcus pyogenes* and *Neisseria gonorrhoeae* bind to fibronectin allowing for the efficient invasion of host cells (12–13). Studies on the *Borrelia burgdorferi* protein BBK32 showed that it also binds fibronectin and that antibodies to BBK32 promote protective immunity in mice (14). Considering this, it is possible that by inhibiting the interaction between the antigen 85 proteins and fibronectin, one of the pathways that *M. tuberculosis* utilizes to enter host cells, would potentially slow down infection.

In an attempt to address some lingering questions concerning the roles of the different antigen 85 proteins, structural studies were extended to include antigen 85A and co-crystal structures with substrate analogs. The structure of antigen 85C with octylthioglucoside (OSG) was solved to gain information concerning substrate binding orientation and interactions important for this binding within the antigen 85 active site. The structure of antigen 85A was solved to allow for a thorough comparison of all three *M. tuberculosis* antigen 85 proteins. This comparison focuses on the substrate-binding site and the active site along with surface residue conformation, which leads to a better understanding of why the tubercle bacillus utilizes three different enzymes to perform the same function.

EXPERIMENTAL PROCEDURES

Expression Construct Creation—The *fbpA* gene was amplified using the PCR from a plasmid containing the gene and inserted into pET30b (Novagen) between the NdeI and XhoI (New England Biolabs) restriction sites. Sequencing of the plasmid indicated the presence of the *fbpA* gene and confirmed that the construct possessed the correct sequence. This plasmid (pDRA1) was used to transform *Escherichia coli* BL21(DE3) cells, and expression tests indicated the production of recombinant antigen 85A.

Protein Purification, Crystallization, and Structure Determination—Both antigens 85A and 85C were purified as previously described (2). However, the treatment of the diffraction data from antigens 85A and 85C are slightly different. Crystallization and crystallographic details for antigen 85A are as follows. Recombinant antigen 85A was crystallized in 100 mM cacodylate buffer (pH 6.5), and 2.4 M sodium acetate. The antigen 85A crystals grew in a $C222_1$ space group with unit cell dimensions a = 131.4 Å, b = 288.7 Å, c = 101.3 Å, and $\alpha = \beta = \gamma = 90°$. Addition of glycerol to a final concentration of 20% v/v was used as a cryoprotectant for data collection at beamline 14BM-C at the Advanced Photon Source at Argonne National Labs. Crystals were flash-cooled to 100 K, and x-rays with a wavelength of 1 Å were utilized. The antigen 85A data were reduced using Denzo and Scalepack (15) and converted to structure factors with CNS (16) or TRUNCATE (17). Because these crystals were not isomorphous with any of the previous crystal structures for antigen 85B or 85C, molecular replacement was utilized to obtain phases for the antigen 85A structure. The molecular replacement protocol from AMoRe (18) was used to obtain phases for the antigen 85A structure. A monomer of the apo form of antigen 85C (PDB accession 1DQZ) was used as a search model (2).

Calculation of the Matthews coefficient assuming 6 molecules within the asymmetric unit resulted in a V_m of 2.669 and an approximate solvent content of 53% (16, 19, 20). However, only two molecular replacement (MR) solutions were obtained from the rotation and translation searches. Rigid body refinement in the program CNS was performed using these two solutions, resulting in an R-factor and R_{free} of 38 and 44%, respectively. $2F_o - F_c$ and $F_o - F_c$ maps were calculated from the refined model and viewed using the program O (21). Both maps indicated that the two MR solutions were indeed correct (Molecules A and B). However, each of the maps indicated the presence of an additional protein molecule that had not been identified using the MR protocol. The third protein molecule (Molecule C) was manually positioned within the additional density. The model, now containing 3 protein molecules was again subjected to rigid body refinement to ensure proper fitting of the backbone atoms of the manually positioned molecule. The second rigid body refinement yielded an R-factor of 28% and an R_{free} of 33%. Because the search model contained the sequence for antigen 85C, mutations were made to the three protein molecules using the program O (21), so the sequence of the model would match the *M. tuberculosis* antigen 85A sequence. The model was then subjected to the simulated annealing protocol in CNS with a starting temperature of 6000 K resulting in an R-factor of 26% and an R_{free} of 30% (16). Further manual and computational refinement reduced the R-factor to 22.4% and the R_{free} to 25.6%. Water molecules were then added using the automated water pick protocol in CNS (16) and manually through analysis of $F_o - F_c$ maps. The final model includes 3 protein molecules containing residues 0–286 for molecules A and B, residues 8–284 of molecule C, and 54 water molecules. The model has good stereochemistry (Table I) with the Ramachandran plot showing 85.3% of the residues in the core allowed regions. As observed in all other α/β hydrolases, the active site serine, Ser-126, is found in the disallowed region.

The antigen 85C-OSG (octylthioglucoside) crystals were produced in a solution containing 100 mM imidazole buffer (pH 6.5), 1.2 M sodium acetate, and 9 mM OSG. Crystals were cryoprotected using 25% polyethylene glycol 400 and flash-cooled in liquid nitrogen. Data were collected using CuK_α radiation from a rotating anode. The crystals grew in a C2 space group with unit cell dimensions of a = 144.818 Å, b = 68.117 Å, c = 82.281 Å, and $\beta = 119.601$. Utilizing the apo structure of antigen 85C (PDB accession no. 1DQZ) as a search model for molecular replacement, the cross-rotation and translation protocols in CNS produced two clear molecular replacement solutions (16). These solutions were subjected to rigid body refinement in CNS, where each monomer was treated as a separate rigid body (16). Subsequent refinement included simulated annealing (starting temperature of 2500 K), energy minimization, and B-factor refinement using CNS (16).

The resulting $2F_o - F_c$ maps were of high quality and matched well with the refined model, indicating that the model was accurate. Inspection of $F_o - F_c$ electron density maps after initial refinement indicated the presence of 2 OSG molecules per molecule of antigen 85C for a total of four OSG molecules per asymmetric unit. The OSG molecules were included into the model for subsequent refinement. Water molecules were added automatically using the water pick protocol in CNS (16) then accepted or rejected manually. The final model of two antigen 85C molecules, 4 OSG molecules, and 262 water molecules, using data to a resolution of 2.02 Å, refined to an R-factor of 19.2% and an R_{free} of 21.5% (Table I) with very good geometry. Only the active site serines, as seen

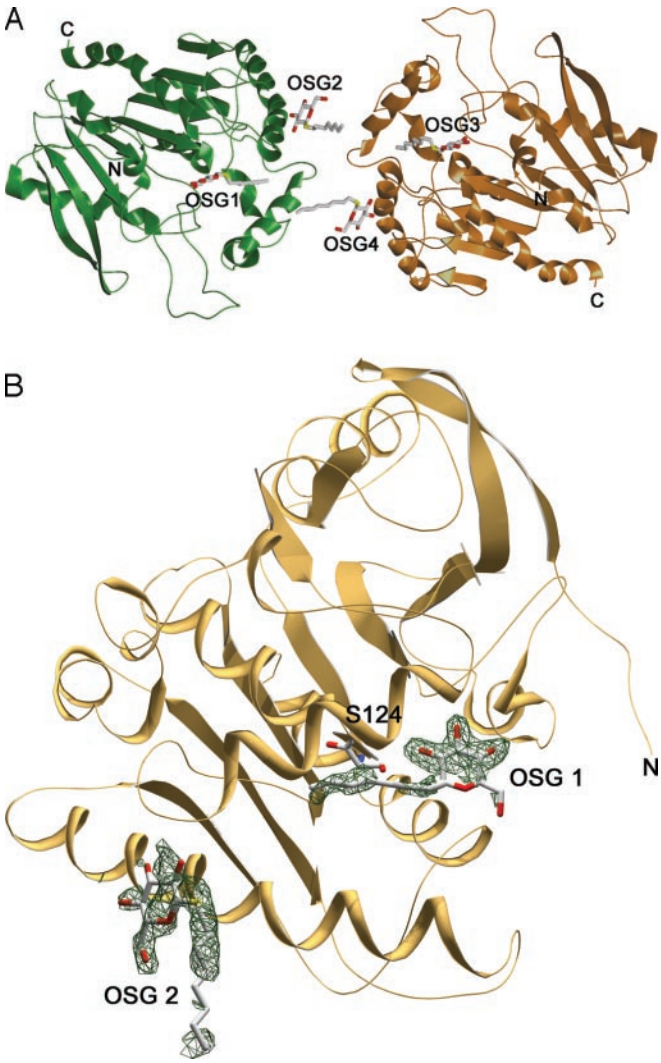

FIG. 1. **Antigen 85C/OSG structure.** *A*, shown is a *ribbon diagram* of the two antigen 85C molecules in the asymetric unit. Molecule 1 is shown in *green*, and molecule 2 is shown in *gold*. The 4 OSG molecules are shown and labeled as are the N and C termini of the protein molecules. *B*, shown as a *gold ribbon* is a single antigen 85C molecule. The *green lines* represent the $F_o - F_c$ map where the OSG molecules were omitted from the map calculation. The 2 OSG molecules are labeled as is the serine nucleophile (*S124*) that represents the enzyme active site and the N terminus of the protein. *White*, *red*, *blue*, and *yellow bonds* represent carbon, oxygen, nitrogen, and sulfur, respectively.

in all other α/β hydrolase enzymes, were found within the disallowed region of the Ramachandran plot.

RESULTS AND DISCUSSION

Antigen 85C/OSG Crystal Structure—We have crystallized antigen 85C in the presence of the detergent OSG (octylthioglucoside) and solved the crystal structure. Molecular replacement with the apo structure of antigen 85C utilized as the search model indicated two protein molecules per asymmetric unit (Fig. 1*A*). Initial refinement and inspection of $F_o - F_c$ electron density maps clearly indicated the presence of 2 OSG molecules per molecule of antigen 85C, or four OSG molecules per asymmetric unit (Fig. 1*B*).

One molecule of OSG is located within the proposed substrate-binding site of each protein molecule (OSG1 and OSG3). The other two OSG molecules (OSG2 and OSG4) are found in the secondary carbohydrate-binding site corresponding to that observed in the trehalose-bound form of antigen 85B (22). No obvious conformational changes were observed in antigen 85C upon binding of OSG. This was confirmed by superimposing, using the Cα atoms of residues 8–282, the apo form of antigen 85C (2) (PDB accession 1DQZ) and the OSG-bound antigen 85C structures.[2] The resulting root mean square deviation of 0.29 Å indicates that these protein structures are nearly identical. This contrasts with the antigen 85C/DEP (diethyl phosphate) structure (PDB accession 1DQY), where a significant conformational change in helix α9 was proposed to result in the stabilization of the acyl-enzyme intermediate (2). This helix forms a lid over the fatty acid bound within the active site to block the hydrolysis of the acyl-enzyme intermediate. It seems clear from the OSG structure that substrate binding does not induce movement of helix α9. The structure suggests that covalent bond formation between Ser-124 and substrate is the trigger for this conformational change.

The affinity of antigen 85C for OSG seems to be due to its similarity to a portion of the natural substrate, trehalose monomycolate. Inspection of the active site in the OSG-bound structure reveals that both van der Waals and hydrogen bonding interactions contribute to analog binding (Fig. 2*A*). Hydrophobic residues within helices α5 and α9 form >70% of the van der Waals contacts between OSG1 (and OSG3) and the protein. Residues within helix α5 are highly conserved, where Trp-158, Leu-161, Ile-163, and Ala-165 form direct van der Waals contacts with the fatty acyl moiety of the substrate analog. Helix α9 forms the opposite side of the acyl binding pocket. Although few of the residues in helix α9 interact directly with the substrate analog, the high level of conservation of these surface hydrophobic side chains implies that they may be important for either "steering" the substrate into the active site or for interacting with a portion of substrate not seen in this structure, specifically the β-chain of the mycolic acid substrate. Other highly conserved residues involved in interactions between the acyl moiety and antigen 85C include Leu-40, Phe-150, Ile-162, and Leu-227 (Fig. 2*B*), where only residue 150 of antigen 85C (152 of 85A and 85B) differs between the three homologs. In 85C, this is a phenylalanine, whereas in 85A and 85B it is a leucine. Because both are hydrophobic and form a small portion of the "floor" of the acyl-binding site, the difference that seems most relevant is the potential π-stacking of phenylalanine side chains. Phe-150 (or Leu-152 in 85A and 85B) forms a portion of the acyl-binding site floor and interacts with substrate near the branch point of the mycolic acid moiety, which does not possess any unsaturated bonds. Thus, variation at this position seems to be of little consequence in substrate specificity, and the three antigen 85 proteins should have very similar affinities for the same substrate.

Residues that form the active site carbohydrate binding pocket are 100% conserved in the *M. tuberculosis* antigen 85 proteins. The binding of OSG1 within the active site shows that side-chain and backbone atoms of residues Gly-39 and Arg-41 form direct hydrogen bonds with the glucose moiety, whereas side chains of residues Asp-38, Asn-52, and Trp-262 are all found within 4.5 Å of the OSG1 sugar moiety and likely interact with the second (distal) ring of the disaccharide of the antigen 85 substrate or product (Fig. 2*B*).

When comparing the OSG-bound form of antigen 85C with the trehalose-bound antigen 85B structure (22), it is observed that the sugar moiety of OSG1 and trehalose possess nearly identical hydrogen bonding interactions within the active sites of antigens 85B and 85C and that the *O*6-hydroxyl of trehalose is oriented in the same direction as that of the fatty acyl moiety of the OSG molecule. Because the mycolyl chains of trehalose dimycolate are attached at the *O*6 and *O*6′ positions of treha-

[2] J. A. Christopher, unpublished material.

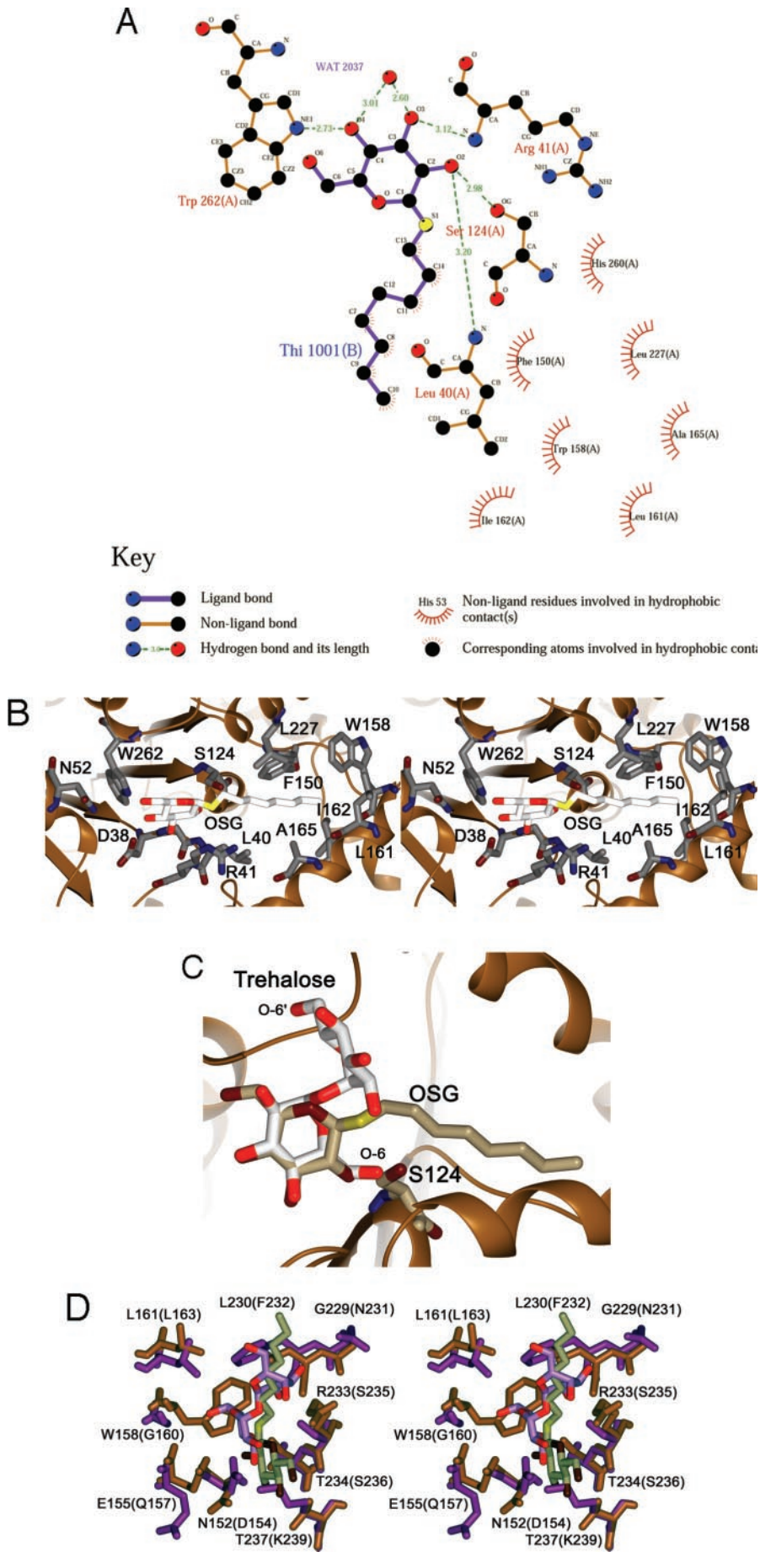

FIG. 2. **Antigen 85C and OSG interactions.** *A*, shown is the output from Ligplot (24). Atoms and bonds of any residues or waters forming hydrogen bonds with OSG are shown along with bond lengths. The van der Waals interactions are shown as mentioned in the figure. *B*, shown is a *stereo diagram* of the antigen 85C active site. The protein backbone is represented by the *gold ribbon*, and side chain bonds are indicated by *gray*, *dark blue*, and *maroon*, representing carbon, nitrogen, and oxygen, respectively. Side chains are also labeled. OSG is labeled, and bonds for carbon, oxygen, and sulfur are shown in *white*, *red*, and *yellow*, respectively. *C*, the antigen 85C/OSG and antigen 85B/trehalose structure were superimposed using the Cα positions of residues 32–282 from antigen 85C. The antigen 85C backbone is shown as a *gold ribbon*. The resulting positions of OSG and trehalose in the enzyme active sites are also shown and labeled as is the Ser-124 nucleophile from antigen 85C. Carbon and oxygen atoms for the trehalose molecule are shown in *white* and *red*, respectively. The *O*6-hydroxyl groups, which are the atoms to which mycolic acids are covalently attached in trehalose dimycolate, are also labeled. For the OSG molecule, carbon, oxygen, and sulfur are shown in *tan*, *maroon*, and *yellow*, respectively. *D*, the atoms of antigens 85B and 85C forming the secondary substrate-binding site are shown in *gold* and *violet*, respectively. Residue numbers for antigen 85C are labeled with the corresponding antigen 85B residue number in *parentheses*. Carbon and oxygen atoms for trehalose are shown in *red* and *light purple*, respectively. OSG carbon, oxygen, and sulfur are shown in *tan*, *maroon*, and *yellow*, respectively.

TABLE I
Crystallographic and refinement data

	Resolution	Completeness (total/highest)	No. of reflections (total/unique)	$I/\sigma I$	R_{sym}[a] (overall/highest)
	Å	%			
Antigen 85A	30–2.7	98.5/98.0	223,878/52,327	15.0	0.069/0.242
Antigen 85C-OSG	30–2.02	95.2/83.9	121,933/44,080	17.5	0.042/0.114

Refinement statistics	Antigen 85A	Antigen 85C-OSG
Resolution (Å)	30–2.7	30–2.02
Reflections in working set	48,101	41,237
R_{cryst} (%)	22.43	19.21
R_{free} (%)	25.57	21.48
r.m.s.d. for bond lengths (Å)	0.008	0.006
r.m.s.d. for bond angles (°)	1.49	1.32
Average *B*-factor (Å^2)	52.1	28.4

[a] $R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma\langle I\rangle$, where I is the observed intensity, and $\langle I\rangle$ is the average intensity of multiple observations of symmetry-related reflections.

lose, this implies that the binding mode of the substrate within the active site of antigen 85B is the same as that seen for antigen 85C (Fig. 2*C*).

The OSG2 (and OSG4)-binding site is in a similar location to that seen in the trehalose/antigen 85B crystal structure (22). However, differences in the interactions at this site indicate that it is not the true substrate-binding site (Fig. 2*D*). The residues of antigen 85C that form direct interactions with the glucose moiety of OSG2 are slightly different than the 85B side chains interacting with trehalose. Asn-152, Asp-155, Trp-158, Asn-152, Thr-234, Thr-237, and Arg-233 of antigen 85C interact directly with OSG2. Although these residues are also involved in binding trehalose in the antigen 85B structure by virtue of the conserved fold, the nature of the residues in this region are not conserved between antigens 85B and 85C. Also, whereas the lipid moiety of OSG2 is well ordered, the protein/analog interactions are mediated by residues from both protein molecules in the asymmetric unit, indicating that a single molecule of antigen 85 would have weak affinity for a fatty acid at this secondary binding site.

Additionally, the OSG2-binding site is only slightly electronegative (~−10 eV).[2] The carbohydrate binding pocket within the active site has a 6-fold stronger negative electrostatic potential (~−60 eV).[2] Although it seems unlikely this is a specific carbohydrate-binding pocket, the binding at this site seems relatively strong. Thus, the docking role hypothesized for this surface region by Anderson *et al.* (22) cannot be ruled out. Enzymatic assays performed *in vitro*, or experiments *in vivo* with a form of antigen 85C containing mutations within this region could ascertain what role, if any, this surface region plays in catalysis or in proper orientation of the antigen 85 active site toward the cell wall surface.

Antigen 85A Crystal Structure—To determine the structural conservation between the three antigen 85 proteins, the crystal structure of the secreted form of *M. tuberculosis* antigen 85A was solved using antigen 85C as the search model for molecular replacement and subsequent refinement to 2.7-Å resolution. Data collection and refinement statistics are detailed in Table I. Three protein molecules are observed in each asymmetric unit. Superposition of 85A with either 85B or 85C using the Cα atoms from residues 37–282 (35–280 in antigen 85C) results in root mean square deviation values of 0.34 and 0.95 Å, respectively. The catalytic triad required for enzymatic activity is intact within antigen 85A and superimposes well with that of the apo structures of 85B and 85C, and residues found within the substrate binding pocket are structurally conserved between the three antigen 85 proteins, indicating that these proteins do indeed bind the same substrate and catalyze the same mycolyl transfer reaction (Fig. 3*A*).

As mentioned previously the substrate-binding site and active site of the antigen 85 proteins are highly conserved. As seen in the antigen 85C/OSG structure, the active site can be divided into two discrete sections (carbohydrate binding and fatty acid binding) with the serine nucleophile, Ser-124 (Ser-126 in 85A and 85B), marking the point of segregation between these regions. The surface residues forming the carbohydrate binding portion of the substrate-binding site are 100% conserved in all three antigen 85 proteins. In the mycolate-binding portion, Phe-150 and Trp-158 of 85C differ from the corresponding residues in 85A and 85B, leading to the conclusion that these differences in the acyl-binding portion of the substrate binding pocket may be enough to alter substrate specificity. However, as stated previously, this seems unlikely.

Various studies have been carried out to try to determine which region of the antigen 85 proteins is responsible for the specific interaction with human fibronectin. The first research to address this question identified residues 71–90, 121–150, and 201–220 of antigen 85B as being responsible for the interaction with fibronectin (10). A second study found that residues 98–108 (FEWYYQSGLSI) of full-length antigen 85B (before signal sequence cleavage) were essential for binding of fibronectin (25). Neither sequence resembled any previously characterized fibronectin-binding motif. Because these two studies do not agree and no other binding studies have been performed, the residues involved in binding fibronectin are still not known.

We have previously hypothesized that a large, conserved surface patch, composed of residues 1–30, 39–66, and 98–111, is responsible for the interaction between antigen 85 proteins and human fibronectin (2). This surface includes those residues identified by Naito and colleagues (25), 56–66 (FEWYDQSGLSV) of the secreted form of antigen 85A, to be important for binding, which is located on an exposed surface loop connecting helix α2 and strand β4. The conserved patch encompasses about 3271 Å^2, which is ~28% of the total accessible surface area of the protein. Because fibronectin is a very large molecule, ~250 kDa, it is probable that the entirety of this conserved patch is involved in direct interactions with fibronectin. Structures of all three antigen 85 proteins give support to this proposal, because backbone conformation is virtually identical in this region for all three enzymes. Additionally, when these side chain conformations of 85A, B, and C within this region are compared, one can see that only residues Tyr-10 and Lys-23 in the antigen 85A structure deviate from the conformation seen in 85B and 85C (Fig. 3*B*). This is simply a result of differences in torsion angles of these two residues in antigen 85A rather than a change in protein structure. In fact, only one of the three molecules in the asymmetric unit of the antigen 85A structure differs at these two positions. Residues Tyr-10 and Lys-23 of the other two antigen 85A molecules share the same conformation as that seen in the antigen 85B and 85C structures. Furthermore, residues Tyr-10 and Lys-23, which

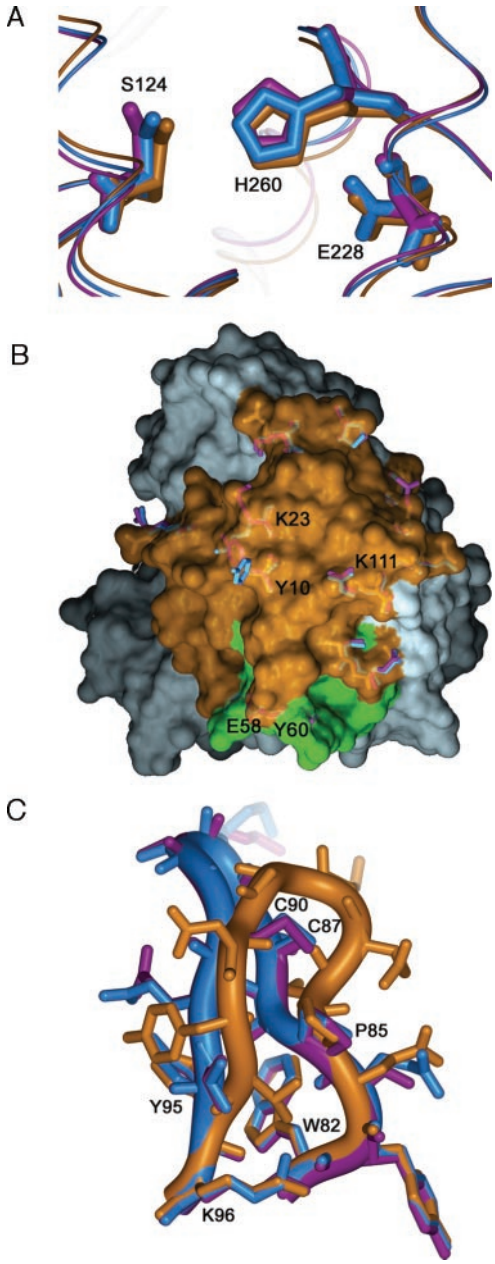

FIG. 3. **Antigen 85 structural comparison.** *A*, superposition of the three antigen 85 structures results in a very good alignment of the antigen 85 catalytic triads. The *numbering* of the residues is that for antigen 85C (in 85A and 85B the numbering is Ser-126, Glu-230, and His-262). The *ribbon* and *atom colors* for 85A, 85B, and 85C are *blue*, *violet*, and *gold*, respectively. *B*, shown is a molecular surface of antigen 85C. The *gold portion* of the surface represents the proposed fibronectin binding surface, and the *green portion* represents the residues shown by Naito *et al.* to be involved in fibronectin interactions. The residues shown in *blue*, *violet*, and *gold* (85A, 85B, and 85C, respectively) are surface residues found in this conserved patch that have altered conformation between the three antigen 85 proteins. Labels for residues mentioned in the text are given. *C*, the *ribbon* and *atom colors* for 85A, 85B, and 85C are *blue*, *violet*, and *gold*, respectively. The *numbering* of the residues is that for antigens 85A and 85B. Shown are residues 82–96, where labels for Trp-82 and Lys-96 mark the edges of the loop. Residues Pro-85 and Tyr-95 are the residues at each extreme end of the loop that occupy a similar position within all three antigen 85 proteins. Cys-87 and Cys-92, which form the only disulfide bond in antigens 85A and 85B, are also labeled.

vary in their conformation, are at the interface of another molecule in the crystal lattice, suggesting that this difference is due to crystal contacts and that the change in conformation of these two residues is an artifact. Other differences on this surface are a result of sequence changes, although these changes result in the conservation of charge or hydrogen bonding potential. Specifically, residue Lys-111 (corresponding to position 113 in antigens 85A and 85B) is an arginine in 85A and 85B, thus the positive charge at this position is maintained. Also, in the region between 56 and 66, which was shown to be necessary for the interaction between the antigen 85 proteins and fibronectin, the W58E mutation in antigen 85C and the Y62D mutation in antigen 85A both allow for the formation of hydrogen bonds that may be important for binding fibronectin. Therefore, because the proposed fibronectin binding region has now been shown to have conserved topology and conserved electrostatics in all three antigen 85 proteins, it seems clear that these residues do represent the point of interaction with fibronectin. Of course, structural studies of an antigen 85/fibronectin complex are required to determine the exact nature of the interactions involved in this important interaction.

Why Three Antigen 85 Enzymes?—Looking closely at the few differences between the antigen 85 proteins there seems to be a selective advantage for *M. tuberculosis* to possess redundant antigen 85 proteins. As mentioned previously, reverse-transcriptase PCR assays of the antigen 85 genes under different growth conditions indicate that mycobacteria adjust antigen 85 expression as a response to the environment (6). Specifically, antigens 85B and 85C were expressed when bacteria were grown in Sauton medium, but antigen 85A is only expressed when the bacteria are inside a macrophage. Furthermore, when macrophages were activated with lipopolysaccharide and interferon-γ, expression of the antigen 85A was turned off while the 85C gene expression was turned on.

To understand the importance of this fact, the slight differences between the antigen 85 proteins must be looked at carefully. For example, the spatial positions of residues 85–91 of the mature antigen 85C structure are noticeably different from the analogous residues conserved between antigens 85A and 85B. As a result, the structure of antigen 85C differs in this region from the structures of antigen 85A and 85B (Fig. 3*C*). Residues 86–95 of antigens 85A and 85B form a short surface-exposed turn between two anti-parallel β-strands. A disulfide bond between residues Cys-87 and Cys-92 forms the base of this short turn. As a result of the conformational restraints placed on this short turn by the presence of the disulfide bond, the backbone conformation is identical between antigens 85A and 85B. In contrast, antigen 85C does not possess a disulfide bond in this region, so no β-sheet is present and this turn contains a noticeable kink. Because this region of antigen 85B is a known T-cell epitope (26), the differences seen here highlight the fact that residues not involved in catalysis or fibronectin binding are not conserved between the three antigen 85 proteins.

To extend this hypothesis, it should be noted that the three antigens possess significant differences in surface-exposed residues within the C-terminal regions. In fact, it seems that almost every surface residue not directly involved in either substrate binding, fibronectin binding, or maintaining the three-dimensional structure (hydrogen bonding, forming salt bridges, etc.) has been mutated in at least one of the antigen 85 proteins. For example, in helix 10, all of the conserved residues face the interior of the protein while none of the solvent-exposed residues are conserved. A similar pattern is also seen in the C-terminal half of helix α9, which encompasses residues 234–248. Of these 15 residues, the conserved residues (Thr-234, Asn-235, Phe-238, and Tyr-242) face the protein interior while the solvent-exposed residues are variable. If a vaccine is to be based on the antigen 85 proteins (27–32), it may be

beneficial to utilize all three antigen 85 proteins, because the redundancy of the antigen 85 proteins promotes a selective advantage for *M. tuberculosis*, possibly through immune system evasion.

CONCLUSION

As the number of people infected with *M. tuberculosis* increases, and the death toll due to tuberculosis continues to rise, research focused on the development of new therapeutics continues to be of utmost importance. The research presented here should further the study in areas of anti-tubercular therapeutics and vaccine design. The co-crystal structure of OSG bound in the active site of antigen 85C illustrates the binding mode of the substrate and extends the knowledge concerning specific protein/substrate interactions. This is the first step toward designing potent inhibitors to the mycolyltransferase activity of the antigen 85 enzymes. Although knowledge of specific protein/carbohydrate interactions was previously known from the antigen 85B trehalose structure (22), information concerning the nature and dimensions of the van der Waals interactions between the protein and acyl chain should allow for the design of strongly binding inhibitors.

The apo structure of the antigen 85A protein, when compared with 85B and 85C, confirms that the active site and substrate-binding site are structurally conserved within all three enzymes. Additionally, this same comparison indisputably confirms the structural conservation of residues near the N terminus proposed to be involved in the interaction of antigen 85 with human fibronectin. Because the fibronectin binding region is important for bacterial adhesion to mucosal surfaces and infection of macrophages, inhibition of this interaction could be important for vaccine design. Antibodies raised against the N-terminal region of the antigen 85 proteins could effectively block the attachment of antigen 85 to fibronectin, thus preventing adhesion, and/or stimulation of phagocytosis. This may eliminate one mode of macrophage entry used by *M. tuberculosis* during the initial stage of infection.

The comparison of the three antigen 85 proteins also illustrates the strikingly high variability of surface residues found over the remaining portion of the protein surface. While the variability of structurally or catalytically unimportant surface residues and the conservation of residues within the core of a protein are not traits specific only to the antigen 85 proteins as divergent evolution dictates that residues not important for maintaining protein structure or function will change over time, this pattern is seen in all homologous proteins. However, the antigen 85 proteins pose an interesting case, because proteins that possess the same enzymatic activity, the same cellular localization, and the same tight binding interaction with human fibronectin have so much variation in the C-terminal surface residues. In such closely related proteins, this seems to indicate that variation of the antigen 85 surfaces is important to the survival of *M. tuberculosis*.

Acknowledgment—We thank Dr. Ken Duncan of GlaxoSmithKline for helpful comments during the preparation of the manuscript.

REFERENCES

1. Belisle, J. T. Vissa, V. D., Sievert, T., Takayama, K., Brennan, P. J., and Besra, G. S. (1997) *Science* **276,** 1420–1422
2. Ronning, D. R., Klabunde, T., Besra, G. S., Vissa, V. D., Belisle, J. T., and Sacchettini, J. C. (2000) *Nat. Struct. Biol.* **7,** 141–146
3. Jackson, M., Raynaud, C., Laneelle, M. A., Guilhot, C., Laurent-Winter, C., Ensergueix, D., Gicquel, B., and Daffe, M. (1999) *Mol. Microbiol.* **31,** 1573–1587
4. Puech, V., Guilhot, C., Perez, E., Tropis, M., Armitige, L. Y., Gicquel, B., and Daffe, M. (2002) *Mol. Microbiol.* **44,** 1109–1122
5. Harth, G., Horwitz, M. A., Tabatadze, D., and Zamecnik, P. C. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 15614–15619
6. Mariani, F., Cappelli, G., Riccardi, G., and Colizzi, V. (2000) *Gene* (*Amst.*) **253,** 281–291
7. Armitige, L. Y., Jagannath, C., Wanger, A. R., and Norris, S. J. (2000) *Infect. Immun.* **68,** 767–778
8. Abou-Zeid, C., Ratliff, T. L., Wiker, H. G., Harboe, M., Bennedsen, J., and Rook, G. A. (1988) *Infect. Immun.* **56,** 3046–3051
9. Naito, M., Fukuda, T., Sekiguchi, K., and Yamada, T. (2000) *Biochem. J.* **347,** 725–731
10. Peake, P., Gooley, A., and Britton, W. J. (1993) *Infect. Immun.* **61,** 4828–4834
11. Schlesinger, L. S., and Horwitz, M. A. (1991) *J. Immunol.* **147,** 1983–1994
12. Talay, S. R., Zock, A., Rohde, M., Molinari, G., Oggioni, M., Pozzi, G., Guzman, C. A., and Chhatwal, G. S. (2001) *Cell. Microbiol.* **2,** 521–535
13. van Putten, J. P., Duensing, T. D., Cole, R. L. (1998) *Mol. Microbiol.* **29,** 369–379
14. Fikrig, E., Feng, W., Barthold, S. W., Telford, S. R., 3rd, and Flavell, R. A. (2000) *J. Immunol.* **164,** 5344–5351
15. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
16. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921
17. Collaborative Computational Project, Number 4 (1994) CCP4 Suite: Programs for protein crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 760–763
18. Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50,** 157–163
19. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491–497
20. Drenth, J. (1994) *Principals of Protein X-ray Crystallography*, p. 71, Springer-Verlag, New York
21. Jones, T. A., Zou, J. Y., Cowtan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
22. Anderson, D. H., Harth, G., Horwitz, M. A., and Eisenberg, D. (2001) *J. Mol. Biol.* **307,** 671–681
23. Deleted in proof
24. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) *Prot. Eng.* **8,** 127–134
25. Naito, M., Ohara, N., Matsumoto, S., and Yamada, T. (1998) *J. Biol. Chem.* **273,** 2905–2909
26. Roche, P. W., Peake, P. W., Billman-Jacobe, H., Doran, T., and Britton, W. J. (1994) *Infect. Immun.* **62,** 5319–5326
27. Horwitz, M. A., Lee, B. W., Dillon, B. J., and Harth, G. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 1530–1534
28. Huygen, K., Content, J., Denis, O., Montgomery, D. L., Yawman, A. M., Deck, R. R., DeWitt, C. M., Orme, I. M., Baldwin, S., D'Souza, C., Drowart, A., Lozes, E., Vandenbussche, P., Van Vooren, J. P., Liu, M. A., Ulmer, J. B. (1996) *Nat. Med.* **2,** 893–898
29. Kamath, A. T., Feng, C. G., Macdonald, M., Briscoe, H., and Britton, W. J. (1999) *Infect. Immun.* **67,** 1702–1707
30. Horwitz, M. A., Harth, G., Dillon, B. J., and Maslesa-Galic, S. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 13853–13858
31. Kamath, A. T., Groat, N. L., Bean, A. G., and Britton, W. J. (2000) *Clin. Exp. Immunol.* **120,** 476–482
32. Miki, K., Nagata, T., Tanaka, T., Kim, Y. H., Uchijima, M., Ohara, N., Nakamura, S., Okada, M., and Koide, Y. (2004) *Infect Immun.* **72,** 2014–2021