THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 32, pp. 21613–21625, August 7, 2009

Substrate-induced Conformational Changes in the Essential Peripheral Membrane-associated Mannosyltransferase PimA from Mycobacteria

IMPLICATIONS FOR CATALYSIS[*] [S]

Received for publication, April 3, 2009, and in revised form, June 3, 2009 Published, JBC Papers in Press, June 11, 2009, DOI 10.1074/jbc.M109.003947

Marcelo E. Guerin[‡1,2], **Francis Schaeffer**[§1], **Alain Chaffotte**[¶], **Petra Gest**[‡], **David Giganti**[§], **Jana Korduláková**[‖], **Mark van der Woerd**[**], **Mary Jackson**[‡], **and Pedro M. Alzari**[§]

From the [‡]*Department of Microbiology, Immunology, and Pathology and* [**]*Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado 80523-1682,* [§]*Unité de Biochimie Structurale (CNRS URA 2185) and* [¶]*Unité de RMN des Biomolécules, Institut Pasteur, 25 rue du Dr. Roux, 75724 Paris Cedex 15, France, and the* [‖]*Department of Biochemistry, Faculty of Natural Sciences, Comenius University, Mlynska dolina, 84215 Bratislava, Slovakia*

Phosphatidyl-*myo*-inositol mannosyltransferase A (PimA) is an essential glycosyltransferase (GT) involved in the biosynthesis of phosphatidyl-*myo*-inositol mannosides (PIMs), which are key components of the mycobacterial cell envelope. PimA is the paradigm of a large family of peripheral membrane-binding GTs for which the molecular mechanism of substrate/membrane recognition and catalysis is still unknown. Strong evidence is provided showing that PimA undergoes significant conformational changes upon substrate binding. Specifically, the binding of the donor GDP-Man triggered an important interdomain rearrangement that stabilized the enzyme and generated the binding site for the acceptor substrate, phosphatidyl-*myo*-inositol (PI). The interaction of PimA with the β-phosphate of GDP-Man was essential for this conformational change to occur. In contrast, binding of PI had the opposite effect, inducing the formation of a more relaxed complex with PimA. Interestingly, GDP-Man stabilized and PI destabilized PimA by a similar enthalpic amount, suggesting that they formed or disrupted an equivalent number of interactions within the PimA complexes. Furthermore, molecular docking and site-directed mutagenesis experiments provided novel insights into the architecture of the *myo*-inositol 1-phosphate binding site and the involvement of an essential amphiphatic α-helix in membrane binding. Altogether, our experimental data support a model wherein the flexibility and conformational transitions confer the adaptability of PimA to the donor and acceptor substrates, which seems to be of importance during catalysis. The proposed mechanism has implications for the comprehension of the peripheral membrane-binding GTs at the molecular level.

Glycans are not only one of the major components of the cell but also are essential molecules that modulate a variety of important biological processes in all living organisms. Glycans are used primarily as energy storage and metabolic intermediates as well as being main structural constituents in bacteria and plants. Moreover, as a consequence of protein and lipid glycosylation, glycans generate a significant amount of structural diversity in biological systems. This structural information is particularly apparent in molecular recognition events including cell-cell interactions during critical steps of development, the immune response, host-pathogen interactions, and tumor cell metastasis. Most of the enzymes encoded in eukaryotic/prokaryotic/archaeans genomes that are responsible for the biosynthesis and modification of glycan structures are GTs[3] (1). Here we have focused in the phosphatidyl-*myo*-inositol mannosyltransferase A (PimA), an essential enzyme of mycobacterial growth that initiates the biosynthetic pathway of key structural elements and virulence factors of *Mycobacterium tuberculosis*, the phosphatidyl-*myo*-inositol mannosides (PIM) lipomannan and lipoarabinomannan (2–5). This amphitropic enzyme catalyzes the transfer of a Man*p* residue from GDP-Man to the 2-position of PI to form phosphatidyl-*myo*-inositol monomannoside (PIM_1) on the cytoplasmic side of the plasma membrane (2) (Fig. 1).

Although considerable progress has been made in recent years in understanding the mode of action of GTs at the molecular level, the mechanisms that govern recognition of lipid acceptors and membrane association of peripheral membrane-binding GTs remains poorly understood. GTs can be classified as either "inverting" or "retaining" enzymes according to the anomeric configuration of the reaction substrates and products. A single displacement mechanism in which a general base assists in the activation of the acceptor substrate for nucleophilic attack by the sugar donor is well

* This work was supported, in whole or in part, by National Institutes of Health Grant AI064798 from NIAID. This work was also supported by the Infectious Disease SuperCluster at Colorado State University, Institut Pasteur Program GPH-5, Contract 01-B-0095 from the Ministère de la Recherche, France, and European Commission Contract LSHP-CT-2005-018923 (New Medicines for Tuberculosis).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1S–4S and Tables 1S and 2S.

[1] Both authors contributed equally to this work.

[2] To whom correspondence should be addressed. Tel.: 970-491-4067; Fax: 970-491-1815; E-mail: mrcguerin@gmail.com.

[3] The abbreviations used are: GT, glycosyltransferase; PIM, phosphatidyl-*myo*-inositol mannoside; PI, phosphatidyl-*myo*-inositol; PIM_1, phosphatidyl-*myo*-inositol monomannoside; *Ms*PimA, PimA from *Mycobacterium smegmatis*; ITC, isothermal titration calorimetry; DSC, differential scanning calorimetry; AUC, analytical ultracentrifugation; Ins-P, *myo*-inositol 1-phosphate; PDB, Protein Data Bank; r.m.s.d., root-mean-square deviation.

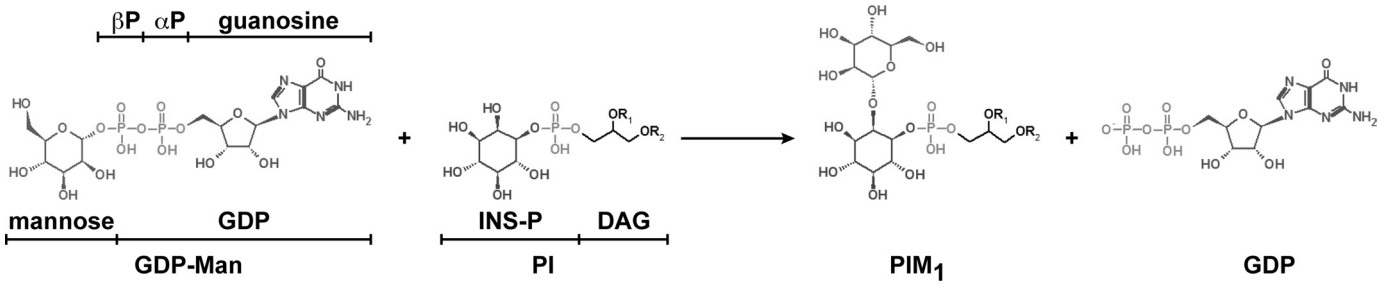

FIGURE 1. **PIM_1 biosynthesis in mycobacteria.** PimA transfers a Man*p* residue from GDP-Man to the 2-position of the *myo*-inositol ring of PI to form PIM_1 (where *DAG* is di-acyl-glycerol, and *INS-P* is 1-L-*myo*-inositol phosphate). The reaction occurs with retention of the anomeric configuration of the sugar donor.

established for inverting enzymes (6, 7). In contrast, the catalytic mechanism for retaining enzymes, including PimA, remains unclear. By analogy with glycosylhydrolases, a double displacement mechanism via the formation of a covalent glycosyl-enzyme intermediate was first proposed (8). However, in the absence of direct evidence of a viable covalent intermediate, an alternative mechanism known as the S_Ni "internal return" has been suggested where phospho-sugar donor bond breakage and sugar-acceptor bond formation occur in a concerted, but necessarily stepwise manner on the same face of the sugar (6, 9). Only two protein topologies have been found for nucleotide-diphospho-sugar-dependent enzymes among the first 30 GT sequence-based families (10) (see the carbohydrate-active enzymes (CAZy) data base) for which three-dimensional structures have been reported (11). These topologies are variations of "Rossmann-like" domains and have been defined as GT-A (12) and GT-B (13). Both inverting and retaining enzymes were found in GT-A and GT-B folds, indicating that there is no correlation between the overall fold of GTs and their catalytic mechanism. The primary sequence of PimA contains the GPGTF (glycogen phosphorylase/GT) motif, a signature present in enzymes of the GT-B fold (14). GT-B proteins do not use divalent cations and consist of two Rossmann-like (β-α-β) domains separated by a deep fissure. Therefore, an important interdomain movement has been predicted in some members of this superfamily during catalysis, including MurG (15), glycogen synthase (16),and the *myo*-inositol 1-phosphate *N*-acetylglucosaminyltransferase MshA (17).

To perform their biochemical functions, membrane-binding GTs interact with membranes by two different mechanisms. Whereas integral membrane GTs are permanently attached through transmembrane regions (*e.g.* hydrophobic α-helices) (18) peripheral membrane-binding GTs temporarily bind membranes by (i) a stretch of hydrophobic residues exposed to bulk solvent, (ii) electropositive surface patches that interact with acidic phospholipids (*e.g.* amphipathic α-helices), and/or (iii) protein-protein interactions (19–22). A close interaction of the enzyme with membranes might be a strict requirement for PI modification by PimA. We recently solved the crystal structure of PimA from *Mycobacterium smegmatis* (*Ms*PimA) in complex with the donor substrate GDP-Man (23, 24). The notion of a membrane-associated protein via electrostatic interactions is consistent with the finding of an amphipathic α-helix and surface-exposed hydrophobic residues in the N-terminal domain of *Ms*PimA. Despite the fact that sugar transfer is catalyzed between the mannosyl group of the GDP-Man donor and the *myo*-inositol ring of PI, the enzyme displays an absolute requirement for both fatty acid chains of PI in order for the transfer reaction to take place. Furthermore, PimA was able to bind monodisperse PI, but its transferase activity was stimulated by high concentrations of nonsubstrate anionic surfactants, indicating that the reaction requires a lipid-water interface. We thus proposed a model of interfacial catalysis in which PimA recognizes the fully acylated substrate PI with its polar head within the catalytic cleft and the fatty acid moieties only partially sequestered from the bulk solvent (24).

This study describes a detailed investigation of the lipid acceptor binding site and the conformational properties of PimA in solution. Using a combination of limited proteolysis, isothermal titration calorimetry (ITC), differential scanning calorimetry (DSC), circular dichroism (CD), analytical ultracentrifugation (AUC) and site-directed mutagenesis, we propose a plausible model for substrates recognition and binding. The implications of this model for the comprehension of the early steps of PIM biosynthesis and the catalytic mechanism of other members of the peripheral membrane-binding GT family are discussed.

EXPERIMENTAL PROCEDURES

Methods—The *pimA* gene from *M. smegmatis* (MSMEG_2935, Rv2610c, genolist.pasteur.fr/TubercuList/) was subcloned from pQE70-*MspimA* DNA by standard PCR using oligonucleotide primers pimA_NdeI_Fwd (5′-GGGAATTC-CATATGCGTATCGGGATGGTCTGCC-3′) and pimA_BamHI_Rev (5′-CGCGGATCCTCAGTGATGGTGATGGT-GATG-3′) with Phusion DNA polymerase (New England Biolabs). The PCR fragment was digested with NdeI and BamHI and ligated to the expression vector pET29a (Novagen) generating pET29a-*MspimA*. The recombinant *Ms*PimA (378 residues) has an additional peptide of 8 amino acids (^{379}RSHHH-HHH^{386}) at the C terminus that includes a histidine tag. *Ms*PimA and *Ms*PimA mutants were purified to apparent homogeneity by a combination of metal ion affinity, anionic exchange, and gel filtration chromatography steps as described previously (23). The enzymatic activity of *Ms*PimA and *Ms*PimA mutants was monitored using a radiometric assay. Briefly, the reaction mixture contained 0.0625 μCi of GDP-[C^{14}]Man (specific activity, 305 mCi mmol^{-1}; Amersham Biosciences), membrane preparations from *M. smegmatis* mc^{2}155

(0.5 mg of proteins) as the source of lipid acceptors, 50 μg of purified *Ms*PimA or *Ms*PimA mutants, and 50 mM Tris-HCl buffer, pH 7.5, in a final volume of 250 μl. Reactions were incubated for 2 h at 37 °C and stopped with 1.5 ml of $CHCl_3/CH_3OH$ (2:1 (v/v)). The PIM-containing organic phase was prepared and analyzed by TLC as described by Korduláková *et al.* (2). The bands recorded by a PhosphorImager Typhoon TRIO (GE Healthcare) were quantified using volume integration with the ImageQuant TL v2005 program (GE Healthcare).

Gel Filtration—Gel filtration chromatography was performed using a BioSuite 250 5-μm HR SEC column (Waters Corp.) equilibrated in 50 mM Tris-HCl, pH 6.8, and 150 mM NaCl at 1 ml min^{-1}. The column was previously calibrated using gel filtration standards (Sigma) including β-amylase (200 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa), carbonic anhydrase (29 kDa), and cytochrome *c* (12.4 kDa).

Limited Proteolysis of MsPimA—Recombinant purified *Ms*PimA (25 μg) was incubated with 0.05 μg of elastase (Sigma) in 100 μl of 50 mM Tris-HCl, pH 7.6, in the presence of 1 mM guanosine, 1 mM GDP (Sigma), 1 mM GDP-Man (Sigma), or 1.25 mM PI (Sigma) for 0–90 min at 37 °C. Aliquots of 12 μl were mixed with 3 μl of 250 mM Tris-HCl, pH 6.8, 10% SDS, 50% glycerol, 500 mM dithiothreitol, and 0.01% bromphenol blue at the indicated times. Samples were boiled for 3 min and run onto an SDS-polyacrylamide gel. Protein bands were visualized by staining the gel with Coomassie Brilliant Blue R-250.

N-terminal Sequence Analyses—Proteolytic fragments were electrotransferred to Hybond-P polyvinylidene difluoride (Amersham Biosciences) in 50 mM Trizma base and 50 mM boric acid buffer for 16 h at room temperature. Bands were stained with 0.1% Coomassie Brilliant Blue R-250 and subjected to N-terminal sequence analysis using an AABI494 protein sequencer (Applied Biosystems).

Isothermal Titration Calorimetry—Ligand binding to *Ms*PimA was assayed using the high precision VP-ITC system (MicroCal Inc.) as described previously (24, 25) with the following modifications. The ITC cell (1.4 ml) contained 10 μM *Ms*PimA in 25 mM Tris-HCl, pH 7.6, 150 mM NaCl, 5% Me_2SO, and the syringe (300 μl) contained 150 μM GDP-Man, 150 μM GDP, or 250 μM guanosine in the same buffer. Binding of PI aggregates to *Ms*PimA-GDP and mutant *Ms*PimA$^{R77E/K78E/K80E/K81E}$-GDP complexes was assayed as follows. The protein was first titrated with GDP, and at the end of the titration, the solution containing the protein complex was kept in the ITC cell and titrated with 1.25 mM solution of PI in the same buffer preparation. Sample solutions were thoroughly degassed under vacuum, and each titration was performed at 25 °C by one injection of 2 μl followed by 29 injections of 10 μl, with 210 s between injections, using a 290 rpm rotating syringe. Raw heat signal collected with a 16-s filter was corrected for the dilution heat of the ligand in the *Ms*PimA buffer and normalized to the concentration of ligand injected. Data were fit to a bimolecular model (26) using the Origin™ software provided by the manufacturer.

Differential Scanning Calorimetry—Thermal unfolding of *Ms*PimA and of the *Ms*PimA complexes was quantified by DSC using the high precision VP-DSC system (MicroCal Inc.) as described previously (27). Sample solutions contained 9.5 μM *Ms*PimA alone or mixed with one of the following ligands at a saturating concentration determined by ITC: GDP-Man, 50 μM; GDP, 30 μM; PI, 150 μM (24); or guanosine, 250 μM (this study). One protein sample contained both GDP and PI added in this order at the same concentrations as above. Protein and ligands were dissolved in the same buffer preparation (50 mM sodium phosphate, pH 7.0, 0.2% Me_2SO) and degassed under vacuum for 10 min with gentle stirring prior to being loaded into the calorimetric cell (0.5 ml). Sample solutions were held *in situ* under a constant external pressure of 25 p.s.i., equilibrated for 30 min at 25 °C, and then heated at a constant rate of 1 °C min^{-1}. Experimental heat capacity functions were collected with a 16-s filter. Data normalization and quantification were carried out with the Origin7 software (28) provided by the manufacturer. The partial molar heat capacity (C_p) was derived from the experimental heat capacity by subtraction of the instrumental base line. The partial molar excess heat capacity ($C_{p,ex}$) was derived from the C_p function by subtraction of the molar intrinsic heat capacity ($C_{p,int}$). $C_{p,int}$ was calculated as an interpolation over the transition region of the folded and unfolded states weighted in proportion to their relative contributions. For free *Ms*PimA, the partial molar heat capacity of the native state was obtained by linear extrapolation of the pretransition base line. The partial molar heat capacity of the unfolded state was calculated by summing up the heat capacities of the amino acid residues constituting the polypeptide chain using the known quadratic dependence of C_p of the unfolded state with temperature (29, 30) (*cf.* Fig. 3*A*, *inset*). Fitting and deconvolution of $C_{p,ex}$ required the use of a non-two-state transition model.

Calculation of MsPimA Intrinsic Unfolding Enthalpy—On free *Ms*PimA unfolding, the observed unfolding enthalpy (ΔH_u) measured as the surface under the heat absorption peak ($C_{p,ex}$) of the DSC endotherm is the intrinsic unfolding enthalpy of the protein (ΔH_u (free) = $\Delta H_{u,int}$ (free)). On unfolding of a *Ms*PImA-substrate complex, the observed unfolding enthalpy equals the intrinsic unfolding enthalpy of the protein inside the protein complex minus the substrate binding enthalpy,

$$\Delta H_u(\text{complex}) = \Delta H_{u,\text{int}}(\text{complex}) - \Delta H_{\text{bind}} \quad \text{(Eq. 1)}$$

The change in $\Delta H_{u,int}$ from the free to the bound state of *Ms*PimA then equals

$$\Delta\Delta H_{u,\text{int}} = \Delta H_u(\text{complex}) - \Delta H_u(\text{free}) + \Delta H_{\text{bind}} \quad \text{(Eq. 2)}$$

Values of ΔH_{bind} for the binding of GDP-Man and GDP to free *Ms*PimA were measured by ITC (Fig. 2*D*). Values of ΔH_{bind} for the binding of PI to free *Ms*PimA ($\Delta H_{bind} = -1.3$ kcal/mol) and to the *Ms*PimA-GDP complex ($\Delta H_{bind} = 13.6$ kcal mol^{-1}) were taken from Guerin *et al.* (24). From repeated experiments, the value of ΔH_{bind} is ± 0.2 kcal/mol.

Circular Dichroism—*Ms*PimA thermal unfolding was determined by recording the ellipticity change at 222 nm as a function of the temperature. An Aviv 215 spectropolarimeter was used. The protein sample at 10 μM in 10 mM potassium phosphate, pH 7.0, with or without ligand(s), was placed in a rectangular fused silica cuvette of 0.2-cm path length. The cuvette was

inserted in the cell holder with a Peltier temperature control. The sample was heated with a slope of 1 °C min^{-1}, and ellipticity was recorded each 0.5 °C after an equilibration time of 15 s. The CD signal was averaged over 10 s. Each set of data collected was then converted to $\Delta q/\Delta T$ using the first derivative converter available in the software of the CD instrument with a window of 11 data points and a second degree polynomial smoothing. The observed difference in $T_{½}$ between the CD and DSC melting curves was shown to result from the difference in the phosphate buffer concentration (data not shown).

Analytical Ultracentrifugation—AUC experiments were performed with a Beckman XL-1 analytical ultracentrifuge using absorbance optics. Velocity measurements utilized a two-sector charcoal-filled Epon centerpieces, quartz windows, 400-μl sample, and 420-μl reference (buffer: 50 mM Tris-HCl, pH 8.0, 150 mM NaCl) volumes. All samples were centrifuged in a Beckman 8-hole An50Ti rotor at 22 °C at 40,000 rpm, and the data were collected at 280 nm with a radial increment of 0.003 cm for ~6 h. Velocity data were edited and analyzed using the boundary analysis method of Demeler and van Holde as implemented in Ultrascan, version 7.3 for Windows (31, 32). Sedimentation coefficients (s) are reported in Svedberg units (S), where $1\ S = 1 \times 10^{-13}\ s$, and were corrected to that of water at 20 °C ($s_{20,w}$). The partial specific volume of full-length *Ms*PimA was calculated from the amino acid sequence within Ultrascan. Modeling of hydrodynamic parameters was performed using Ultrascan. The frictional ratio (f/f_0) was calculated from the known molecular mass and the measured sedimentation coefficient using Ultrascan.

Structural Alignment—Structural alignments of *Ms*PimA (PDB code 1GEJ) with *Cg*MshA (PDB code 3C4V), WaaG (PDB code 2IV7), WaaF (PDB code 1PSW), WaaC (PDB code 2H1F), MurG (PDB code 1NLM) and GumK (PDB code 3CV3) were performed by the distance alignment matrix method using DALI Lite (33). The images were generated with PyMOL, version 0.99 (34).

Molecular Docking Calculations—We used the program ICM, version 3.6 (35), to dock Ins-P to the structure of the *Ms*PimA-GDP-Man complex (PDB code 2GEJ). Interaction grids were calculated in the acceptor binding site after the preliminary selection of key residues chosen within 6 Å of the reactive axial oxygen of the mannose moiety. Ins-P and GDP-Man input files were extracted from the PubChem data base (pubchem.ncbi.nlm.nih.gov). Both structures were protonated and optimized, and all water molecules were removed from the receptor using standard ICM protocols. The thoroughness parameter was set to a value of 2 to increase the Monte Carlo steps and ensure a greatest exploration. The best solution showed a predicted energy below 60 kcal M^{-1} and was used for structural analysis.

Site-directed Mutagenesis—*M. smegmatis pimA* was subcloned in the pET29a expression vector (Novagen) using the NdeI and BamHI restriction sites, allowing for the production of a C-terminal His_6-tagged fusion protein. The mutants *Ms*PimA^{Q18A}, *Ms*PimA^{Y62A}, *Ms*PimA^{N63A}, *Ms*PimA^{S65A}, *Ms*PimA^{R68A}, *Ms*PimA^{R70A}, *Ms*PimA^{N97A}, *Ms*PimA^{T119A}, *Ms*PimA^{K123A}, *Ms*PimA^{R196A}, *Ms*PimA^{E199A}, and *Ms*PimA$^{R77E/K78E/K80E/K81E}$ were generated by the two-step PCR overlap method using Phusion high-fidelity DNA polymerase (New England Biolabs) and pET29a-*MspimA* as the DNA template. The flanking oligonucleotides that annealed with the NdeI and BamHI sites were pimA_NdeI_Fwd and pimA_BamHI_Rev, respectively. The overlapping oligonucleotides for each mutant are described in supplemental Table 1S. The 1100-bp fragments obtained were digested with NdeI/BamHI and ligated to pET29a treated with the same enzymes. All plasmids were sequenced at the Proteomics and Metabolomics Facility at Colorado State University.

Functionality of MsPimA$^{R77E/K78E/K80E/K81E}$ Mutant in M. smegmatis—A two-step homologous recombination procedure was used earlier to provide evidence of the essentiality of *pimA* in *M. smegmatis* (2). In the course of this experiment, strains having undergone a single crossover at the *pimA* locus were easily isolated, but it was shown that a second crossover leading to gene knock-out was achievable only in the presence of a rescue wild-type copy of *pimA* carried by a replicative plasmid. To determine whether the *MsPimA*$^{R77E/K78E/K80E/K81E}$ mutant encoded an active enzyme in mycobacteria, a mycobacteria/*Escherichia coli* shuttle plasmid (pVV16) expressing this gene was tested for its ability to serve as a rescue plasmid in the gene knock-out experiment. A pVV16 construct carrying a wild-type copy of the *pimA* gene from *M. smegmatis* was used as a positive control, and the empty pVV16 plasmid served as the negative control in this experiment (2).

RESULTS AND DISCUSSION

Conformational Flexibility Studies Using Limited Proteolysis—To study the effect of substrate binding on the conformation of PimA, we first determined the oligomeric state of *Ms*PimA in solution by subjecting the purified protein to size exclusion chromatography (supplemental Fig. 1*S*). Results showed that *Ms*PimA is a monomer with no sign of higher order oligomers in the purified preparation. The monomeric state of the full-length protein was independent of the presence of substrates (data not shown).

Limited proteolysis has proven to be a powerful technique to study ligand-induced conformational changes in proteins (36). In the case of *Ms*PimA, a monomeric globular protein, the reaction is expected to occur primarily at flexible exposed loops. However, the presence of ligands can have considerably effect on the susceptibility of the protein to the protease. When incubated with elastase, *Ms*PimA was rapidly degraded (Fig. 2*A*). Microsequencing of the two predominant small species of 23 and 15 kDa revealed the sequences SAMRS, located in α9 (Fig. 2*B* and supplemental Fig. 2*S*) and SFADA, in the connecting loop β7-β8 at the junction between N- and C-terminal domains (Fig. 2*B*). It is worth noting that helix α9 contains two critical residues involved in donor substrate recognition: Asp^{253}, which interacts with N_2 of the guanidyl group of GDP-Man, conferring *Ms*PimA its specificity for the nucleoside; and Lys^{256}, which participates in ribose binding. Similar proteolysis experiments performed in the presence of GDP or GDP-Man suggest, in contrast, a conformational rearrangement, as *Ms*PimA was protected from the action of elastase even after 90 min incubation (Fig. 2*A*). The crystal structure of the *Ms*PimA-GDP-Man complex revealed that the enzyme crystallizes in a

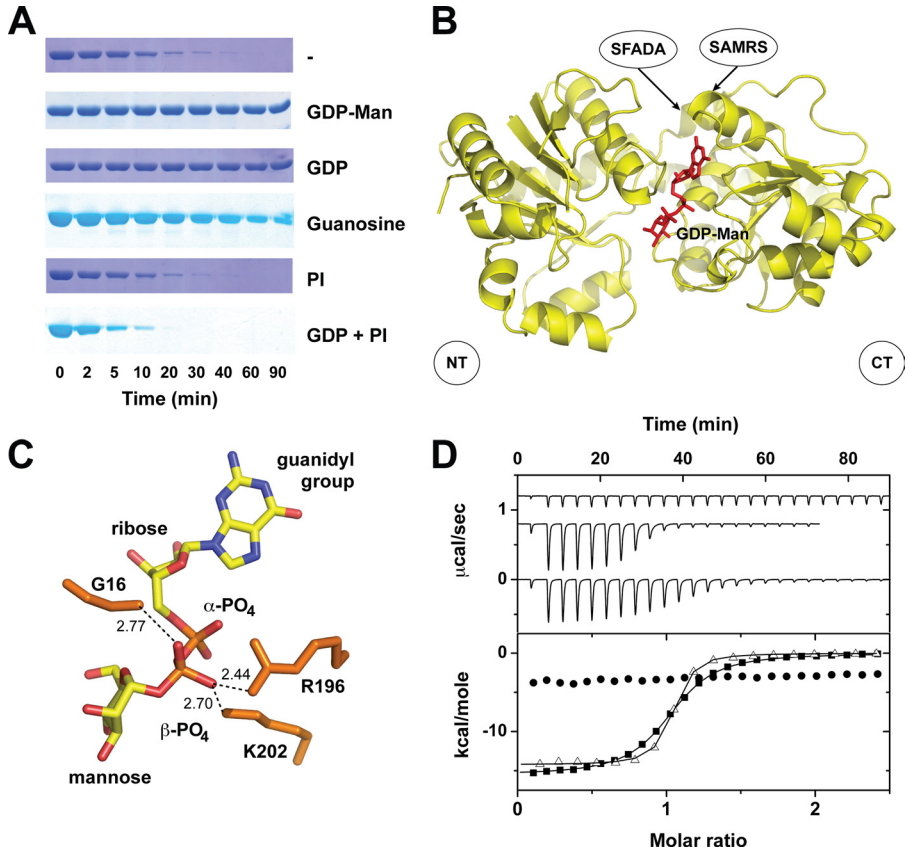

FIGURE 2. **Substrate-induced conformational changes in *Ms*PimA.** *A*, elastase cleavage of *Ms*PimA preincubated with different ligands. *B*, surface representation of *Ms*PimA three-dimensional structure in a closed conformation. The GDP-Man is buried at the N- (*NT*) and C-terminal (*CT*) domains interface. The N-terminal residues of selected proteolytic fragments are indicated. *C*, *Ms*PimA residues Gly^{16} (N-terminal) and Arg^{196}/Lys^{202} (C-terminal) involved in the β-PO_4 stacking interaction. *D*, ITC measurements of *Ms*PimA-ligand interactions. The *upper panel* shows the raw heat signal of the titration of *Ms*PimA with GDP-Man (■), GDP (△), and guanosine (●). The *lower panel* shows the integrated heats of injections of the above titrations corrected for the ligand heat of dilution and normalized to the ligand concentration. *Solid lines* correspond to the best fit of data using a bimolecular interaction model. GDP-Man and GDP bind to *Ms*PimA with a 1:1 stoichiometry, dissociation constants (K_d) of 0.23 and 0.04 μM, and heats of binding (ΔH_{bind}) of -15.6 kcal mol^{-1} and -14.2 kcal mol^{-1}, respectively. Guanosine binds with a weak binding affinity generating a linear variation of ΔH_{bind} with guanosine concentration, $K_a \leq 10^4\ M^{-1}$ and $-\Delta H_{bind} \leq 4$ kcal mol^{-1}.

"closed" conformation with its active site located in a deep cleft formed between the N- and C-terminal domains (Ref. 24, Fig. 2*B*). The GDP-Man binding site resides mainly in the C-terminal domain, where it makes a number of hydrogen bonds with the protein. However, two residues of the N-terminal domain were found to interact with GDP-Man: Pro^{14}, in the connecting loop $\beta1$-$\alpha1$ stabilizing the guanine heterocycle by a van der Waals stacking interaction; and Gly^{16}, at the beginning of $\alpha1$, establishing a hydrogen bond with the β-PO_4 of GDP (Fig. 2*C*). Gly^{16} is located in the Gly-Gly loop, a conserved nucleoside-diphospho-sugar binding motif that has been found to be essential for enzymatic activity in other GT-B enzymes (37). Interestingly, two other residues, Arg^{196} and Lys^{202}, located in the C-terminal domain, also make electrostatic interactions with the β-PO_4, thereby restricting its position in the active site of *Ms*PimA (Fig. 2*C*). It is worth mentioning that the α-PO_4 of GDP-Man does not interact with any particular residue from the enzyme (24). To determine the relevance of the β-PO_4 in the stabilization of the closed conformation, *Ms*PimA was incubated with elastase in the presence of guanosine, in which the α-PO_4 and β-PO_4 of GDP are missing. As shown in Fig. 2*A*, guanosine only partially protected *Ms*PimA from degradation by the protease. Moreover, ITC measurements revealed that GDP-Man and GDP bind to *Ms*PimA with high association constants and favorable binding enthalpies (Fig. 2*D*, Ref. 24). In contrast, guanosine binds to *Ms*PimA with a binding constant $\sim10^3$-fold smaller than that of GDP and with a ~3-fold reduced binding enthalpy (Fig. 2*D*). This observation was confirmed by the small differences in the unfolding parameters of the protein in a guanosine-saturated solution as measured by DSC (supplemental Fig. 3*S*). These results clearly demonstrate that the interactions between *Ms*PimA and the β-PO_4 of GDP-Man are essential for generating a closed state, which is a critical event for maintaining a functional active site.

The effect of PI binding to *Ms*PimA and the *Ms*PimA-GDP complex was analyzed by proteolysis experiments. Surprisingly, in both cases, *Ms*PimA became highly sensitive to elastase, indicating that PI triggers a yet significant conformational change that modifies the closed GDP-Man-induced conformation (Fig. 2*A*).

Unfolding Thermodynamics of Free and Substrate-bound MsPimA—To further characterize the effect of substrate binding on *Ms*PimA conformation, the thermal stabilities of the free and bound protein were compared in solution by DSC and CD. The temperature dependence of the partial molar heat capacity (C_p) of free *Ms*PimA revealed a significant heat absorption peak occurring at a low temperature ($T_{½}$ = 46.5 °C; Fig. 3*A* and Table 1). C_p for the native state was linear with temperature ($dC_p/dT \sim 0.27$ kcal K^{-1} mol^{-1}) as expected (29, 30). As shown in Fig. 3*A*, at high temperatures the C_p of *Ms*PimA approaches the value of the completely unfolded state calculated from the known sequence, assuming that all amino acid residues are exposed to water (29). Because the partial heat capacity of a protein is a very sensitive indicator of the exposure of protein groups to water (29, 30), it can be concluded that upon completion of the large heat absorption peak, *Ms*PimA is completely unfolded. Thus, *Ms*PimA unfolds with a small total heat capacity increment (ΔC_p(60 °C) = 2.4 kcal K^{-1} mol^{-1} at 60 °C) (Fig. 3*A*, Ref. 38).

Both GDP-Man and PI induced large and opposite effects on the C_p curve (Fig. 3*A*) upon binding to *Ms*PimA. Whereas GDP-Man increases the thermal stability of the protein, as well as its cooperativity of unfolding, PI acts in a reversed way (compare

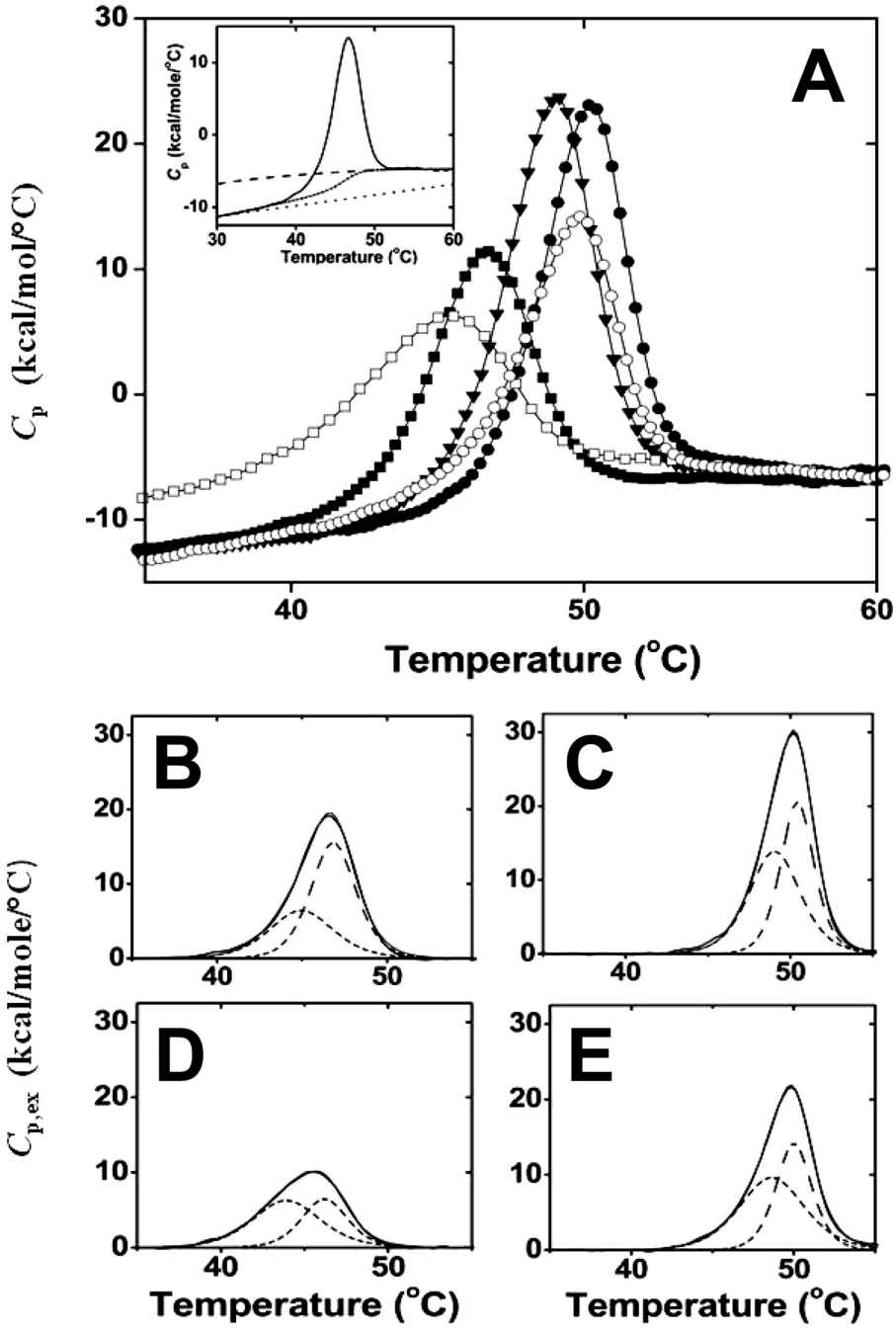

FIGURE 3. **Calorimetric study of the temperature-induced unfolding of *Ms*PimA and of *Ms*PimA-substrate complexes.** *A*, temperature dependence of the partial molar heat capacity (C_p) of *Ms*PimA (■) and *Ms*PimA complexes with PI (□), GDP-Man (▼), GDP (●), and GDP and PI (○). *Inset*, the intrinsic molar heat capacity ($C_{p,\mathrm{int}}$) of *Ms*PimA (■) connecting the C_p function of the native state (····) and the calculated C_p function of the unfolded state (-----). *B–D*, deconvolution of the molar excess heat capacity ($C_{p,\mathrm{ex}}$) of *Ms*PimA (*B*) and *Ms*PimA complexes with GDP (*C*), PI (*D*), and GDP and PI (*E*) using a non-two-state model. The experimental (—) and calculated (– – –) $C_{p,\mathrm{ex}}$ functions of free and bound *Ms*PimA are shown together with the $C_{p,\mathrm{ex}}$ functions of the low temperature (-----) and high temperature (– – –) unfolding transitions.

the temperatures and the peak heights and widths of the various molecular species in Fig. 3*A*). The same opposing trends were observed upon binding of product GDP to free *Ms*PimA or PI to the *Ms*PimA-GDP complex, respectively. Differences in the C_p curve of *Ms*PimA could be quantified easily by considering the molar excess heat capacity function ($C_{p,\mathrm{ex}}$) (Fig. 3, *B–D*). The $C_{p,\mathrm{ex}}$ curve of free *Ms*PimA is asymmetric with a gradual increase in heat uptake at low temperature followed by a sharper transition peak at high temperature, consistent with a complex unfolding process (Fig. 3*B*). Calculation of $C_{p,\mathrm{ex}}$ showed two overlapping unfolding transitions separated by 1.9 °C with a 1.6-fold difference in unfolding enthalpy, in agreement with the unfolding of the two domains of *Ms*PimA (Fig. 3*B* and Table 1). The binding of GDP-Man (and GDP) shifts the unfolding of the two domains toward higher temperatures while largely increasing the unfolding enthalpies (*e.g.* T_1 increased by 3.9 °C, whereas $\Delta H_{u,1}$ increased by 23.5 kcal mol^{-1} on GDP binding; Table 1). Conversely, PI binding to free *Ms*PimA (and to the *Ms*PimA-GDP complex) shifted the unfolding of the two domains toward lower temperatures while largely decreasing the unfolding enthalpies (*e.g.* T_2 decreased by 0.7 °C while $\Delta H_{u,2}$ decreased by 29.5 kcal mol^{-1}; Table 1).

The simultaneous changes of the two unfolding transitions show that the two domains of *Ms*PimA unfold cooperatively. However, a comparison of the enthalpy changes upon GDP and PI binding clearly shows that although both ligands affect the unfolding enthalpy of the entire protein, GDP predominantly interacts with the low temperature (and PI with the high temperature) melting domain of *Ms*PimA ($\Delta\Delta H_{u,1}$ = 94% and $\Delta\Delta H_{u,2}$ = −97% of ΔH_u of the GDP and PI complexes, respectively; Table 1). GDP-Man increases, almost evenly, the unfolding enthalpies of the two transitions.

Taken together, the large differences between the unfolding cooperativity and unfolding enthalpy of free and bound *Ms*PimA clearly demonstrate large conformational changes in the protein upon substrate binding. The importance of these conformational changes can be estimated in terms of energy. The difference in *Ms*PimA intrinsic unfolding enthalpy between the free and substrate-bound states ($\Delta\Delta H_{u,\mathrm{int}}$, Equation 2) reflects the number of non-covalent interactions lost or gained by the protein on substrate binding. $\Delta\Delta H_{u,\mathrm{int}}$ was obtained by subtracting the binding enthalpy (ΔH_{bind} (24) from the unfolding enthalpy difference (between the bound and free states of the protein (Table 1). Values of $\Delta\Delta H_{u,\mathrm{int}}$ are large as a result of GDP-Man and PI

TABLE 1

Unfolding parameters of free and bound *Ms*PimA measured by DSC

T_1 and T_2 are temperatures at half denaturation.

Ligand	T_1	$DH_{u,1}$	$DDH_{u,1}$[a]	T_2	$\Delta H_{u,2}$	$\Delta DH_{u,2}$[a]	DH_u[b]	$\Delta DH_{u,int}$[c]
	°C	*kcal mol^{-1}*	*%*	*°C*	*kcal mol^{-1}*	*%*	*kcal mol^{-1}*	*kcal mol^{-1}*
	45.0 ± 0.2	33.8 ± 0.7		46.9 ± 0.2	53.1 ± 0.7		86.9 ± 1.4	
GDP-Man	47.9 ± 0.1	55.6 ± 0.5	58.4 ± 6.9	49.2 ± 0.1	68.6 ± 0.5	41.6 ± 5.9	124.2 ± 1.0	21.7 ± 2.6
GDP	48.9 ± 0.1	57.3 ± 0.5	94.4 ± 5.9	50.4 ± 0.1	54.5 ± 0.5	5.6 ± 4.8	111.8 ± 1.0	10.7 ± 2.6
PI	43.9 ± 0.2	32.9 ± 1.0	−3.0 ± 6.0	46.2 ± 0.2	23.6 ± 1.0	−97.0 ± 16.4	56.5 ± 2.0	−31.7 ± 3.6
GDP + PI	48.7 ± 0.1	51.5 ± 0.5	−31.2 ± 8.7[d]	50.0 ± 0.1	41.7 ± 0.5	−68.8 ± 12.8[d]	93.2 ± 1.0	−5.0 ± 2.6

[a] $\Delta\Delta H_u$ is the ΔH_u difference of an unfolding transition between the bound and free *Ms*PimA protein. It is expressed as % of total unfolding enthalpy difference between the bound and free protein.

[b] ΔH_u is the total unfolding enthalpy, $\Delta H_u = \Delta H_{u,1} + \Delta H_{u,2}$.

[c] $DDH_{u,int}$ is the change of intrinsic unfolding enthalpy of *Ms*PimA in the protein complexes (Equation 2; see "Experimental Procedures"). Mean values ± S.E. are reported for $T_{1/2}$ and ΔH_u based on three independent DSC experiments per protein samples. Errors for $\Delta\Delta H_u$ and $DDH_{u,int}$ were calculated literally.

[d] $\Delta\Delta H_u$, same as in footnote *a*, taking the unfolding of *Ms*PimA-GDP complex as reference.

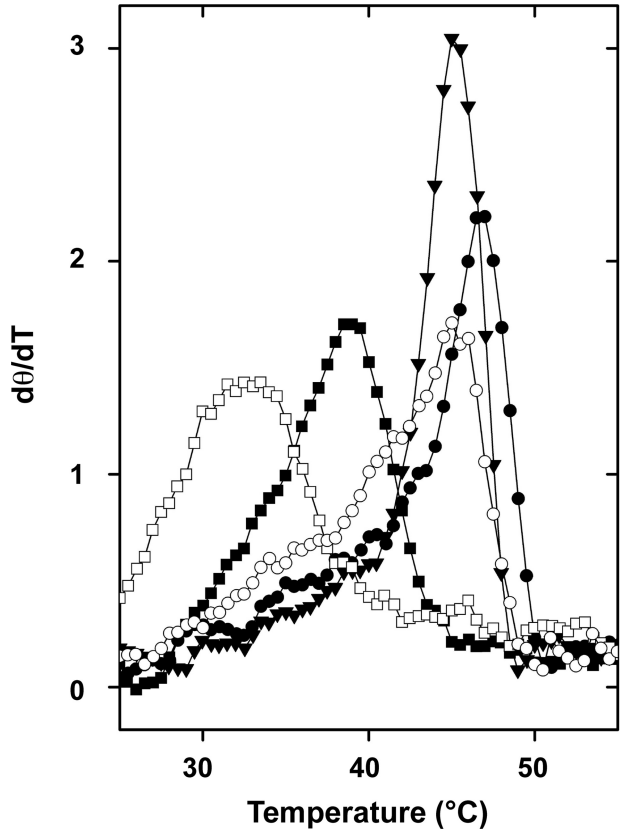

FIGURE 4. **Thermal unfolding of *Ms*PimA alone and in complex with its substrates as monitored by circular dichroism.** The first derivative of the ellipticity change at 222 nm is recorded as a function of the temperature. *Ms*PimA (■), and *Ms*PimA complexes with PI (□), GDP-Man (▼), GDP (●), and GDP and PI (○).

binding and, of note, almost equal and of opposite sign ($\Delta\Delta H_{u,int}$ = 28 kcal/mol and −32 kcal/mol for the GDP-Man and PI complexes, respectively; Table 1).

CD experiments were performed, confirming the two-domain unfolding process of *Ms*PimA as well as the changes in thermal stability of the *Ms*PimA complexes (Fig. 4). Changes in the ellipticity of free and bound *Ms*PimA at 222 nm as a function of temperature clearly reproduced the asymmetry of the C_p curves as well as the $T_{1/2}$ and unfolding cooperativity differences evidenced by DSC.

Hydrodynamic Properties of Free and Substrate-bound MsPimA—AUC, an accurate tool for the determination of hydrodynamic properties of proteins in solution, has been used successfully for detecting and characterizing substrate-induced conformational changes (39). To monitor the conformational changes in PimA in the absence and presence of GDP-Man, GDP, guanosine, and PI, sedimentation velocity AUC studies of pure *Ms*PimA were performed. Velocity data were analyzed and corrected for diffusion using the boundary analysis method of Demeler and van Holde (31, 32). The diffusion-corrected sedimentation coefficient distributions are plotted in Fig. 5. The nearly vertical distribution (*s*) indicates that *Ms*PimA sedimented as a single homogeneous species with an average sedimentation coefficient of 3.22 *s*, which is consistent with a moderately asymmetric (f/f_0 = 1.4) monomeric protein (Table 2, Fig. 5*A*). Upon the addition of equimolar GDP or GDP-Man, the sedimentation coefficients increased slightly to 3.53 and 3.50 *s*, respectively. Taking into the account the apparent 1:1 stoichiometry of binding and the relatively insignificant increase in the molecular weight of the complexes due to binding of these ligands, this change indicates the formation of a slightly more symmetrical and compact structure (Table 2, Fig. 5*A*). In contrast, the presence of guanosine did not significantly affect the sedimentation coefficient value of *Ms*PimA, consistent with the requirement of the enzyme for the β-PO_4 to achieve the closed conformation.

The addition of PI to *Ms*PimA resulted in a more significant change in the distribution and a resulting change in the average *s*-value from 3.22 to 2.79 *s*, consistent with the formation of a significantly less compact structure. PI also induced a conformational change of the *Ms*PimA-GDP complex, in agreement with the proteolysis experiments (Table 1, Fig. 5*B*). It is noteworthy that the differences in *Ms*PimA compactness among the various substrate complexes paralleled the variation of unfolding energy of the protein substrate complexes ($\Delta\Delta H_{u,int}$; Table 1). As shown in Fig. 5*C*, a linear correlation is observed between the sedimentation coefficient variations and the intrinsic unfolding enthalpy variations of *Ms*PimA inside the various protein complexes (Tables 1 and 2). These observations are in agreement with the existence of large conformational changes of *Ms*PimA between a more relaxed or "open" PI-bound conformation and a more compact or closed GDP-bound conformation. Furthermore, the linear correlation between ΔS_{AV} and $\Delta\Delta H_{u,int}$ together with the opposite values of ΔS_{AV} and $\Delta\Delta H_{u,int}$ on GDP-Man and PI binding strongly suggests that the formation of the open and closed protein conformations

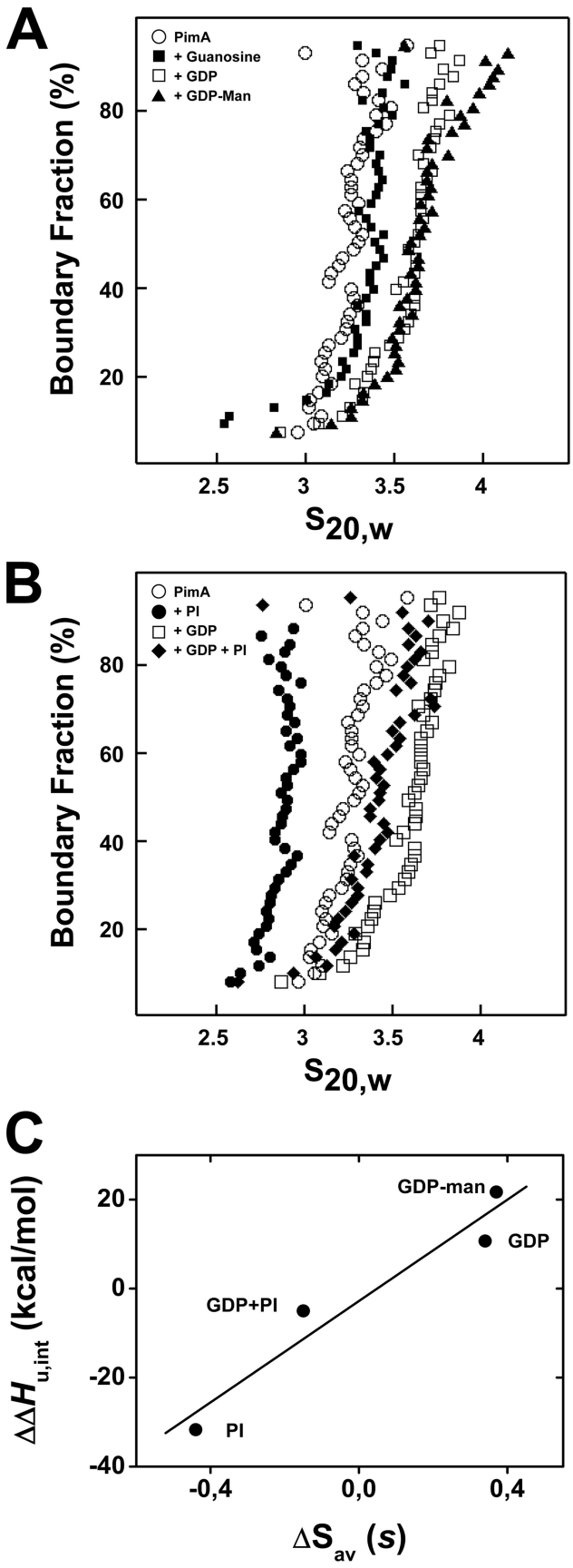

FIGURE 5. **AUC studies of *Ms*PimA and of *Ms*PimA-substrate complexes.** *A* and *B*, sedimentation velocity. *Ms*PimA was incubated alone (○) or with equimolar amounts of guanosine (■), GDP (□), GDP-Man (▲), PI (●), and GDP-PI (◆) prior to sedimentation velocity. The resulting integral distribution of *s* (corrected for water at 20 °C ($s_{20,w}$)) is shown. *C*, correlation between *Ms*PimA compaction and unfolding energy variations inside the various protein-substrate complexes. For each *Ms*PimA complex, the intrinsic unfolding enthalpy variation of the protein, $\Delta\Delta H_{u,int}$ (Table 1), is plotted against the sedimentation coefficient variation of the protein complex, ΔS_{AV} (Table 2). Linear fit of the data gives $\Delta\Delta H_{u,int} = -1.995 + 54.129 \times \Delta S_{AV}$ with $r = 0.958$.

TABLE 2
Hydrodynamic properties of free and bound *Ms*PimA measured by AUC

	S_{AV}[a]	DS_{AV}[a,b]	f/f_0
*Ms*PimA	3.22 ± 0.05		1.40 ± 0.02
*Ms*PimA + GDP-Man	3.50 ± 0.14	0.28 ± 0.14	1.29 ± 0.01
*Ms*PimA + GDP	3.53 ± 0.08	0.31 ± 0.09	1.27 ± 0.01
*Ms*PimA + GUA	3.21 ± 0.11	−0.01 ± 0.12	1.40 ± 0.01
*Ms*PimA + PI	2.79 ± 0.04	−0.43 ± 0.06	1.61 ± 0.03
*Ms*PimA + GDP + PI	3.33 ± 0.13	−0.20 ± 0.14[c]	1.35 ± 0.01

[a] S_{AV} and ΔS_{AV} are in Svedberg units (S).
[b] S_{AV} variation with respect to free *Ms*PimA.
[c] S_{AV} variation with respect to the *Ms*PimA-GDP complex. Mean values ± S.E. are reported based on two independent AUC experiments using two different preparations of *Ms*PimA.

results from the disruption and formation of the same intra protein contacts.

Molecular Architecture of Ins-P Binding Site—Despite much effort, we were unable to crystallize *Ms*PimA-GDP-PI ternary complexes even when assaying PI derivatives such as Ins-P, glycerophosphoryl-*myo*-inositol, 1,2-dibutyryl-*sn*-glycero-3-phosphoinositol, or 1,2-dioctanoyl-*sn*-glycero-3-phosphoinositol. To describe the architecture of the *Ms*PimA Ins-P binding site, we thus generated a three-dimensional model of the *Ms*PimA-GDP-Man-Ins-P complex by *in silico* molecular docking approaches. The Ins-P ring was placed in a well-defined pocket with its O_2 atom favorably positioned to receive the mannosyl residue from GDP-Man (Fig. 6, *A* and *B*). Interestingly, the overall structure of *Ms*PimA is very similar to that of the recently solved MshA from *Corynebacterium glutamicum* (*Cg*MshA) even though the sequence identity of the two proteins is low (24%, after structural alignment) (Fig. 6*A*). MshA is involved in the biosynthesis of mycothiols where it catalyzes the transfer of *N*-acetylglucosamine from UDP-*N*-acetylglucosamine to Ins-P. The structure of MshA has been determined both in the absence of substrates and in complex with UDP and Ins-P (17). A detailed comparison of the *Cg*MshA-UDP-Ins-P complex with the three-dimensional model of *Ms*PimA-GDP-Man-Ins-P revealed common features in the Ins-P binding site (Fig. 6*B*). Thus, the C-terminal domain of *Ms*PimA (residues 1–169 and 349–373) superimposes well with the equivalent domain in *Cg*MshA (residues 1–196 and 390–409) (r.m.s.d. of 1.8 Å for 171 aligned residues, 31% identity). In addition, the central β-sheet of the N-terminal domains of *Ms*PimA (residues 170–348) and *Cg*MshA (residues 197–389) display a similar topology (r.m.s.d. of 2.8 Å for 162 aligned residues, 19% identity).

However, two important regions that interact with the Ins-P ring in *Cg*MshA have a different conformation in *Ms*PimA. The first one is the region including the connecting loop between β4 and α2 (residues 68–79) and the amphipatic helix α2 (residues 80–99) (Fig. 6*A*). This loop contains the Lys^{78}, which as a side chain that forms an important salt bridge with the phosphate

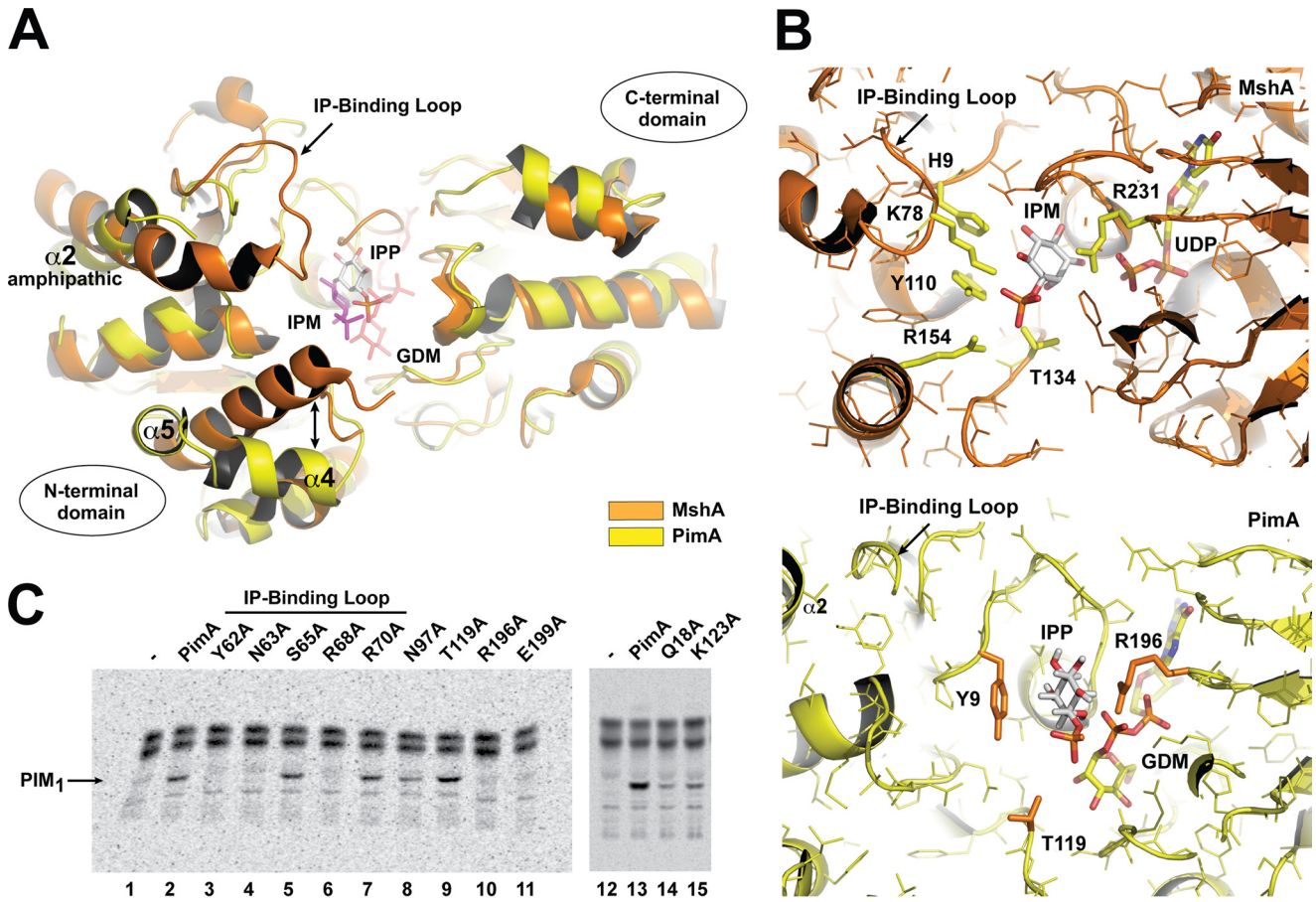

FIGURE 6. **Molecular architecture of the inositol binding site.** *A*, structural alignment between *Ms*PimA (PDB code 1GEJ (*yellow*)) and *Cg*MshA (PDB code 3C4V (*orange*)). *B*, detailed comparison of the Ins-P binding sites from *Ms*PimA and *Cg*MshA. *C*, enzymatic activities of wild-type *Ms*PimA and selected mutants (see "Experimental Procedures"). TLC autoradiograph of the incorporation of GDP-[^{14}C]Man into mycobacterial membrane mannolipids from *M. smegmatis* mc^{2}155. Enzymatic activity is shown as a percentage of *Ms*PimA activity (set at 100%). *Lane 1*, membranes alone; *lane 2*, membranes + *Ms*PimA^{WT}; *lane 3*, membranes + *Ms*PimA^{Y62A} (0%); *lane 4*, membranes + *Ms*PimA^{N63A} (0%); *lane 5*, membranes + *Ms*PimA^{S65A} (100%); *lane 6*, membranes + *Ms*PimA^{R68A} (0%); *lane 7*, membranes + *Ms*PimA^{R70A} (100%); *lane 8*, membranes +*Ms*PimA^{N97A} (71%); *lane 9*, membranes + *Ms*PimA^{T119A} (100%); *lane 10*, membranes + *Ms*PimA^{R196A} (0%); *lane 11*, membranes + *Ms*PimA^{E199A} (0%); *lane 12*, membranes alone; *lane 13*, membranes + *Ms*PimA^{WT}; *lane 14*, membranes + *Ms*PimA^{Q18A} (9%); *lane 15*, membranes + *Ms*PimA^{K123A} (23%).

moiety of Ins-P in *Cg*MshA (Fig. 6*B*, Ref. 17). We have previously shown that the equivalent loop in *Ms*PimA (residues ^{59}PIPYNGSVARLR^{70}) is disordered in the crystal structure, indicating conformational flexibility. This loop is strictly conserved in all mycobacterial PimA orthologues, and its deletion drastically impairs the interaction of the protein with PI aggregates *in vitro* (24). To investigate the role of this loop in Ins-P binding in *Ms*PimA, we prepared a series of five point mutants: *Ms*PimA^{Y62A}, *Ms*PimA^{N63A}, *Ms*PimA^{S65A}, *Ms*PimA^{R68A}, and *Ms*PimA^{R70A} (Fig. 6, *B* and *C*). Using a radiometric assay (Fig. 6*C*) (2), no detectable activity could be observed for the *Ms*PimA^{Y62A}, *Ms*PimA^{N63A}, and *Ms*PimA^{R68A} mutants, strongly supporting a role for these residues in Ins-P binding. The second major difference between MshA and PimA involves *Cg*MshA α4 (residues 148–165), which has a similar length (17 residues) as the sum of two consecutive α-helices in *Ms*PimA, α4 (residues 123–131) and α5 (residues 134–140) (Fig. 6*A*). The phosphate group of Ins-P makes a hydrogen bond interaction with the guanidinium group of Arg^{154} in *Cg*MshA (Fig. 6*B*). To investigate whether the homologous *Ms*PimA residue (Lys^{123}) is involved in Ins-P binding, this residue was replaced by alanine. As shown in Fig. 6*C*, the mutant *Ms*PimA^{K123A} was 23% less active than *Ms*PimA^{WT}, suggesting an involvement of Lys^{123} in substrate binding.

Other important residues and their interactions in the Ins-P binding pocket are preserved in *Ms*PimA/*Cg*MshA (Fig. 6*B*), further supporting a common Ins-P binding mechanism. The residue Gln^{18} is structurally equivalent to Asn^{25} in *Cg*MshA, which interacts with the O-4 of the Ins-P ring. Substitution of Gln^{18} by alanine also resulted in a strong reduction (91%) of the enzymatic activity (Fig. 6*C*). The O-5 hydroxyl interacts with His^{9} in *Cg*MshA, which is structurally equivalent to the essential Tyr^{9} in *Ms*PimA (24). In addition to the previously described Arg^{78} and Arg^{154}, the side chain of Tyr^{110} (Thr^{134} in *Ms*PimA) also forms hydrogen-bonding interactions with the phosphate moiety of Ins-P in *Cg*MshA. This Thr^{134} corresponds to Thr^{119} in *Ms*PimA. Moreover, Arg^{231}, which was found to be a very important residue in delineating the substitution and conformation of the Ins-P ring in *Cg*MshA, has an equivalent in *Ms*PimA, Arg^{196}, located in the β10-α8 loop (residues 194–200) in the C-terminal domain. This arginine was also found to change its conformation upon substrate binding in other GT-B enzymes (9, 16, 40). Its substitution by alanine in *Ms*PimA completely abolished enzymatic activity (Fig. 6*C*).

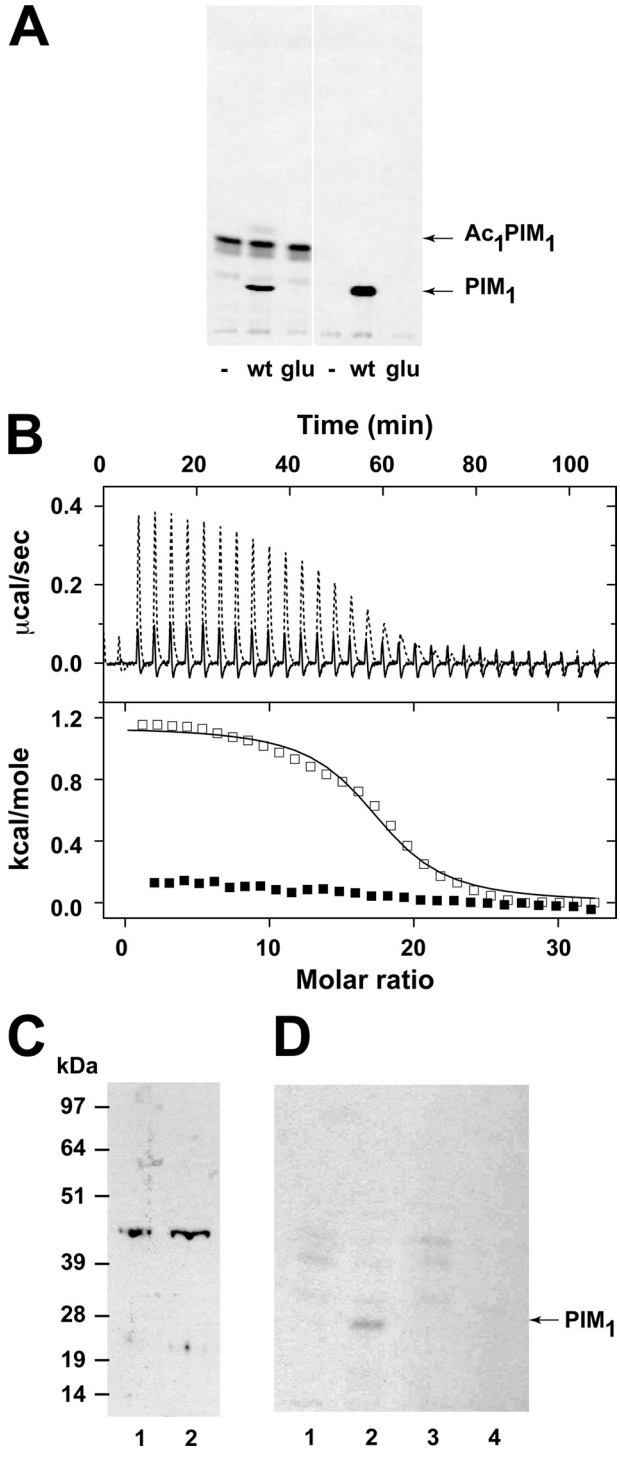

FIGURE 7. **The *Ms*PimA amphiphatic α2 is essential for membrane binding *in vivo* and *in vitro*.** *A*, enzymatic activity of *Ms*PimA and *Ms*PimA$^{R77E/K78E/K80E/K81E}$ mutant expressed in *E. coli* cells. TLC autoradiograph of the incorporation of GDP-[^{14}C]Man into mycobacterial membrane mannolipids from *M. smegmatis* mc^{2}155. *Lane 1*, membranes alone; *lane 2*, membranes + *Ms*PimA; *lane 3*, membranes + *Ms*PimA$^{R77E/K78E/K80E/K81E}$; *lane 4*, PI alone; *lane 5*, PI + *Ms*PimA; *lane 6*, PI + *Ms*PimA$^{R77E/K78E/K80E/K81E}$. *B*, ITC binding isotherms for the binding of PI aggregates to the *Ms*PimA-GDP (□) and *Ms*PimA$^{R77E/K78E/K80E/K81E}$-GDP (■) complexes at 25 °C. Heats of injections were corrected for the heat of dilution of PI and normalized to the concentration of PI injected. *Solid* and *dotted lines* correspond to best fit of data using a bimolecular interaction model. *C*, Western blot analysis of the cytoplasmic fractions of *M. smegmatis* mc^{2}155 cells expressing: *lane 1*, *Ms*PimA; *lane 2*, *Ms*PimA$^{R77E/K78E/K80E/K81E}$. The recombinant proteins were detected using anti-His-tag antibodies. *D*, cytosol fractions were used for the enzymatic activity assays of *Ms*PimA and *Ms*PimA$^{R77E/K78E/K80E/K81E}$ mutants expressed in *M. smegmatis* mc^{2}155 cells. TLC autoradiograph of the incorporation of GDP-[^{14}C]Man into PIM$_1$. *Lane 1*, *Ms*PimA cytosol alone; *lane 2*, *Ms*PimA cytosol + PI; *lane 3*, *Ms*PimA$^{R77E/K78E/K80E/K81E}$ cytosol alone; *lane 4*, *Ms*PimA$^{R77E/K78E/K80E/K81E}$ cytosol + PI.

Furthermore, the *Ms*PimA-GDP-Man-Ins-P model predicted that Glu^{199}, which is located in the same loop as the essential residues Arg^{196} and Arg^{201}, might be positioned favorably to interact with the phosphate group of Ins-P. The role of Glu^{199} in acceptor substrate binding was confirmed by directed mutagenesis; its replacement by alanine completely inactivated the enzyme.

Altogether, these site-directed mutants validated our model and provided evidence that Ins-P binding is defined by key interactions with residues located in the β1-α1 loop (residues 7–16), the β3-α2 loop (residues 59–70), the β6-α4 loop (residues 117–123), two α-helices (α4 and α5) from the N-terminal domain, and the β10-α8 loop (residues 194–200) from the C-terminal domain. These results also suggest that an important conformational change most likely takes place upon PI binding, in agreement with the limited proteolysis, DSC, CD, and AUC data.

Amphipathic Helix α2 Is Essential for Membrane Binding in Vivo—Amphipathic α-helices represent a common membrane-binding motifs in proteins. They bind to lipid membranes by a physical adsorption mechanism in which electrostatic and hydrogen bond interactions are dominant (21). This process usually follows three thermodynamic steps: (i) the binding is initiated by the electrostatic attraction of a cationic peptide to the anionic membrane head group; (ii) the peptide is then disposed into the plane of binding where its exact position depends on the hydrophobic/hydrophilic balance of the molecular groups and forces involved; and (iii) the bound peptide changes its conformation. In some cases peptides that are in random coil conformation in solution can adopt an α-helix conformation when associated with the lipid membrane (22). There is evidence suggesting that PimA binds to the membrane in a process mediated by electrostatic interactions (24). Although the enzyme lacks hydrophobic transmembrane α-helices, it has been localized to a subfraction of the plasma membrane, termed PM$_f$. A salt wash of the PM$_f$ fraction significantly reduces the synthesis of PIM$_1$ (41). Moreover, the addition of negatively charged lipids (*e.g.* cardiolipin or 1,2-dipalmitoyl-*sn*-glycero-3-phosphate) to a reaction mixture that included sub-critical mycelle concentrations of the acceptor 1,2-dioctanoyl-*sn*-glycero-3-phosphoinositol drastically increased enzymatic activity (24). We have previously postulated that the amphiphatic α2 might be an important recognition element to determine binding of PimA to the membrane. To further investigate this possibility, four amino acid substitutions were introduced by replacing Arg^{77}, Lys^{78}, Lys^{80}, and Lys^{81} with glutamic acid residues (*Ms*PimA$^{R77E/K78E/K80E/K81E}$). These basic residues are well conserved in other PimA orthologues including *Mycobacterium bovis*, *Mycobacterium leprae*, and *M. tuberculosis* (24). These substitutions by negatively charged residues completely inactivated *Ms*PimA (Fig. 7*A*). The binding properties of *Ms*PimA$^{R77E/K78E/K80E/K81E}$ were fur-

ther compared with those of the wild-type protein using ITC. The mutant protein was observed to bind GDP with a binding affinity in the submicromolar range in an enthalpy-driven reaction ($\Delta H = -12.7$ kcal mol^{-1}, $\Delta H/\Delta G = 149\%$) similar to that observed for the wild-type protein ($\Delta H = -14.0$ kcal mol^{-1}, $\Delta H/\Delta G = 136\%$), strongly suggesting that the integrity of the protein was not affected by the amino acid substitutions. Instead, significant differences were observed in the membrane binding properties of the wild-type and mutant enzymes. *Ms*PimA binds phospholipid aggregates (micelles or liposomes), and these interactions induce significant increase in the enzyme activity (24). When *Ms*PimA$^{R77E/K78E/K80E/K81E}$ was assessed for PI binding at a high PI concentration (Fig. 7*B*), no detectable binding ($K_d > 200$ μM) could be observed in conditions where the wild-type protein bound to PI aggregates of 15–20 PI molecules with high affinity ($K_d \sim 3.0$ μM). Taken together, the amphipathic helix α2 seems to be a major element for *Ms*PimA membrane interaction.

The mutated *MsPimA*$^{R77E/K78E/K80E/K81E}$ gene was then expressed in *M. smegmatis* mc^{2}155, and the cytosolic and membrane fractions from the recombinant strain were prepared and used in *in vitro* assays (2). Although production of PIM_1 was observed in the cytosol in the strain overproducing *Ms*PimA^{WT} upon addition of PI, no PIM_1 was formed under the same conditions in the strain overproducing *Ms*PimA$^{R77E/K78E/K80E/K81E}$ (Fig. 7, *C* and *D*). Because *M. smegmatis* produced similar amounts of PimA when transformed with *Ms*PimA$^{R77E/K78E/K80E/K81E}$ or with *Ms*PimA^{WT} (Fig. 7*C*), this result was clearly not due to a defect in the production of the recombinant protein. That *Ms*PimA$^{R77E/K78E/K80E/K81E}$ encodes an inactive protein was also confirmed by complementation experiments. As mentioned above, it has been established that PimA is an essential enzyme of *M. smegmatis* and *M. tuberculosis* (2).[4] Using a homologous recombination strategy, we found that whereas the wild-type chromosomal locus of *pimA* in *M. smegmatis* could be disrupted in the presence of a rescue copy of *Ms*PimA, *Ms*PimA$^{R77E/K78E/K80E/K81E}$ failed to rescue the knock-out mutant (data not shown). The amphipathic α2 helix is thus required for activity *in vivo*. Taken together, these results strongly argue in favor of the amphipathic helix α2 as being a major *Ms*PimA component in membrane interaction, which seems to be essential for the growth of mycobacteria.

A Model of Action for PimA—The substrate-induced conformational changes and the mutagenesis studies described above support a mechanism for the binding of the donor and lipid acceptor substrates that seems to be of importance during catalysis. According to this model, the donor GDP-Man induces a striking conformational change from an open to a closed state wherein the β-PO_4 plays a critical role in the stabilization of the enzyme (Fig. 8*A*). The open to closed motion brings together critical residues from the N- and C-terminal domains, allowing the formation of a functionally competent active site. When the lipid acceptor PI binds to PimA, an opposite structural rearrangement takes place that destabilizes the

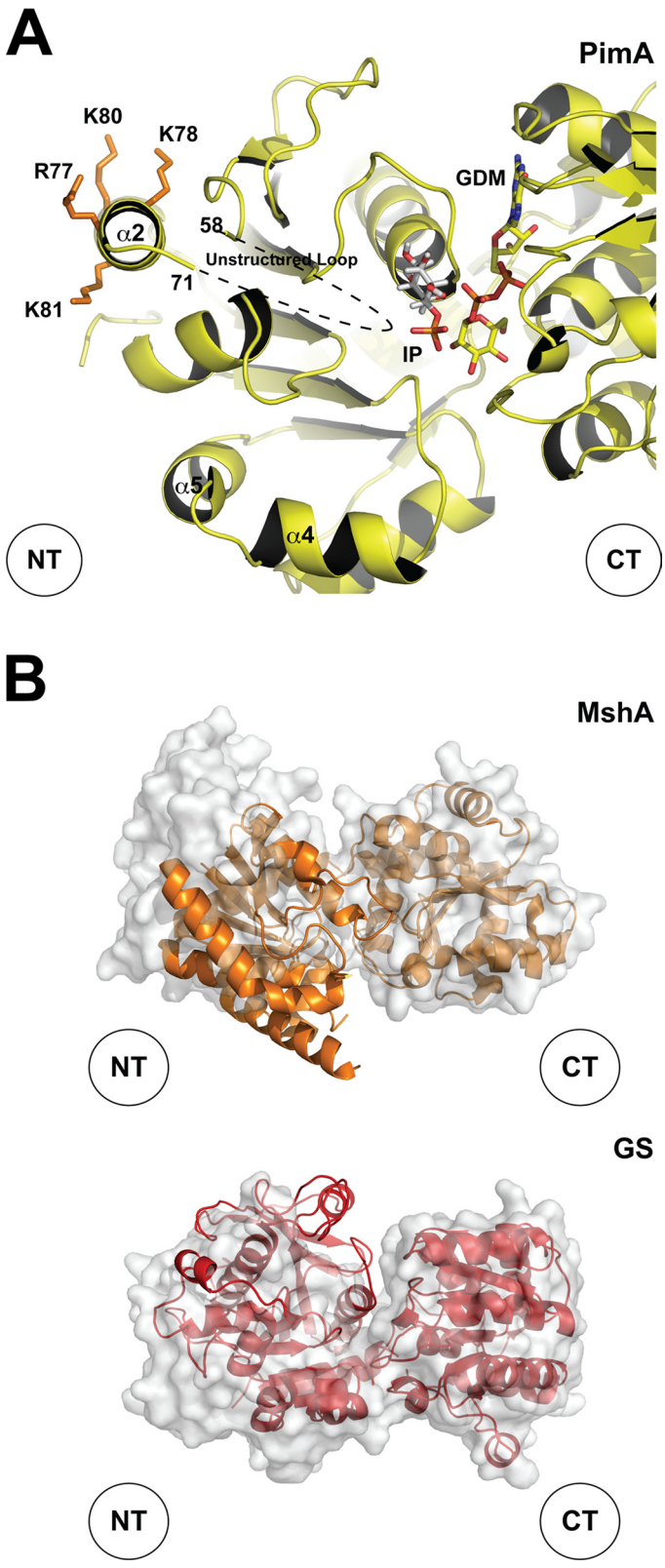

FIGURE 8. **A proposed model of action for PimA.** *A*, schematic representing the secondary structure of a selected region of *Ms*PimA including the GDP-Man binding site, the Ins-P binding site, and the amphipathic α2 helix involved in membrane association. *B*, structural comparison between the open (molecular surface representation) and closed (schematic) states of MshA from *C. glutamicum* and GS from *A. tumefaciens*.

[4] M. Jackson, unpublished results.

enzyme. It could be speculated that PI binding allows the initiation of the enzymatic reaction and induces the "opening" of the protein in order to release the products.

Several lines of evidence support this model. The crystal structures of the *Ms*PimA-GDP and *Ms*PimA-GDP-Man complexes revealed that the proteins crystallize in a closed conformation. Both structures superimpose well (r.m.s.d. of 0.3Å for 361 identical residues), clearly showing that residues Gly^{16}, from the N-terminal domain, and Arg^{196}/Lys^{202}, from the C-terminal domain, participate in the stacking interaction of β-PO_4 upon GDP or GDP-Man binding (24). These residues are largely conserved not only in PimA orthologues from other mycobacterial species but also in other GT-B enzymes (14). For instance, structural comparisons of the free and nucleotide-diphospho-sugar-bound forms of *Cg*MshA and the glycogen synthase from *Agrobacterium tumefaciens* (*At*GS) revealed that an important subdomain rotation is required to achieve the closed state (Fig. 8*B*) (16, 17). Residues Gly^{23}, Arg^{231}, and Lys^{236} in MshA and Gly^{18}, Arg^{299}, and Lys^{304} in *At*GS interact with the β-PO_4 of uridine 5′-diphosphate *N*-acetylglucosamine (UDP-GlcNAc) and adenosine 5′-diphosphate glucose (ADP-Glc), respectively. Interestingly, these amino acids occupy the same location as the residues Gly^{16}, Arg^{196}, and Lys^{202} in the closed conformation of *Ms*PimA. The catalytic mechanism of retaining GT-B enzymes is still a controversial matter of debate. Recent work suggests the direct implication of the β-PO_4 in catalysis in the context of the proposed S_Ni internal return mechanism. This model proposes that the β-PO_4 is responsible for the nucleophilic attack of the acceptor substrate and the cleavage of the sugar-nucleotide bond (17).

How does PimA recognize PI? The answer to this question is still incomplete, in part because no direct structural information is available for any peripheral membrane-binding GT in the presence of its lipid acceptor substrate. In general, peripheral membrane-binding proteins tend to penetrate one leaflet of the lipid bilayer membrane. In the case of PimA, this interaction is mediated mainly by electrostatic forces located in the essential amphiphatic helix α2 on the N-terminal domain. Our site-directed mutagenesis studies not only validate the position of Ins-P in a pocket in close proximity with the α2 helix and the mannose ring of GDP-Man but also suggest an important rearrangement of two α-helices (α2 and α4) after PI binding. Furthermore, our biophysical studies strongly argue in favor of the formation of a new open state after PI binding.

Structural Similarity with Peripheral Membrane-binding GT-B Superfamily—GTs have been classified into distinct families based on amino acid sequence similarities. Within each family, GTs are expected to adopt a unique three-dimensional fold (see the CAZy data base). On the basis of this classification, the three-dimensional structures of 12 GT-B family representatives, including GT1, GT4, GT5, GT9, GT10, GT20, GT23, GT28, GT35, GT63, GT70, and GT80, have been reported. As shown in supplemental Table 2S, GTs from families GT5 (*e.g.* GS (16)), GT20 (*e.g.* OtsA (9)), GT35 (*e.g.* GP (42)), GT63 (*e.g.* BGT (13)), and GT72 (*e.g.* AGT (40)) are soluble proteins with no indication of membrane association and include both retaining and inverting enzymes. In contrast, families GT9 (*e.g.* WaaC (43)), GT10 (*e.g.* FucT (44)), GT23 (*e.g.* FUT8 (45)), GT28 (*e.g.* MurG (15)), GT70 (*e.g.* GumK (46)), and GT80 (*e.g.* ST1 (47)) comprise exclusively peripheral membrane-associated proteins, all of which are inverting enzymes. Peripheral membrane-associated GTs were also observed or predicted in the retaining GT4 family (*e.g.* PimA), indicating that the membrane association process is independent of the catalytic mechanism. Interestingly, the GT1 family, like the GT4 family, includes both soluble and membrane-associated proteins, where members of family GT1 bind membranes using a variety of mechanisms including transmembrane α-helices, amphipathic α-helices, and protein-protein interactions (20, 48).

It was reported previously that the calculation of the isoelectric points in GT-B enzymes lacking transmembrane domains proved to be useful in discriminating soluble from peripheral membrane-bound proteins (19). Typically, N-terminal domains have higher pI values than C-terminal domains, reflecting the presence of a positively charged surface, including potential amphipathic α-helices. The electrostatic surface potential of *Ms*PimA reveals a polar protein with a positively charged N-terminal domain (pI 8.1) in agreement with the presence of the amphipathic helix α2 (24). A detailed comparison of *Ms*PimA with the available three-dimensional structures of GT-B enzymes reveals the presence of amphipathic α-helices in the N-terminal domains of some homologues including WaaG (GT4), WaaC and WaaF (GT9), MurG (GT28), and GumK (GT70), all of which are peripheral membrane-associated proteins (supplemental Fig. 4*S*). Interestingly, a similar mechanism of membrane association was also recently described for GT-A enzymes (49). The close proximity between the amphipathic α-helices and the acceptor binding site may be required to improve the efficient glycosylation of membrane-bound compounds such as PI. As lipid acceptors exhibit a huge diversity of structures, the corresponding N-terminal recognition domains in GT-B GTs are not expected to display significant similarities in their tertiary structures. In fact, the N-terminal domains in PimA, WaaG, WaaC, WaaF, MurG, and GumK are structurally more distant (r.m.s.d. of 2.5–3.4 Å) than the nucleotide-diphospho-sugar-binding C-terminal domains (r.m.s.d. of 2.2–2.9 Å) as a consequence of the different rearrangement of secondary structural elements. Although the superimposition of the N-terminal domains shows that the amphipathic α-helices display very different orientations, all of them have been found to be exposed to the solvent and disposed almost perpendicular to the entrance of the interdomain cleft. Inspection of these structures also indicates the presence of a pocket compatible with a putative binding site for lipid acceptor in close proximity to the positively charged clusters. These observations suggest a common molecular mechanism of membrane association and lipid recognition for this family of peripheral membrane-binding GTs.

Conclusions—PimA catalyzes an essential step in the biosynthesis of PIMs, key structural elements of the cell envelope of mycobacteria that also play a role in host-pathogen interactions. The enzyme is a paradigm of a large family of peripheral membrane-binding GTs for which the molecular mechanism of substrate/membrane recognition and catalysis is unknown. This article reports a detailed investigation of the substrate-induced conformational changes of PimA in solution. Alto-

gether, our findings support the notion that flexibility and conformational transitions confer adaptability of PimA to the donor/acceptor substrates and to the membrane, which is essential for the reaction to take place. The proposed mechanism has fundamental implications for the comprehension of the early steps of PimA-mediated PIM biosynthesis as well as of peripheral membrane-binding GTs at the molecular level.

Substrate-induced conformational changes in proteins are known to be critically important in drug discovery and development strategies (50). Given the fact that PimA is an essential enzyme for mycobacterial growth, the information presented here may thus prove valuable in the rational design of specific PimA inhibitors.

Acknowledgments—We gratefully acknowledge Dr. Karolin Luger (Department of Biochemistry and Molecular Biology, Colorado State University) for providing us with full access to their protein purification facilities. We also thank Dr. Otto Pritsch (Unit of Protein Biophysics, Pasteur Institute of Montevideo, Uruguay) and Dr. Silvia Altabe (Department of Microbiology, Institute of Molecular and Cell Biology of Rosario, Argentina) for valuable scientific discussions and Dr. Steven McBryant (Department of Biochemistry and Molecular Biology, Colorado State University) for help with AUC experiments.

REFERENCES

1. Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) *Ann. Rev. Biochem.* **77,** 521–555
2. Korduláková, J., Gilleron, M., Mikusova, K., Puzo, G., Brennan, P. J., Gicquel, B., and Jackson, M. (2002) *J. Biol. Chem.* **277,** 31335–31344
3. Briken, V., Porcelli, S. A., Besra, G. S., and Kremer, L. (2004) *Mol. Microbiol.* **53,** 391–403
4. Berg, S., Kaur, D., Jackson, M., and Brennan, P. J. (2007) *Glycobiology* **17,** 35R–56R
5. Gilleron, M., Jackson, M., Nigou, J., and Puzo, G. (2008) *The Mycobacterial Cell Envelope* (Daffe, M., and Reyrat, J. M., eds) pp. 75–105, ASM Press, Washington, DC
6. Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G., and Strynadka, N. C. (2001) *Nat. Struct. Biol.* **8,** 98–100
7. Lairson, L. L., and Withers, S. G. (2004) *Chem. Commun.* **20,** 2243–2248
8. Lairson, L. L., Chiu, C. P., Ly, H. D., He, S., Wakarchuk, W. W., Strynadka, N. C., and Withers, S. G. (2004) *J. Biol. Chem.* **279,** 28339–28344
9. Gibson, R. P., Turkenburg, J. P., Charnock, S. J., Lloyd, R., and Davies, G. J. (2002) *Chem. Biol.* **9,** 1337–1346
10. Coutinho, P. M., Deleury, E., Davies, G. J., and Henrissat, B. (2003) *J. Mol. Biol.* **328,** 307–317
11. Davies, G. J., Gloster, T. M., and Henrissat, B. (2005) *Curr. Opin. Struct. Biol.* **15,** 637–645
12. Charnock, S. J., and Davies, G. J. (1999) *Biochemistry* **38,** 6380–6385
13. Vrielink, A., Rüger, W., Driessen, H. P., and Freemont, P. S. (1994) *EMBO J.* **13,** 3413–3422
14. Wrabl, J. O., and Grishin, N. V. (2001) *J. Mol. Biol.* **314,** 365–374
15. Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., Mokhtarzadeh, M., and Walker, S. (2003) *Proc. Natl. Acad. Sci.* **100,** 845–849
16. Buschiazzo, A., Ugalde, J. E., Guerin, M. E., Shepard, W., Ugalde, R. A., and Alzari, P. M. (2004) *EMBO J.* **23,** 3196–3205
17. Vetting, M. W., Frantom, P. A., and Blanchard, J. S. (2008) *J. Biol. Chem.* **283,** 15834–15844
18. Morita, Y. S., Sena, C. B., Waller, R. F., Kurokawa, K., Sernee, M. F., Nakatani, F., Haites, R. E., Billman-Jacobe, H., McConville, M. J., and Kinoshita, T. (2006) *J. Biol. Chem.* **281,** 25143–25155
19. Lind, J., Rämö, T., Rosen Klement, M., Bárány-Wallje, E., Epand, R., Epand, R. F., Mäler, L., and Wieslander, A. (2007) *Biochemistry* **46,** 5664–5677
20. Wang, X., Weldeghiorghis, T., Zhang, G., Imperiali, B., and Prestegard, J. H. (2008) *Structure* **16,** 965–975
21. Seelig, J. (2004) *Biochim. Biophys. Acta* **1666,** 40–50
22. Wieprecht, T., Apostolov, O., Beyermann, M., and Seelig, J. (2000) *Biochemistry* **39,** 191–201
23. Guerin, M. E., Buschiazzo, A., Korduláková, J., Jackson, M., and Alzari, P. M. (2005) *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **61,** 518–520
24. Guerin, M. E., Kordulakova, J., Schaeffer, F., Svetlikova, Z., Buschiazzo, A., Giganti, D., Gicquel, B., Mikusova, K., Jackson, M., and Alzari, P. M. (2007) *J. Biol. Chem.* **282,** 20705–20714
25. Schaeffer, F., Matuschek, M., Guglielmi, G., Miras, I., Alzari, P. M, and Béguin, P. (2002) *Biochemistry* **41,** 2106–2114
26. Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989) *Anal. Biochem.* **179,** 131–137
27. Hible, G., Renault, L., Schaeffer, F., Christova, P., Zoe Radulescu, A., Evrin, C., Gilles, A. M., and Cherfils, J. (2005) *J. Mol. Biol.* **352,** 1044–1059
28. Plotnikov, V. V., Brandts, J. M., Lin, L. N., and Brandts, J. F. (1997) *Anal. Biochem.* **250,** 237–244
29. Makhatadze, G. I., and Privalov, P. L. (1990) *J. Mol. Biol.* **213,** 375–384
30. Griko, Y. V., Freire, E., Privalov, G., Van Dael, H., and Privalov, P. L. (1995) *J. Mol. Biol.* **252,** 447–459
31. Demeler, B., Behlke, J., and Ristau, O. (2000) *Methods Enzymol.* **321,** 38–66
32. Demeler, B., and van Holde, K. E. (2004) *Anal. Biochem.* **335,** 279–288
33. Holm, L., and Park, J. (2000) *Bioinformatics* **16,** 566–567
34. DeLano, W. L. (2003) *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA
35. Abagyan, R. A., Totrov, M. M., and Kuznetsov, D. N. (1994) *J. Comp. Chem.* **15,** 488–506
36. Hubbard, S. J. (1998) *Biochim. Biophys. Acta* **1382,** 191–206
37. Furukawa, K., Tagaya, M., Tanizawa, K., and Fukui, T. (1993) *J. Biol. Chem.* **268,** 23837–23842
38. Privalov, P. L., and Dragan, A. I. (2007) *Biophys. Chem.* **126,** 16–24
39. Lebowitz, J., Lewis, M. S., and Schuck, P. (2002) *Protein Sci.* **11,** 2067–2079
40. Moréra, S., Larivière, L., Kurzeck, J., Aschke-Sonnenborn, U., Freemont, P. S., Janin, J., and Rüger, W. (2001) *J. Mol. Biol.* **311,** 569–577
41. Morita, Y. S., Velasquez, R., Taig, E., Waller, R. F., Patterson, J. H., Tull, D., Williams, S. J., Billman-Jacobe, H., and McConville, M. J. (2005) *J. Biol. Chem.* **280,** 21645–21652
42. O'Reilly, M., Watson, K. A., Schinzel, R., Palm, D., and Johnson, L. N. (1997) *Nat. Struct. Biol.* **4,** 405–412
43. Grizot, S., Salem, M., Vongsouthi, V., Durand, L., Moreau, F., Dohi, H., Vincent, S., Escaich, S., and Ducruix, A. (2006) *J. Mol. Biol.* **363,** 383–394
44. Sun, H. Y., Lin, S. W., Ko, T. P., Pan, J. F., Liu, C. L., Lin, C. N., Wang, A. H., and Lin, C. H. (2007) *J. Biol. Chem.* **282,** 9973–9982
45. Ihara, H., Ikeda, Y., Toma, S., Wang, X., Suzuki, T., Gu, J., Miyoshi, E., Tsukihara, T., Honke, K., Matsumoto, A., Nakagawa, A., and Taniguchi, N. (2007) *Glycobiology* **17,** 455–466
46. Barreras, M., Salinas, S. R., Abdian, P. L., Kampel, M. A., and Ielpi, L. (2008). *J. Biol. Chem.* **283,** 25027–25035
47. Ni, L., Chokhawala, H. A., Cao, H., Henning, R., Ng, L., Huang, S., Yu, H., Chen, X., and Fisher, A. J. (2007) *Biochemistry* **46,** 6288–6298
48. Miley, M. J., Zielinska, A. K., Keenan, J. E., Bratton, S. M., Radominska-Pandya, A., and Redinbo, M. R. (2007) *J. Mol. Biol.* **369,** 498–511
49. Leipold, M. D., Kaniuk, N. A., and Whitfield, C. (2007) *J. Biol. Chem.* **282,** 1257–1264
50. Teague, S. J. (2003) *Nat. Rev. Drug Discov.* **2,** 527–541