JBC Papers in Press. Published on August 20, 2013 as Manuscript M113.462705
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.M113.462705

CONFORMATIONAL PLASTICITY OF THE ESSENTIAL MEMBRANE-ASSOCIATED MANNOSYLTRANSFERASE PIMA FROM MYCOBACTERIA*

David Giganti[1,2,3], Jorge Alegre-Cebollada[3], Saioa Urresti[1,2], David Albesa-Jové[1,2], Ane Rodrigo-Unzueta[1,2], Natalia Comino[1,2], Michael Kachala[4], Sonia López-Fernández[1,2], Dmitri I. Svergun[4], Julio M. Fernández[3, §] and Marcelo E. Guerin[1,2,5, §].

From the [1] Unidad de Biofísica, Consejo Superior de Investigaciones Científicas - Universidad del País Vasco/Euskal Herriko Unibertsitatea (CSIC-UPV/EHU), Barrio Sarriena s/n, Leioa, Bizkaia, 48940, Spain; [2] Departamento de Bioquímica, Universidad del País Vasco, Spain; [3] Department of Biological Sciences, Columbia University, Northwest Corner Building, 550 West 120 Street, New York, New York 10027, USA; [4] European Molecular Biology Laboratory, Hamburg Outstation, c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany, [5] IKERBASQUE, Basque Foundation for Science, 48011, Bilbao, Spain.

Running head: Structural plasticity of the GT-B mannosyltransferase PimA

[§] To whom correspondence should be addressed: Marcelo E. Guerin, Unidad de Biofísica, Centro Mixto Consejo Superior de Investigaciones Científicas - Universidad del País Vasco/Euskal Herriko Unibertsitatea (CSIC-UPV/EHU), Barrio Sarriena s/n, Leioa, Bizkaia, 48940, Spain, Tel: +34 94 601 8052; Fax : +34 94 601 3360; E-mail: mrcguerin@gmail.com; and Julio M. Fernandez, Department of Biological Sciences, Columbia University, Northwest Corner Building, 550 West 120 Street, New York, New York 10027, USA, Tel: (212)-854-9141; E-mail: jfernandez@columbia.edu.

Background: Knowledge of conformational changes and dynamics occurring in glycosyltransferases is limited.
Results: PimA unfolds at low force, following heterogeneous multiple-step mechanical unfolding pathways.
Conclusions: With the elucidation of the solution structure of PimA, we conclude the enzyme displays remarkable structural plasticity that seems to be important for substrate/membrane association.
Significance: This work constitutes the first conformational study of a glycosyltransferase at the single-molecule level.

Phosphatidyl-*myo*-inositol mannosyltransferase A (PimA) is an essential glycosyltransferase (GT) that initiates the biosynthetic pathway of phosphatidyl-*myo*-inositol mannosides (PIMs), lipomannan (LM) and lipoarabinomannan (LAM), which are key glycolipids/lipoglycans of the mycobacterial cell envelope. PimA belongs to a large family of peripheral membrane-associated GTs for which the understanding of the molecular mechanism and conformational changes that govern substrate/membrane recognition and catalysis remains a major challenge. Here we used single molecule force spectroscopy techniques to study the mechanical and conformational properties of PimA. In our studies, we engineered a polyprotein containing PimA flanked by four copies of the well characterized I27 protein which provides an unambiguous mechanical fingerprint. We found that PimA exhibits weak mechanical stability albeit displaying β-sheet topology expected to unfold at much higher forces. Notably, PimA unfolds following heterogeneous multiple-step mechanical unfolding pathways at low force, akin to molten globule states. Interestingly, the *ab initio* low-resolution envelopes obtained from small-angle X-ray scattering of the unliganded PimA and the PimA-GDP complexed forms clearly demonstrate that not only the 'open' and 'closed' conformations of the GT-B enzyme are largely present in solution but, in addition, PimA experiences remarkable flexibility that undoubtedly corresponds to the N-terminal 'Rossmann fold' domain, which has been proved to participate in protein-membrane interactions. Based on these results and on our previous experimental data we propose a model where the conformational transitions are important for the mannosyltransferase to interact with the donor and acceptor substrates/membrane.

Glycosyltransferases (GTs) play a central role in nature. GTs catalyze the stereo- and regiospecific transfer of a sugar moiety from nucleotide-sugar or lipid-phospho-sugar donors to a wide range of acceptor substrates including mono-, oligo- and

polysaccharides, lipids, proteins, small organic molecules and deoxyribonucleic acids (1). As a consequence, GTs generate a significant amount of structural diversity in biological systems. This structural information is particularly apparent not only in the maintenance of the structural integrity of the cell but also in the modulation of molecular recognition events including cell-signaling, cell-cell and cell-pathogens interactions (2). Glycosyl transfer can proceed with either 'inversion' or 'retention' of the anomeric configuration with respect to the sugar donor. 'Inverting' GTs seem to follow a single-displacement mechanism with an oxocarbenium ion-like transition state and an asynchronous SN2 mechanism, analogous to that observed for inverting glycosyl hydrolases (1). In contrast, the catalytic mechanism for 'retaining' enzymes is actually a matter of strong debate. 'Retaining' GTs are believe to follow the 'internal return-like' mechanism, in which leaving group departure and nucleophilic attack occur on the same face of the sugar (3), involving either a short-lived oxocarbenium ion intermediate (SNi-like; 4) or an oxocarbenium ion transition state (SNi; 5). However recent reports suggest the presence of a covalent intermediate in the α3-glycosyltransferase and the human blood group synthesizing α-(1→3)-N-acetylgalactosaminyltransferase (GTA) and α-(1→3)-galactosyltransferase (GTB) (6, 7). Whether or not 'retaining' GTs could proceed via different catalytic mechanisms is a notion that will certainly need further experimental support. Importantly, to achieve such enzyme-transition state complexes, a spatial rearrangement of the active site or more distant regions of the GTs would be often required. However to date, the impact of substrate and enzyme-dynamics and conformational changes in GT-mediated catalysis is largely unknown (1).

Phosphatidyl-*myo*-inositol mannosyltransferase A (PimA) is a 'retaining' membrane-associated GT that initiates the biosynthetic pathway of essential phosphatidyl-*myo*-inositol mannosides (PIMs), lipomannan (LM) and lipoarabinomannan (LAM) in mycobacteria (8, 9, 10). PIM, LM and LAM are considered not only essential structural components of the cell but also important molecules implicated in host-pathogen interactions (10, 11, 12, 13). PimA catalyzes the transfer of a mannose residue from guanosine-5'-diphosphate-mannose (GDP-Man*p*) to the 2-position of phosphatidyl-*myo*-inositol (PI) to form PI monomannoside (PIM_1) on the cytoplasmic side of the plasma membrane. The crystal structure of PimA from *Mycobacterium smegmatis* has been recently solved in complex with both the donor substrate GDP-Man*p,* and GDP, one of the reaction products. Both structures superimpose well (r.m.s.d. of 0.3Å for 361 identical residues), and reveal that the enzyme displays the typical GT-B-fold of GTs, one of the two major structural folds described for the nucleotide sugar-dependent enzymes (14; Figure 1A). This fold was first described for the 351-amino acid DNA-modifying β -glucosyltransferase from family GT63, an inverting GT from bacteriophage T4, and was found to be structurally related to the catalytic core of glycogen phosphorylase (15, 16, 17, 18). The GT-B fold displays two 'Rossmann-fold' domains separated by a deep cleft that includes the catalytic center. Therefore, an important interdomain movement has been predicted/demonstrated in some members of this superfamily during substrate binding and catalysis including MurG (19), glycogen synthase (20, 21, 22) and MshA (23). It is generally accepted that in GT-B enzymes, the nucleotide-sugar donors mainly bind to the C-terminal domain of the protein, whereas the N-terminal domain is involved in acceptor substrate recognition. As acceptors exhibit a markedly diversity of chemical structures compared to nucleotide-sugar donors, the N-terminal domains reflect this variability by showing different rearrangements of secondary structural elements (24). On the basis of primary sequence homology analysis a GPGTF (glycogen phosphorylase/glycosyltransferase family) motif have been suggested to be present in many GT-B enzymes (25, 18). However, GT-B enzymes do not seem to share any strictly conserved residue (26).

Structural, biochemical and biophysical evidence indicate PimA undergoes significant conformational changes upon substrate binding and catalysis, involving loop flexibility and domain motions (14, 27). In particular, the binding of the sugar donor substrate, GDP-Man*p*, triggers an important interdomain rearrangement from an 'open' to a 'closed' state that stabilizes the enzyme and markedly increases its affinity for the acceptor substrate, PI. According to the crystal structures of the PimA-GDP and PimA-GDP-Man*p* complexes, the interaction of the enzyme with the β-phosphate of GDP resulted to be essential for this conformational change to occur. 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. In contrast, the binding of PI to the enzyme had the opposite effect, inducing the formation of a

more relaxed complex with PimA (27). Membrane association seems to be mediated by a combination of positively charged residues on the amphipathic α-helix 2, and surface-exposed hydrophobic residues on the N-terminal domain of the protein (14, 27).

This study describes a detailed investigation of the mechanical properties of PimA by using single-molecule force spectroscopy techniques. In combination with the solution structure of PimA by small-angle X-ray scattering we propose a model for substrate binding wherein protein flexibility and conformational transitions play a prominent role.

EXPERIMENTAL PROCEDURES

Methods — Recombinant PimA from *M. smegmatis* (*Ms*PimA, MSMEG_2935) was produced in *Escherichia coli* and purified to apparent homogeneity as previously described (27).

$I27_2$-PimA-$I27_2$ polyprotein production — The polyprotein $I27_2$-PimA-$I27_2$ was generated by using recombinant DNA technology based in an iterative process of digestion and ligation of DNA fragments, as previously described (28). The *pimA* gene from *M. smegmatis* was amplified by standard PCR using oligonucleotide primers pimA_BamHI_Fwd (5'-CGCGGATCCGGCTCTGGCGCGATGCGTATCGG GATGGTCTGCC-3') and pimA_SacI_Rev (5'-ATAATAGAGCTCTTAGCAACAAGATCTGACCG ATTCTCCGGCCGTCTG-3'), Phusion DNA polymerase (Finnzymes, Thermo Fisher Scientific) and the plasmid pET29a-MspimA as the DNA template (27). The PCR fragment was digested with BamHI and SacI, and ligated to the vector pT7-$I27_2$ harboring two consecutive copies of the immunoglobulin (Ig) domain of the giant muscle protein titin, I27. The resulting plasmid, pT7-$I27_2$-PimA, was further digested with BglII and SacI and the fragment LF-pT7-$I27_2$-PimA purified. The pT7-$I27_2$ was digested with BamHI and SacI and the fragment SF-$I27_2$ purified. Both fragments LF-pT7-$I27_2$-PimA and SF-$I27_2$ were ligated and the resulting construct named pT7-$I27_2$-PimA-$I27_2$. Digestion of pT7-$I27_2$-PimA-$I27_2$ with BamHI and SacI allowed in-frame cloning of the $I27_2$-PimA-$I27_2$ into the expression vector pQE80L (Qiagen, Chatsworth, CA). As depicted in Figure 1S, the full-length construct resulted in the following amino acid additions: (i) the N-terminal sequence is Met-Arg-Gly-Ser-$(His)_6$-Gly-Ser-I27; (ii) the junction between the I27 domains (BamHI-BglII hybrid site) is Arg-Ser; (iii) the junction between the second I27 domain and PimA is Arg-Ser-Gly-Ser-Gly-Ala; (iv) the junction between PimA and the third I27 domain is Arg-Ser. In addition, the resulting construct contained two Cys residues at the C-terminus allowing the covalent attachment of the polyprotein to the gold-covered slips. *E. coli* BL3 cells transformed with pQE80L-$I27_2$-PimA-$I27_2$ were grown in 400 ml LB medium supplemented with 30 μg/ml kanamycin and 50 μg/ml carbenicillin at 37°C. When the culture reached an OD_{600} value of 0.6, $I27_2$-PimA-$I27_2$ expression was induced by adding 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; MIP). After overnight incubation at 25 °C, cells were harvested and resuspended in 16 ml of 50 mM Tris-HCl pH 8.0 (solution A) containing protease inhibitors (Complete EDTA-free; Roche). Cells were then disrupted by sonication (five cycles of 1 min each) and the suspension was centrifuged for 20 min at 10000 g. The supernatant was applied onto Talon resin column (Clontech, Mountain View, CA), followed by a gel filtration step in a Superdex 200 column (GE Healthcare) equilibrated in 10 mM Hepes pH 7.2, 1 mM EDTA, 150 mM NaCl (solution B). The resulting $I27_2$-PimA-$I27_2$ preparation displayed a single protein band when run on a 12% NuPAGE® Bis-Tris precast gel stained with SimplyBlue™ SafeStain (Invitrogen). The purified $I27_2$-PimA-$I27_2$ polyprotein was stored at -80°C for further use in single molecule experiments.

I27-PimA protein production — The *pimA* gene from *M. smegmatis* was amplified as described for the $I27_2$-PimA-$I27_2$ construct. The PCR fragment was digested with BamHI and SacI, and then ligated to the vector pT7-I27 harboring one copy of the I27 module. The resulting plasmid, pT7-I27-PimA, was further digested with BamHI and SacI and ligated to the expression vector pQE80L (Qiagen, Chatsworth, CA). As depicted in Figure 2S, the full-length construct resulted in the following amino acid additions: (i) the N-terminal sequence is Met-Arg-Gly-Ser-$(His)_6$-Gly-Ser-I27; (ii) the junction between the I27 domain and PimA is Arg-Ser-Gly-Ser-Gly-Ala. In addition, the final construct contained two Cys residues at the C-terminus. *E. coli* DH5α cells (Invitrogen) transformed with pQE80L-I27-PimA were grown in 3000 ml of 2xYT medium supplemented with 50 μg/ml carbenicillin at 37°C. When the culture reached an OD_{600} value of 0.6, I27-PimA expression was induced by adding 1 mM IPTG (MIP). After overnight incubation at 18 °C, cells were harvested and

resuspended in 40 ml of solution A containing protease inhibitors (Complete EDTA-free; Roche). Cells were then disrupted by sonication (five cycles of 1 min each) and the suspension was centrifuged for 20 min at 10000 g. The supernatant was applied onto HisTrap column (GE Healthcare), followed by a gel filtration step in a Superdex 75 column (GE Healthcare) equilibrated in buffer 50 mM Tris-HCl pH 7.5, 150 mM NaCl (solution C). The resulting I27-PimA preparation displayed a single protein band when run on a 12% NuPAGE® Bis-Tris precast gel stained with SimplyBlue™ SafeStain (Invitrogen). The purified I27-PimA fusion protein was stored at 4°C and then concentrated for SAXS measurements by using a Vivaspin 20 spin concentrator (Vivascience) with a 10-kDa molar mass (MM) cutoff. The homogeneity of I27-PimA was confirmed by dynamic light scattering (DLS; data not shown).

Single Molecule Force Spectroscopy Measurements — 5-15 µL of $I27_2$-PimA-$I27_2$ from the most concentrated fraction obtained after gel filtration were deposited onto an evaporated gold coverslip. Veeco silicon nitride MLCT (Bruker, Camarillo, CA) cantilevers were mounted in a fluid cell carrier. The laser beam was centered at the tip of the cantilever and carefully calibrated. Spring constants around 15 pN/nm were obtained using the equipartition theorem (29). The cantilever was kept in contact with the gold slip for 1–2 seconds before retraction. We used a custom-built atomic force microscopy (AFM) setup that allows a strict control of the force based on an electronic feedback. Experiments were performed at room temperature in 10 mM Hepes pH 7.2, 1 mM EDTA, 150 mM NaCl. Only traces displaying 4 steps corresponding to I27 unfolding were considered. Due to the low stability of the construct $I27_2$-PimA-$I27_2$, single molecule experiments were performed within 3-4 days after purification. During that time, the purified $I27_2$-PimA-$I27_2$ polyprotein construct demonstrated to conserve fully enzymatic activity *in vitro* (8; data not shown). To calculate the contour length of the pulled polypeptide, we applied the Worm Like Chain (WLC) model with a persistence length of 0.4. Hence, the Worm Like Chain curves in the Figure 3 corresponds to the theoretical extension of the N- and C-terminal Rossmann-fold domains, respectively 163 and 119 amino acids, and I27. For a more accurate estimation of the contour length, the length of the folded N- and C-terminal Rossmann-fold domains (1.8nm and 2.3nm respectively) was subtracted. WLC was fitted in force extension curves as well to estimate the initial extension before the rupture of the first I27 domain (30).

Small Angle X-ray Scattering Measurements — Synchrotron X-ray diffraction data for recombinant purified *Ms*PimA *apo* and in the presence of GDP (27) were collected on a pixel Pilatus 1M detector at the EMBL X33 beamline (DESY, Hamburg; 31, 32), whereas data for the fusion I27-PimA protein were collected at EMBL P12 beamline (DESY, Hamburg) on a pixel Pilatus 2M detector. The scattering patterns were measured with a 2 min exposure time for protein samples at a minimum of 3 different protein concentrations ranging from 1-10 mg/ml at the X33 and with 1 sec exposure time at the P12. Samples were incubated with 2 mM dithiothreitol before data collection. To check for radiation damage, four 30 sec exposures were compared at the X33, and twenty 50 millisecond exposures at the P12, and no radiation damage was observed. Using the sample-to-detector distance of 2.7 m (3.1 m at the P12), the range of momentum transfer values is $0.009 < q < 0.5$ Å−1 at the X33 and $0.008 < q < 0.47$ at the P12 ($q = 4\pi\sin(\theta)/\lambda$, where 2θ is the scattering angle, and $\lambda = 1.5$ Å (1.24 Å for the P12 beamline) is the X-ray wavelength). Data was processed using standard procedures and extrapolated to infinite dilution by the program package PRIMUS (33). The forward scattering $I(0)$ was evaluated using the Guinier approximation (34) assuming the intensity is represented as $I(q) = I(0)\exp(-(qR_g)^2/3)$ for a very small range of momentum transfer values ($q < 1.3/R_g$). The maximum dimensions D_{max}, the interatomic distance distribution functions $P(r)$ and the radii of gyration R_g were computed using GNOM (35). The molecular mass of the protein was evaluated by comparison to the forward scattering with that from a reference solution of bovine serum albumin (66 kDa).

Ab initio shape determination — The low resolution structures of *Ms*PimA *apo* and *Ms*PimA-GDP complex were calculated *ab initio* by using the program DAMMIF (36). Structure clustering and averaging was carried out with the program DAMCLUST (37). The results and statistics are summarized in Table 1. For each dataset, pairwise alignment of structures is carried out and the Normalized Spatial Discrepancy (NSD) calculated (38). The clustering process groups structures with lower NSD discrepancy. In general, NSD values close to one indicate that the two

structures are similar. Structures that deviate by more than 2 are taken as outliers and removed. For all calculations, 3 outlier models were rejected. Average NSD values within clusters are between 0.5-7 indicating a very good structural agreement for each cluster. Furthermore, average NSD values for structures of different clusters are between 0.7-1.0, indicating similarity between clusters.

The *ab initio* reconstruction of the fusion protein I27-PimA was carried out by MONSA. MONSA is based on DAMMIN and allows bead modeling of macromolecular complexes by simultaneously fitting multiple experimental data (39, 40). The search volume was defined as a sphere of radius equal to half the maximum dimension of I27-PimA (D_{max}=120 Å) as computed by GNOM. Within the search volume, a grid of points is defined with a constant distance of 2 Å between points, resulting in dense packing of spheres. Each sphere within the search volume might correspond to one of the following components: solvent, I27 or PimA, represented by an integer numbers 0, 1 or 2. To facilitate computation and to keep physical sense with respect to the experimental data (i.e. low resolution of data and connectivity of each component), a number of conditions are applied during structural calculations, accounting for the likely connectivity of two arbitrarily selected atoms belonging to a given component, and its compactness. The minimization procedure fits the two experimental data sets, PimA and I27-PimA, to the corresponding calculated data obtained from computed structures. It is carried out by simulated annealing (41), starting from a random configuration X_0, where the assignment of a single bead to a given component is changed randomly (a move from X to X'). In the case the change produces an improvement of fitness ($\Delta = f(X') - f(X) < 0$), it is accepted. Otherwise, it can still be accepted with a probability of $\exp(-\Delta/T)$, where T is the annealing temperature, which decreases as the calculation progresses towards a minimum. The described approach allowed us to unambiguously differentiate the relative location of I27 and PimA within the fusion protein and to correctly orientate the N- and C-terminal domains of PimA as observed in the crystal structure, in the solution structures obtained by SAXS techniques.

Molecular Modeling – SITUS (42) was used to perform rigid body docking of high-resolution crystal structures into solution structures. To this end, the bead models obtained by SAXS techniques were transformed to 3D volumes using the pdb2vol program (part of the SITUS package), followed by exhaustive search docking, including all six degrees of freedom (three rotation angles and three translation coordinates). The fitting improvement is assessed by a cross-correlation coefficient function (CC). Values of CC range between 0 to 1, where 1 indicates excellent agreement between the 3D volume (obtained by SAXS) and the crystal structures (42).

RESULTS AND DISCUSSION

PimA displays an intrinsic low mechanical stability — Single-molecule force spectroscopy has been proven to be a powerful technique not only to understand the role of force in mechanical stability, but also to study protein conformational dynamics and enzyme catalysis (43, 44) For instance, applying a calibrated force to the substrate of an enzyme can probe the dynamics of the active site during catalysis with sub-Angstrom resolution (45). In order to study the mechanical properties of PimA under force, we engineered a polyprotein composed of a single PimA copy flanked on either side by two copies of an immunoglobulin (Ig) domain of the giant muscle protein titin; I27 ($I27_2$-PimA-$I27_2$) (see Experimental Procedures; 46).

We applied force to this polyprotein with an Atomic Force Microscope (AFM) as shown in Figure 1B. I27 modules provide an unambiguous mechanical fingerprint to identify single-molecule events. Two cysteines at the C-terminal allow a covalent binding of $I27_2$-PimA-$I27_2$ to a gold surface on top of a piezoelectric positioner (Figure 1B; 28). To calculate the expected initial extension, we considered the sequences linking the I27 domains and PimA in the polyprotein. The N-terminal before the first I27 domain contains 11 amino acids, whereas the C-terminal after the last I27 contains 4 amino acids. Twelve extra residues separate the I27 modules in the polyprotein (Figure 1S). Therefore, the initial extension of $I27_2$-PimA-$I27_2$, including the unfolding of the 384 amino acids of PimA, involves 384+11+12+4 = 411 amino acids and 4 folded I27 domains (4.4 nm each). Considering a contour length per amino acid of 0.4 nm (47), this corresponds to a maximal contour length of 182 nm. The polyprotein was first stretched at a constant rate of 400 nm/s. The opposing force exerted by the polyprotein was measured from the deflection of the AFM cantilever. The resulting force extension traces displayed an

initial extension region characterized by the presence of marginal unfolding peaks. The initial extension was followed by a sawtooth pattern of three or four consecutive unfolding events and a final detachment peak (Figure 2). The observed unfolding peaks of ~200 pN spaced by ~28-29 nm (number of events = 138) fit well to the Worm-Like Chain model of polymer elasticity, and correspond to the characteristic fingerprint of the I27 module (48). Based on the contour length of the first I27 peak, we determined the length of the initial extension as ~180 nm (Figure 1S and 2), which is in agreement with a fully extended PimA and sequences linking the I27 domains (182 nm, see above). The addition of 5 mM GDP gives similar curves without any distinct pic during the extension of PimA (data not shown). Altogether, these results demonstrate that $I27_2$-PimA-$I27_2$ is tethered from both ends and PimA is subject to force. It is worth noting that the addition of protein domains with known mechanical features has been successfully used to study new protein folds under force (46, 49). In the case of proteins without uniform and discernable mechanical response to force, this indispensable fingerprint guaranties that the initial extension is relevant to the pulled polyprotein (50). PimA seems to unfold at low force, without significant energy barriers limiting its extensibility, indicating that the mannosyltransferase displays an intrinsic weak mechanical resistance to force.

PimA unfolds following heterogeneous multiple-step mechanical unfolding pathways — In order to further characterize PimA unfolding, we used a force ramp protocol with a linearly increasing force (51). This protocol explores a wide range of forces and is therefore more suitable to examine and segregate different unfolding pathways. We increased the force from 0 pN to 250 pN over a time period of 5 s obtaining 50 well-defined traces in the presence and absence of GDP. Extension of $I27_2$-PimA-$I27_2$ was identified by the presence of four uniform I27 unfolding steps (Figure 3A and 3B; 50). For each distinguishable step appearing in force ramp, the step size and the force at which the event occurs are reported in the scatter-plots in figure 3. Remarkably, PimA unfolds following multiple steps at low force. The steps observed during the ramp are well defined and can be characterized by its size and force. However, the large dispersion in step sizes and unfolding forces for PimA unfolding reflects the heterogeneity of mechanical unfolding pathways, analogous to the mechanical unfolding of molten globule states (52, 53). Similar results were obtained in the presence of 5 mM GDP (dissociation constant K_d of 0.03 μM (ΔH=-14.0 kcal.mol^{-1}; Figure 2), which is a well-known thermal stabilizer of the enzyme. GDP binds tightly to the C-terminal domain through a network of hydrogen bonds and stabilize the enzyme by coupling the two N- and C-terminal lobes in a 'closed' state (27).

Figure 4 reveals a particular well-resolved trace illustrating a possible pathway in which the theoretical extension coincide with the one-step rupture of 163 amino acids forming the N-terminal β-sheet and the 119 amino acids forming the C-terminal β -sheet (Figure 4A). We propose that the first four shorter steps bear out the detachment of the C-terminal helices (yellow, Figure 4B). Then the linker and the N-terminal domain unfold in a two-step manner. The last step is compatible with the full extension of the C-terminal β-sheet. This single trace depicts one of the multiple routes that are available to fully extend PimA under force. Step sizes larger than the theoretical unfolding length of the two β-sheets would suggest the co-rupture of larger structural elements.

The mechanical properties of proteins are strongly influenced by their structural elements (28). To date, the strongest mechanical element corresponds to the β-sandwich topology characteristic of Ig and Ig-like proteins (54). As depicted in Figures 3 and 4, I27 unfolds in an all-or-none unfolding pathway identically to proteins with similar topology such as ubiquitin and protein L (51, 55). In all these proteins, two parallel β-strands linking the N- and C-terminus, firmly protect the majority of native contacts before unraveling the rest of the protein. As soon as the parallel β -strands interactions break, the remaining structural elements unfold without any force resistance. Indeed, mutations in this structural motif in Ig or Ig-like proteins markedly reduce mechanostability (56). Interestingly, different unfolding forces have been reported for homologous Ig domains. The presence of parallel β-strands motifs do not fully explain the mechanostability of I27 and other properties have to be considered (48).

PimA topology is remarkably more complex than that of I27. PimA shares with mechanically stable proteins the presence of β-sheets in both 'Rossmann-fold' domains (14) with numerous parallel β -strands (Figure 1 and 4), suggesting a strong mechanical stability. The N-terminal end folds into a β-strand (β1,

residues 1-7) and make strong interactions with β2 and β4 strands of the central β-sheet core of the first domain. In contrast, the C-terminal region terminates with the classical GT-B pattern of three long α-helices, α12 (residues 315-328), α13 (residues 330-347) and α14 (residues 349-363); with α14 crossing over to the N-terminal domain (Figure 1 and 4). This hinge region is completed by residues 165-190, which include α7. Because α-helices are mechanically weak elements (57, 58), it is expected all helices to unfold first at low force, allowing PimA to reach a conformation where the two 'Rossmann-fold' domains are separated. In this scenario, the GDP binding site is dislocated very early in the unfolding pathway. Therefore, when the force is applied to a PimA-GDP complex, the affinity for GDP is expected to drop following the mechanical disruption of the binding site. This model agrees with the observation that mechanical unfolding of PimA in the presence or absence of GDP gives rise to similar distributions of events (Figure 3B).

Once the hinge has extended, parallel β strands in both 'Rossmann-fold' domains are exposed to force (Figure 4B). This situation induced by force is equivalent to native mechanostable proteins such as I27. However, contrary to I27, PimA unfolds with remarkable wide distribution of step sizes at low force, where no common structure dominates. Our results suggest that the presence of parallel β-strands is not sufficient to explain the mechanical properties of PimA, where the presence of marginal stable structural elements within the protein allows the 'Rossmann-fold' domains to unfold along a variety of pathways. We propose that the partial unfolding of helical regions and the hinge at low force leads to a marked perturbation of PimA energy landscape, where a significant number of native contacts are disrupted before the two β-sheets are under force.

The 'open' and 'closed' states of PimA as visualized by SAXS — To gain insight into the conformational changes occurred in PimA, the solution structure of the enzyme in its unliganded and GDP bound forms were solved by small-angle X-ray scattering (SAXS). SAXS proved to be a powerful technique capable of providing structural information of flexible and dynamic proteins in solution (59). The interatomic distance distribution function $P(r)$ computed for the PimA *apo* and in the presence of GDP is shown in Figure 5. The molecular mass determined from the scattering data confirmed that the protein is monomeric in solution, in agreement with size exclusion chromatography and analytical ultracentrifugation experiments (see Experimental Procedures; Table 1; 14, 27). It is worth noting that the protein remained in a monomeric state even in the presence of GDP.

The radii of gyration (R_g) obtained for PimA *apo* and the PimA-GDP complex revealed a reduction in R_g (ΔR_g) of −1.0(1) Å, indicating that GDP binding leads to a compaction of PimA, clearly supporting the existence of the 'open' and 'closed' states of the enzyme (Table 1; 27). Several lines of experimental evidence provide strong confidence in this model. The relevance of the β-PO_4 in the stabilization of the closed conformation of PimA, was first studied by limited proteolysis (27). PimA was rapidly degraded after incubation with elastase. N-terminal sequencing of the two predominant species of 23 and 15 kDa revealed two exposed sites located in α9 and the connecting loop β7-β8 at the junction between N- and C-terminal domains. The α9 contains two important residues involved in GDP-Man*p* recognition: Asp253 and Lys256. Interestingly, when PimA was incubated in the presence of GDP, the enzyme was protected from the action of elastase even after 90 min suggesting a conformational rearrangement. A close inspection of the crystal structure of the PimA-GDP complex revealed that the enzyme crystallizes in a 'closed' conformation with the active site buried between both Rossmann-fold domains (14). Two residues of the N-terminal domain were found to interact with GDP: Pro14, in the connecting loop β1-α1 stabilizes the guanine heterocycle by a van der Waals stacking interaction, whereas Gly16, at the top of α1, hydrogen bonds the β-PO_4 of GDP (14). Moreover, Arg196 and Lys202, largely conserved in PimA orthologues but located in the C-terminal domain, make electrostatic interactions with the β-PO_4, thus restricting its position into the catalytic site. Importantly, the α-PO_4 of GDP does not interact with any particular residue from the enzyme (14). Structural comparisons of the unliganded and nucleotide diphospho-sugar-bound forms of MshA from *Corynebacterium glutamicum* (*Cg*MshA) and the glycogen synthase from *Agrobacterium tumefaciens* (*At*GS) and *E. coli* (*Ec*GS) revealed that a substantial subdomain rotation is required to achieve the 'closed' state (20, 21, 23). The residues Gly23, Arg231, and Lys236 in *Cg*MshA, Gly18, Arg299, and Lys304 in *At*GS and Gly18, Arg300 and Lys305 in *Ec*GS interact with the β-PO_4 of uridine 5-diphosphate (UDP) and

adenosine 5-diphosphate (ADP), respectively, being structurally equivalent to Gly16, Arg196, and Lys202 in the 'closed' conformation of *Ms*PimA.

The 'open to closed' motion has also been monitored in the absence and presence of GDP and guanosine by sedimentation velocity analytical ultracentrifugation studies on pure *Ms*PimA (27). The unliganded form of the enzyme sedimented as a single homogeneous species with an average sedimentation coefficient of 3.22 s, which is consistent with an asymmetric monomeric protein. Upon the addition of equimolar GDP, the sedimentation coefficient increased to 3.53 s, indicating the formation of a more symmetrical and compact structure. The addition of guanosine, where the α-PO_4 and β-PO_4 are missing, did not significantly affect the sedimentation coefficient value of the mannosyltransferase, supporting the requirement of the β-PO_4 to trigger the 'closed' conformation. Moreover, ITC measurements revealed that guanosine bound to PimA with a binding constant ~10^3-fold smaller than that of GDP. The GDP binding also produce a stabilizing effect on *Ms*PimA characterized by an increment of 3.5 ^oC in the melting temperature value as observed by differential scanning calorimetry and circular dicroism (27). In summary, GDP induces the closing movement of PimA, with a clear increment of the sedimentation coefficient value which is correlated with a reduction of the radii of gyration, and accompanied by a markedly stabilization of the enzyme.

Conformational flexibility of PimA — The *ab initio* low-resolution calculated envelopes for PimA *apo* and PimA-GDP complex displayed the characteristic two-domain organization of the GT-B enzymes with the deep fissure between the N- and C-terminal lobes including the catalytic center (see Experimental Procedures; Table 1, Figure 6 and 7). However, a comparison of the crystal structure of PimA-GDP complex with the solution structures obtained by SAXS revealed significant differences in the conformation of one of the 'Rossmann-fold' domains of the protein. To unambiguously identify the N- and C-terminal domains in the *ab initio* models, SAXS data was collected on a I27-PimA polyprotein, where the 10 kDa I27 module was fused to the N-terminal end of *Ms*PimA (see Table 1 and experimental section for details). The experimental data clearly demonstrates the position of the I27 protein in close proximity to a protruding lobe on the N-terminal domain (Table 1; Figure 6). It is worth noting that the different structures depicted in Figure 7 do not directly prove the flexibility of PimA, but rather stem from the potential ambiguity of *ab initio* reconstructions in SAXS. Nevertheless they provide inside into the possible structural flexibility depicting possible protein conformations compatible with the scattering data.

The N-terminal domain of PimA has been proved to participate in acceptor recognition, involving important conformational changes in the protein (14, 27). Specifically, when incubated with PI, both the unliganded and the GDP bound forms of the enzyme became highly sensitive to elastase treatment, indicating that PI triggers a significant conformational change able to modify the 'closed' GDP-induced conformation. Furthermore, analytical ultracentrifugation experiments demonstrated that the addition of PI to the enzyme resulted in a significant change in the sedimentation coefficient values of both PimA *apo* and PimA-GDP complex consistent with the formation of less compact structures, which correlates with a reduction of the melting temperature by 1.5 and 0.4 ^oC, respectively. The crystal structure of the PimA-GDP complex displayed two regions on the N-terminal domain comprising residues 118 to 140 and 149 to 165, with particularly high B-factor values, suggesting structural flexibility. These residues are located in three α-helices, α4, α5 and α6, and primarily define a hydrophobic tunnel, which is compatible with the binding of fatty acyl chains of PI or phospholipids of the mycobacterial plasma membrane. Thus, this particular region of PimA could certainly experience conformational flexibility, and might explain the presence of the extra lobe in the SAXS structure.

The conformation of membrane-associated proteins can be certainly affected by binding to the lipid bilayer, thus modulating protein function. These conformational changes can be quite different in nature ranging from subtle modifications such as the reorientation of lateral chain residues to more complex structural rearrangements (60). For example, the orientation of transmembrane domains in polytopic integral membrane-associated proteins has been proved to be dependent on membrane phospholipid composition during initial assembly as well as on changes in lipid composition post-assembly (61). A dramatic conformational change has also been reported for the pore forming toxin pneumolysin where two of the four domains refold along with the deformation of the membrane (62). Interestingly, recent studies suggest that the molten globule states

play an important role in cellular processes that involve membrane interaction. The translocation of the N-terminal domain of the anthrax lethal factor across the membrane requires mechanical unfolding of a molten globule-like state (63). Moreover, the interaction of the peripheral membrane-associated α-lactalbumin to phospholipid bilayers of various compositions revealed the membrane bound state to resemble a molten globule-like state (64). It is worth noting that the amphipathic helix C resulted to be a critical determinant for the adsorption and further integration of the α-lactalbumin into the membrane. In that sense, PimA displays an amphipathic helix on its N-terminal domain, which has been proved to be essential for membrane binding and activity *in vitro* and *in vivo* (27). The occurrence of metastable states on PimA might be reflective of conformational variability of the enzyme, restricted to the N-terminal domain, and the association to the acceptor substrate and/or lipid bilayer. However, further studies will be required to define in detail the residues responsible for the observed structural variability.

CONCLUDING REMARKS

Remarkable progress has been made in recent years in our understanding of the catalytic mechanism and structural basis of glycosyl transfer. However, the study of the conformational changes and dynamics that govern substrate recognition and catalysis remains a major challenge in the field of GTs. The combination of single-molecule force spectroscopy and SAXS techniques outline the high variability of the PimA *apo* and GDP bound states. The structural data indicate that the 'open' and 'closed' conformations of the enzyme are largely present in solution but, in addition, PimA experiences remarkable flexibility that undoubtedly correspond to the N-terminal 'Rossmann fold' domain of the protein. This domain has been proved to participate in acceptor/membrane interaction. Unfolding of PimA under force leads to heterogeneous unfolding trajectories that seem to be reflective of its intrinsic dynamic nature. According to our previously reported and current experimental data we propose a model wherein the flexibility and conformational transitions confer adaptability of PimA to the substrates and the membrane, which seems to be of importance during the catalytic cycle.

Finally, protein dynamics and conformational changes are known to be critically important in drug discovery and development strategies (65). Although GTs play a central role in a broad range of human diseases, there is a clear lack of potent and specific GT inhibitors (66). The information presented herein highlights the conformational plasticity of PimA, information that should be considered in the rational design of GTs inhibitors.

REFERENCES

1. Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) Glycosyltransferases. Structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77**, 521–5552.
2. Rini, J., Esko, J., and Varki, A. (2009) in Essentials of Glycobiology (Varki, A., Cummings, R., Esko, J., Freeze, H., Stanley, P., Bertozzi, C. R., Hart, G., and Etzler, M. E., eds) 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
3. Lee, S. S., Hong, S. Y., Errey, J. C., Izumi, A., Davies, G. J., and Davis, B. G. (2011) Mechanistic evidence for a front-side, SNi-type reaction in a retaining glycosyltransferase. *Nat. Chem. Biol.* **7**, 631–638.
4. Ardévol, A., and Rovira, C. (2011) The molecular mechanism of enzymatic glycosyl transfer with retention of configuration. Evidence for a short lived oxocarbenium-like species. *Angew. Chem. Int. Ed. Engl.* **50**, 10897–10901.
5. Gómez, H., Polyak, I., Thiel, W., Lluch, J. M., and Masgrau, L. (2012) Retaining glycosyltransferase mechanism studied by QM/MM methods. Lipopolysaccharyl-α-1,4-galactosyltransferase C transfers α-galactose via an oxocarbenium ion-like transition state. *J. Am. Chem. Soc.* **134**, 4743–47526.
6. Monegal, A. and Planas, A. (2006) Chemical rescue of α **3**galactosyltransferase. Implications in the mechanism of retaining glycosyltransferases. *J. Am. Chem. Soc.* **128**, 16030–16031.
7. Soya, N., Fang, Y., Palcic, M. M., Klassen, J. S. (2011) Trapping and characterization of covalent intermediates of mutant retaining glycosyltransferases. *Glycobiology* **21**, 547-52.
8. Kordulakova, J., Gilleron, M., Mikusova, K., Puzo, G., Brennan, P. J., Gicquel, B., and Jackson, M. (2002) Definition of the first mannosylation step in phosphatidylinositol mannoside synthesis. PimA is essential for

growth of mycobacteria. *J Biol Chem.* **277**, 31335-31344.
9. Guerin, M. E., Kaur, D., Somashekar, B. S., Gibbs, S., Gest, P., Chatterjee, D., Brennan, P. J., Jackson, M. (2009) New insights into the early steps of phosphatidylinositol mannoside biosynthesis in mycobacteria: PimB' is an essential enzyme of *Mycobacterium smegmatis*. *J. Biol. Chem.* **284**, 25687-25696.
10. Guerin, M. E.; Kordulakova, J.; Alzari, P. M.; Brennan, P. J.; Jackson, M. (2010) Molecular basis of phosphatidyl-myo-inositol mannoside biosynthesis and regulation in mycobacteria. *J. Biol. Chem.* **285**, 33577-33583.
11. Morita, Y. S, Fukuda, T., Sena, C. B., Yamaryo-Botte, Y., McConville, M. J., Kinoshita, T. (2011) Inositol lipid metabolism in mycobacteria: biosynthesis and regulatory mechanisms. *Biochim Biophys Acta.* **1810**, 630-41.
12. Cala-De Paepe, D., Layre, E., Giacometti, G., Garcia-Alles, L. F., Mori, L., Hanau, D., de Libero, G., de la Salle, H., Puzo. G., Gilleron, M. (2012) Deciphering the role of CD1e protein in mycobacterial phosphatidyl-myo-inositol mannosides (PIM) processing for presentation by CD1b to T lymphocytes. *J Biol Chem.* **287**, 31494-502.
13. Mishra, A. K., Alves, J. E., Krumbach, K., Nigou, J., Castro, A. G., Geurtsen, J., Eggeling, L., Saraiva, M., Besra, G. S. (2012) Differential arabinan capping of lipoarabinomannan modulates innate immune responses and impacts T helper cell differentiation. *J. Biol. Chem.* **287**, 44173-83.
14. Guerin, M. E., Kordulakova, J., Schaeffer, F., Svetlikova, Z., Buschiazzo, A., Giganti, D., Gicquel, B., Mikusova, K., Jackson, M., Alzari, P. M. (2007) Molecular recognition and interfacial catalysis by the essential phosphatidylinositol mannosyltransferase PimA from mycobacteria. *J. Biol. Chem.* **282**, 20705-20714.
15. Vrielink, A., Ruger, W., Driessen, H. P., Freemont, P. S. (1994) Crystal structure of the DNA modifying enzyme β-glucosyltransferase in the presence and absence of the substrate uridine diphosphoglucose. *EMBO J.* **13**, 3413–3422.
16. Barford, D., Johnson, L. N. (1989) The allosteric transition of glycogen phosphorylase. *Nature* **340**, 609–616.
17. Artymiuk, P. J., Rice, D. W., Poirrette, A. R., Willett, P. (1995) β-glucosyltransferase and phosphorylase reveal their common theme. *Nat. Struct. Biol.* **2**, 117 - 120.
18. Wrabl, J. O., Grishin, N. V. (2001) Homology between O-linked GlcNAc transferases and proteins of the glycogen phosphorylase superfamily. *J. Mol. Biol.* **314**, 365-374.
19. Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., Mokhtarzadeh, M., Walker, S. (2003) Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases. *Proc. Natl. Acad. Sci. U S A.* **100**, 845-849.
20. Buschiazzo, A., Ugalde, J. E., Guerin, M. E., Shepard, W., Ugalde, R. A., Alzari, P. M. (2004) Crystal structure of glycogen synthase: homologous enzymes catalyze glycogen synthesis and degradation. *EMBO J.* **23**, 3196-205.
21. Sheng, F., Jia, X., Yep, A., Preiss, J., Geiger, J. H. (2009) The crystal structures of the open and catalytically competent closed conformation of *Escherichia coli* glycogen synthase. *J. Biol. Chem.* **284**, 17796-17807.
22. Baskaran, S., Roach, P. J., Depaoli-Roach, A. A., Hurley, T. D. (2010) Structural basis for glucose-6-phosphate activation of glycogen synthase. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 17563-17568.
23. Vetting, M. W., Frantom, P. A., Blanchard, J. S. (2008) Structural and enzymatic analisis of MshA from *Corynebacterium glutamicum*: substrate-assisted catalysis. *J. Biol. Chem.* **283**, 15834–15844.
24. Breton, C., Snajdrová, L., Jeanneau, C., Koca, J., Imberty, A. (2006) Structures and mechanisms of glycosyltransferases. *Glycobiology* **16**, 29R-37R.
25. Abdian, P. L., Lellouch, A. C., Gautier, C., Ielpi, L., Geremia, R. A. (2000) Identification of essential amino acids in the bacterial alpha-mannosyltransferase AceA. *J. Biol. Chem.* **275**, 40568-40575.
26. Hu, Y., Walker, S. (2002) Remarkable structural similarities between diverse glycosyltransferases. *Chem. Biol.* **9**, 1287-1296. Review.
27. Guerin, M. E.; Schaeffer, F.; Chaffotte, A.; Gest, P.; Giganti, D.; Kordulakova, J.; van der Woerd, M.; Jackson, M.; Alzari, P. M. (2009) Substrate-induced conformational changes in the essential peripheral membrane-associated mannosyltransferase PimA from mycobacteria. Implications for catalysis. *J. Biol. Chem.* **284**, 21613-21625.

28. Carrion-Vazquez, M., Oberhauser, A. F., Fowler, S. B., Marszalek, P. E., Broedel, S. E., Clarke, J., Fernandez, J. M. (1999) Mechanical and chemical unfolding of a single protein: a comparison. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3694-3699.
29. Florin, E. L., Rief, M., Lehmann, H., Ludwig, M., Dornmair, C., Moy, V. T., and Gaub, H.E. (1995) Sensing specific molecular interactions with the atomic force microscope. *Biosen. Bioelectron.* **10**, 895-901.
30. Bustamante, C., Marko, J.F., Siggia, E.D. and Smith S. Entropic elasticity of lambda-phage DNA. (1994) *Science* **265**, 1599-1600.
31. Roessle, M. W., Klaering, R., Ristau, U., Robrahn, B., Jahn, D., Gehrmann, T., Konarev, P., Round, A., Fiedler, S., Hermesa, C., and Svergun, D. (2007) Upgrade of the small-angle X-ray scattering beamline X33 at the European Molecular Biology Laboratory, Hamburg. *J. Appl. Cryst.* **40**, S190–S194.
32. Round, A.R., Franke, D., Moritz, S., Huchler, R., Fritsche, M., Malthan, D., Klaering, R., Svergun, D.I. & Roessle, M. (2008) Automated sample-changing robot for solution scattering experiments at the EMBL Hamburg SAXS station X33. *J. Appl. Cryst.* **41**, 913-917.
33. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. and Svergun, D. I. (2003) *J. Appl. Cryst.* **36**, 1277-1282.
34. Guinier, A. (1939) La diffraction des rayon's X aux trés petits angles: application a l'étude de phénoménes ultramicroscopies. *Ann. Phys.* (Paris) **12**, 161-237.
35. Svergun, D.I. (1992) Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria. *J. Appl. Cryst.* **25**, 495-503.
36. Franke, D. and Svergun, D.I. (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Cryst.* **42**, 342-346.
37. Volkov, V. V., and Svergun, D. I. (2003). Uniqueness of ab-initio shape determination in small-angle scattering. *J. Appl. Cryst.* **36**, 860-864.
38. Kozin, M. B., and Svergun, D. I. (2001) Automated matching of high- and low-resolution structural models. *J. Appl. Cryst.* **34**, 33-41.
39. Svergun, D. I. (1999) Restoring low-resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J.* **76**, 2879-86.
40. Svergun, D.I., and Nierhaus, K. H. (2000) A map of protein-rRNA distribution in the 70 S *Escherichia coli* ribosome. *J. Biol. Chem.* **275**, 14432-9.
41. Kirkpatrick, S., Gelatt, C. D. Jr, Vecchi, M. P. (1983) Optimization by simulated annealing. Science **220**, 671-80.
42. Wriggers, W. (2012) Conventions and workflows for using Situs. *Acta Cryst.* **D68**, 344-351.
43. Larson, M. H., Landick, R., and Block, S. M. (2011) Single-molecule studies of RNA polymerase: one singular sensation, every little step it takes. *Mol. Cell* **41**, 249-262.
44. Alegre-Cebollada, J., Perez-Jimenez, R., Kosuri, P., Fernandez, J.M. Single-molecule force spectroscopy approach to enzyme catalysis. (2010) *J. Biol. Chem.* **285**, 18961-6.
45. Wiita, A. P., Perez-Jimenez, R.; Walther, K. A., Grater, F., Berne, B. J., Holmgren, A., Sanchez-Ruiz, J. M., Fernandez, J. M. (2007) Probing the chemistry of thioredoxin catalysis with force. *Nature* **450**, 124-127.
46. Li, H., Oberhauser, A. F., Redick, S. D., Carrion-Vazquez, M., Erickson, H. P., Fernandez, J. M. (2001) Multiple conformations of PEVK proteins detected by single-molecule techniques. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10682-10686.
47. Ainavarapu, S. R., Brujic, J., Huang, H. H., Wiita, A. P., Lu, H., Li, L., Walther, K. A., Carrion-Vazquez, M., Li, H., Fernandez, J. M. Contour length and refolding rate of a small protein controlled by engineered disulfide bonds. (2007) *Biophys. J.* **92**, 225-233.
48. Li, H., Linke, W. A., Oberhauser, A. F., Carrion-Vazquez, M., Kerkvliet, J. G., Lu, H., Marszalek, P. E., Fernandez, J. M. (2002) Reverse engineering of the giant muscle protein titin. *Nature* **418**, 998-1002.
49. Ainavarapu, S. R., Li, L., Badilla, C. L., Fernandez, J. M. (2005) Ligand binding modulates the mechanical stability of dihydrofolate reductase. *Biophys. J.* **89**, 3337-3344.
50. Alegre-Cebollada, J., Badilla, C. L., Fernandez, J. M. (2010) Isopeptide bonds block the mechanical extension of pili in pathogenic *Streptococcus pyogenes*. *J. Biol. Chem.* **285**, 11235-11242.
51. Schlierf, M., Li, H., Fernandez, J. M. (2004) The unfolding kinetics of ubiquitin captured with single-

molecule force-clamp techniques. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7299-7304.
52. Garcia-Manyes, S., Dougan, L., Badilla, C. L., Brujic, J., Fernandez, J. M. (2009) Direct observation of an ensemble of stable collapsed states in the mechanical folding of ubiquitin *Proc. Natl. Acad. Sci. U.S.A.* **106**, 10534-10539.
53. Elms PJ, Chodera JD, Bustamante C, Marqusee S. (2012) The molten globule state is unusually deformable under mechanical force. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 3796-801.
54. Valbuena, A., Oroz, J., Hervas, R., Vera, A.M., Rodriguez, D., Menendez, M., Sulkowska, J.I., Cieplak, M., Carrion-Vazquez, M. On the remarkable mechanostability of scaffoldins and the mechanical clamp motif. (2009) *Proc Natl Acad Sci U S A* **106**, 13791-6.
55. Liu, R., Garcia-Manyes, S., Sarkar, A., Badilla, C. L., Fernandez, J. M. (2009) Mechanical characterization of protein L in the low-force regime by electromagnetic tweezers/evanescent nanometry. *Biophys. J.* **96**, 3810-3821.
56. Li, H., Carrion-Vazquez, M., Oberhauser, A. F., Marszalek, P. E., Fernandez, J. M. (2000) Point mutations alter the mechanical stability of immunoglobulin modules. *Nat. Struct. Biol.* **7**, 1117-1120.
57. Law, R., Carl, P., Harper, S., Dalhaimer, P., Speicher, D. W., Discher, D. E. (2003) Cooperativity in forced unfolding of tandem spectrin repeats. *Biophys. J.* **84**, 533-544.
58. Rief, M.; Pascual, J.; Saraste, M.; Gaub, H. E. (1999) Single molecule force spectroscopy of spectrin repeats: low unfolding forces in helix bundles. *J. Mol. Biol.* **286**, 553-561.
59. Blanchet, C. E., and Svergun, D. I. (2013) Small-angle X-ray scattering on biological macromolecules and nanocomposites in solution. *Annu Rev Phys Chem.* **64**, 37-54.
60. Seelig, J. (2004) Thermodynamics of lipid–peptide interactions. *Biochim. Biophys. Acta.* **1666**, 40-50.
61. Dowhan, W., Bogdanov, M. (2009) Lipid-dependent membrane protein topogenesis. *Annu. Rev. Biochem.* **78**, 515-40.
62. Tilley, S. J., Orlova, E. V., Gilbert, R. J.C., Andrew, P. W., and Saibil, H.R. (2005) Structural basis of pore formation by the bacterial toxin pneu- molysin. *Cell* **121**, 247-256.
63. Thoren, K. L., Worden, E. J., Yassif, J. M., Krantz, B. A. (2009) Lethal factor unfolding is the most force-dependent step of anthrax toxin translocation. *Proc. Natl. Acad. Sci. USA* **106**, 21555–21560.
64. Agasøster, A. V., Halskau, Ø., Fuglebakk, E., Frøystein, N. A., Muga, A., Holmsen, H., Martínez, A. (2003) The interaction of peripheral proteins and membranes studied with alpha-lactalbumin and phospholipid bilayers of various compositions. *J. Biol. Chem.* **278**, 21790-21797.
65. Surade, S., Blundell, T. L. (2012) Structural biology and drug discovery of difficult targets: the limits of ligandability. *Chem Biol.* **19**, 42-50.
66. Breton, C., Fournel-Gigleux, S., Palcic, M. M. (2012) Recent structures, evolution and mechanisms of glycosyltransferases. *Curr Opin Struct Biol.* **22**, 540-549.

FOOTNOTES

* This work was supported by the European Commission contracts no. LSHP-CT-2005-018923 and HEALTH-F3-2011-260872, the Spanish Ministry of Science and Innovation (contract SAF2010-19096), IKERBASQUE and the Basque Government (to M.E.G.); HL61228/HL66030, NIH United States (to J.M.F.). J.A.C and D.A.J. acknowledge the support from Fundación Ibercaja (Zaragoza, Spain) and the Fundación Biofísica Bizkaia, respectively. M.K. acknowledges the support from European Comission (the 7th Framework Programme) Marie Curie grant IDPbyNMR (contract No 264257).

ACKNOWLEDGMENT

We acknowledge the EMBL for synchrotron beamtime allocation at DORIS and PETRA storage rings, EMBL/DESY Hamburg and the X33 and P12 beamlines staff for technical assistance during data collection. We specially thank the BioStruct-X project to support access to structural biology facilities. We gratefully acknowledge all members of the

Structural Glycobiology Group (Spain) and the Single Protein Mechanics and Engineering Laboratory (USA) for valuable scientific discussions.

ABBREVIATIONS

GT, glycosyltransferase; GDP-Man*p*, guanosine 5'-diphosphate mannose; PI, phosphatidyl-*myo*-inositol; PimA, phosphatidyl-myo-inositol mannosyltransferase A; SAXS, small-angle X-ray scattering;

FIGURE LEGENDS

Figure 1. Schematic representation of $I27_2$-PimA-$I27_2$ polyprotein. *A*. The crystal structure of PimA in complex with GDP displays the typical GT-B-fold of GTs. The N-terminal domain contains the residues 1-168 (orange) and the helical C-terminal extremity 348-373 (yellow) whereas the C-terminal domain includes residues 169-310 (red) and the segment 311-347 (yellow). The two last C-terminal helices (yellow) participate in the hinge. *B*. Configuration of the atomic force microscope experiments. The polyprotein is tethered between the tip of a cantilever and a gold slide on a piezo positioner. PimA (red and orange) is flanked by four immunoglobulin I27 domains (blue).

Figure 2. Mechanical properties of $I27_2$-PimA-$I27_2$ as visualized by the force-extension mode of AFM. *A*. Typical force-extension profile of $I27_2$-PimA-$I27_2$ essentialy showing the length of the initial extension of PimA as ~180 nm. The GDP-bound form of PimA displayed a similar behavior (data not shown). *B*. The histogram includes experimental results from both PimA-apo and PimA-GDP conditions.

Figure 3. Extension of $I27_2$-PimA-$I27_2$ during a linear increase in force. *A-B*. Typical force-ramp traces of $I27_2$-PimA-$I27_2$ in the absence and presence of GDP, respectively. Step sizes events of PimA are annotated with red. *C-D*. Distribution of step sizes and unfolding forces of I27 and PimA unfolding events from 50 unambiguous traces. Theoretical extension following the Worm Like Chain of I27, the N-terminal and C-terminal β-sheets of PimA are plotted in blue, red dashed and red dotted lines, respectively.

Figure 4. A single unfolding pathway of $I27_2$-PimA-$I27_2$. *A*. A well-resolved trace illustrating a PimA unfolding pathway. The force applied increases from 0 pN to 250 pN in 5 seconds. The first steps describe extensions of 6, 4, 6, 6, 22, 35 and 42nm as the force increases revealing intermediates along the unfolding pathway of PimA. The 42nm step, colored in red, correlates with the unfolding of the C-terminal Rossmann fold. The four 24nm steps occurring at higher force are colored in blue and correspond to I27 unfolding. Grey arrows display the direction of the force applied to PimA and I27 in their native states. B. Topology of the proposed intermediates after unraveling the hinge (α7) and the last three alpha helices (α12, α13 and α14) of PimA. The N-terminal domain unfolds first followed by the C-terminal domain. In contrast to PimA, I27 unfolds in an all-or-none unfolding pathway.

Figure 5. Solution SAXS data statistics. *A.* Scattering curves of PimA *apo*, PimA-GDP complex and I27-PimA fusion polyprotein. *B*. *P*(r) function distributions of PimA *apo*, the PimA-GDP complex and I27-PimA fusion polyprotein. *C*. Plot of Radius of gyration (R_g) as a function of protein concentration. SAXS measurements were performed at concentrations 2.2, 3.5 and 5.5 mg/ml of PimA. R_g values at concentration 0 mg/ml of 28.5 Å and 27.5 Å were calculated for PimA *apo* and PimA-GDP complex, respectively.

Figure 6. Low resolution models of PimA in solution. *A*. Average low-resolution structure of PimA *apo* with the high-resolution crystal structure of PimA-GDP complex (PDB code: 2GEK) fitted by rigid body docking. *B*. Average low-resolution structure of PimA-GDP complex with the high-resolution crystal structure of PimA-GDP complex fitted by rigid body docking. *C*. Average low-resolution structure of I27-PimA fuison polyprotein with the high-resolution crystal structures of I27 and PimA-GDP complex fitted by rigid body docking.

Figure 7. Gallery of selected PimA *ab initio* SAXS models. *A-B*. Individual *ab initio* reconstruction runs of PimA *apo* and PimA-GDP complex showing structural variability, respectively.

Table 1. Parameters of the unliganded and GDP-bound forms of PimA and the I27-PimA fusion protein calculated from the SAXS data

	*Ms*PimA *apo*-form	*Ms*PimA GDP	I27-*Ms*PimA	
***P*(r) function calculation** [a]				
Effective *q*-range ($Å^{-1}$) [b]	0.011-0.250	0.014-0.253	0.007-0.380	
R_g (Å)[c]	28.5(5)	27.5(4)	37.9(5)	
Molecular mass (kDa) [c]	45(1)	50(1)	57(1)	
***Ab initio* modelling** [d]	DAMMIF	DAMMIF	MONSA*	
			I27-PimA	**PimA *apo***
Final *R*-factor (%)	0.15(1)	0.14(1)	1.30(1)	1.01(1)
			I27	**PimA**
Average NSD [e] between clusters	0.73(6)	0.8(1)	0.55(2)	0.56(1)
Average NSD within clusters	0.60(3)	0.63(2)	0.37(4)	0.42(4)

[a] Values in parenthesis are estimated errors approximated to the last decimal place
[b] *q*-range used for calculation of *P*(r) function
[c] R_g value extrapolated to 0 mg/ml concentration
[c] Molecular mass estimated by calculation of *I*(0) and comparison against bovine serum albumin
[d] Values in parenthesis are standard deviations calculated for 20 *ab initio* calculations
[e] NSD: Normalized Spatial Discrepancy
* MONSA *R*-factors are reported for each experimental curve employed in the calculation (i.e. SAXS data of I27-PimA and PimA *apo*). NSD values are reported for the individual phases reconstructed (i.e. I27 and PimA *apo*).

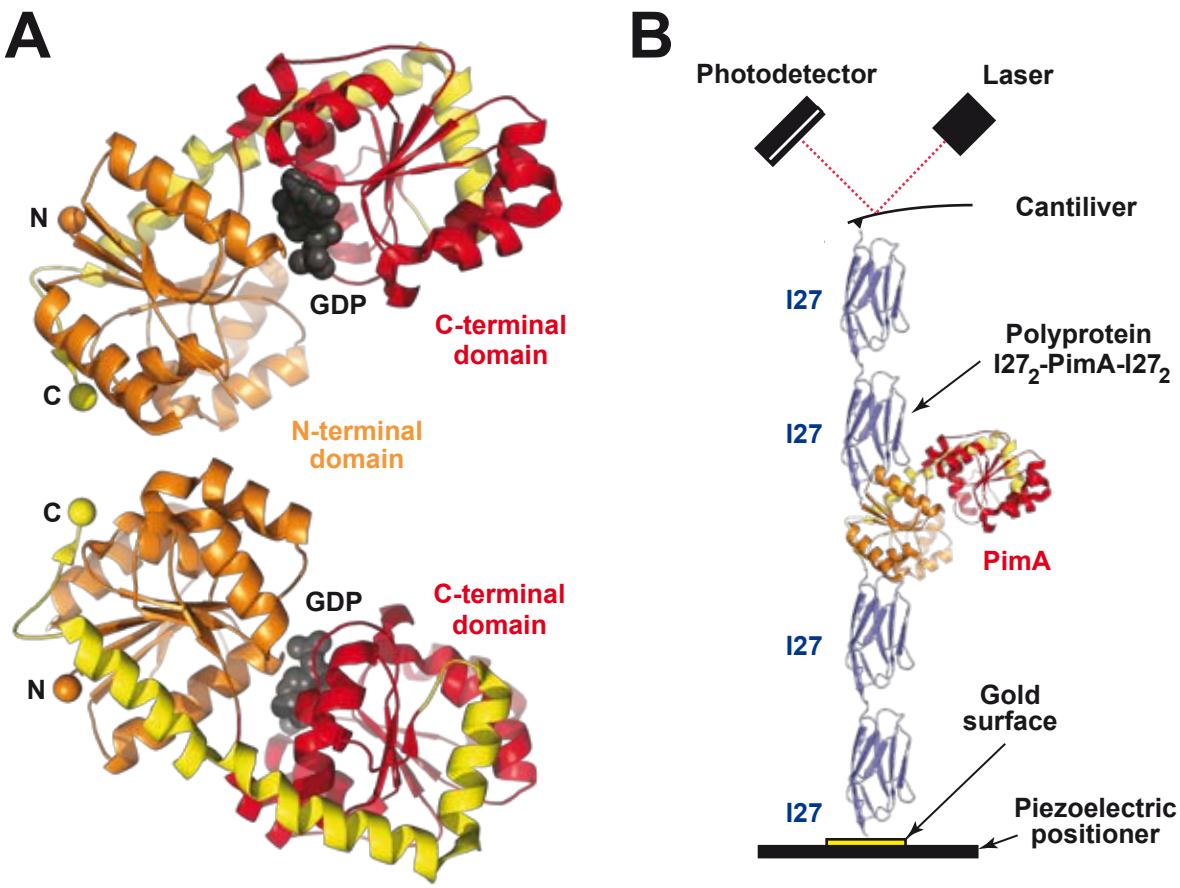


Figure 1

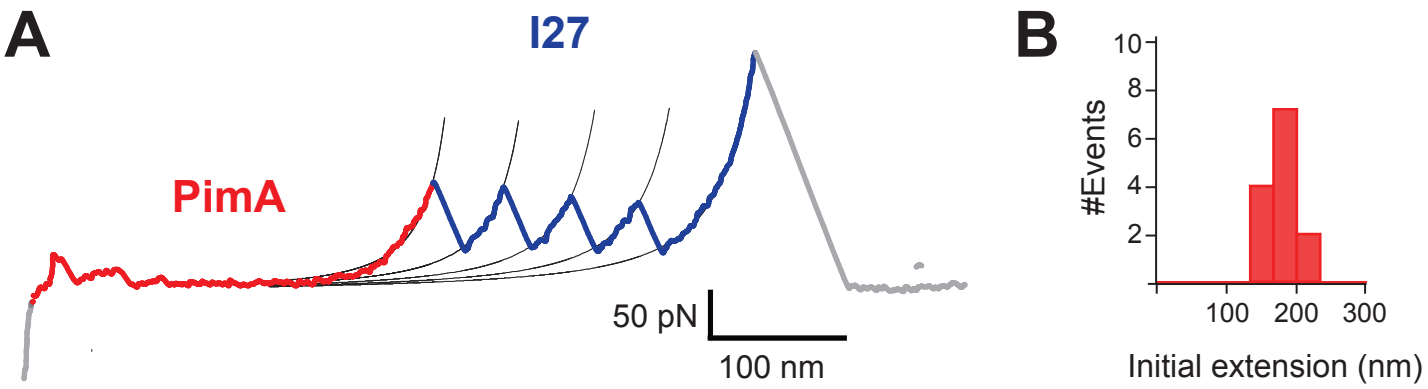


Figure 2

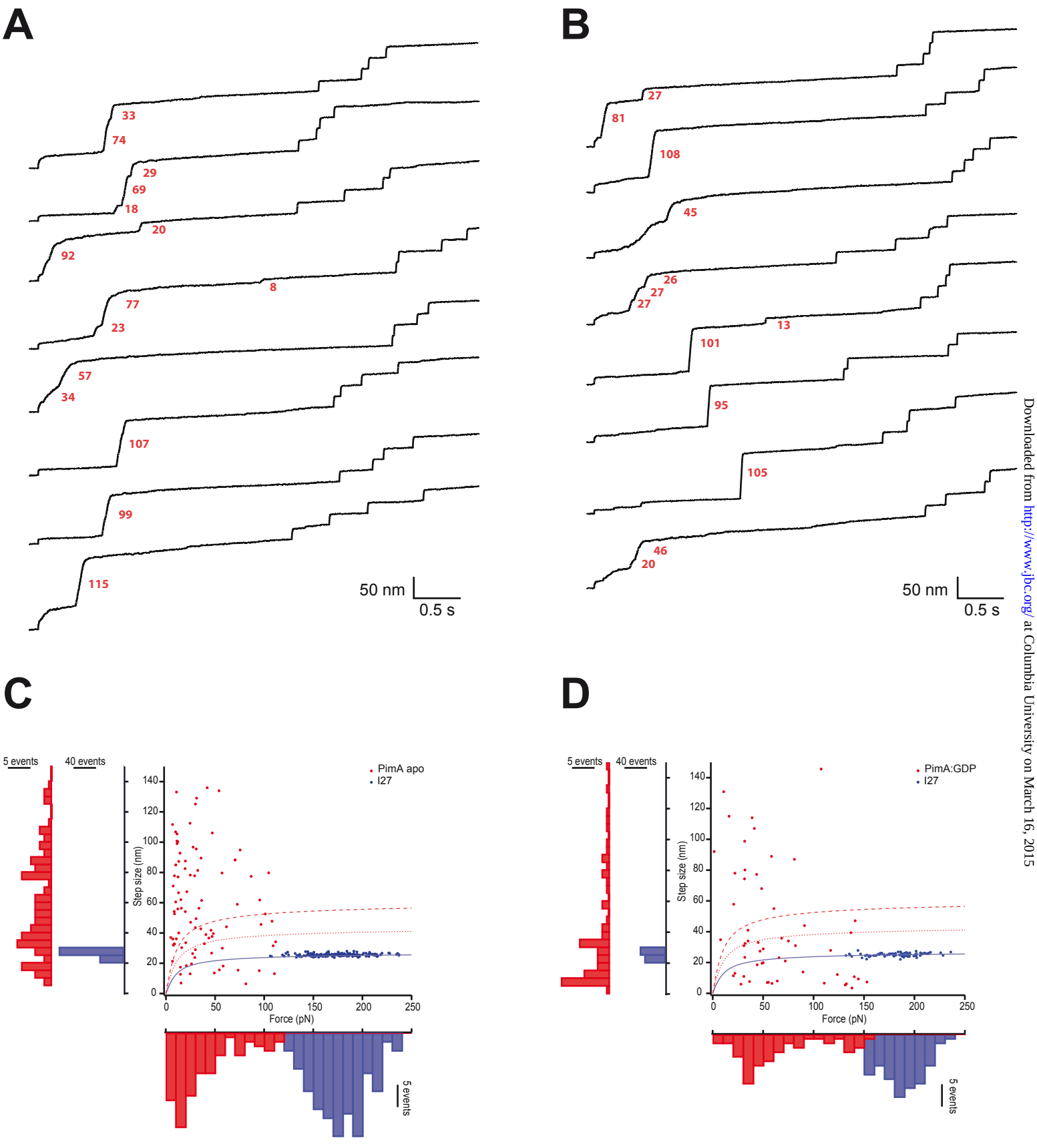


Figure 3

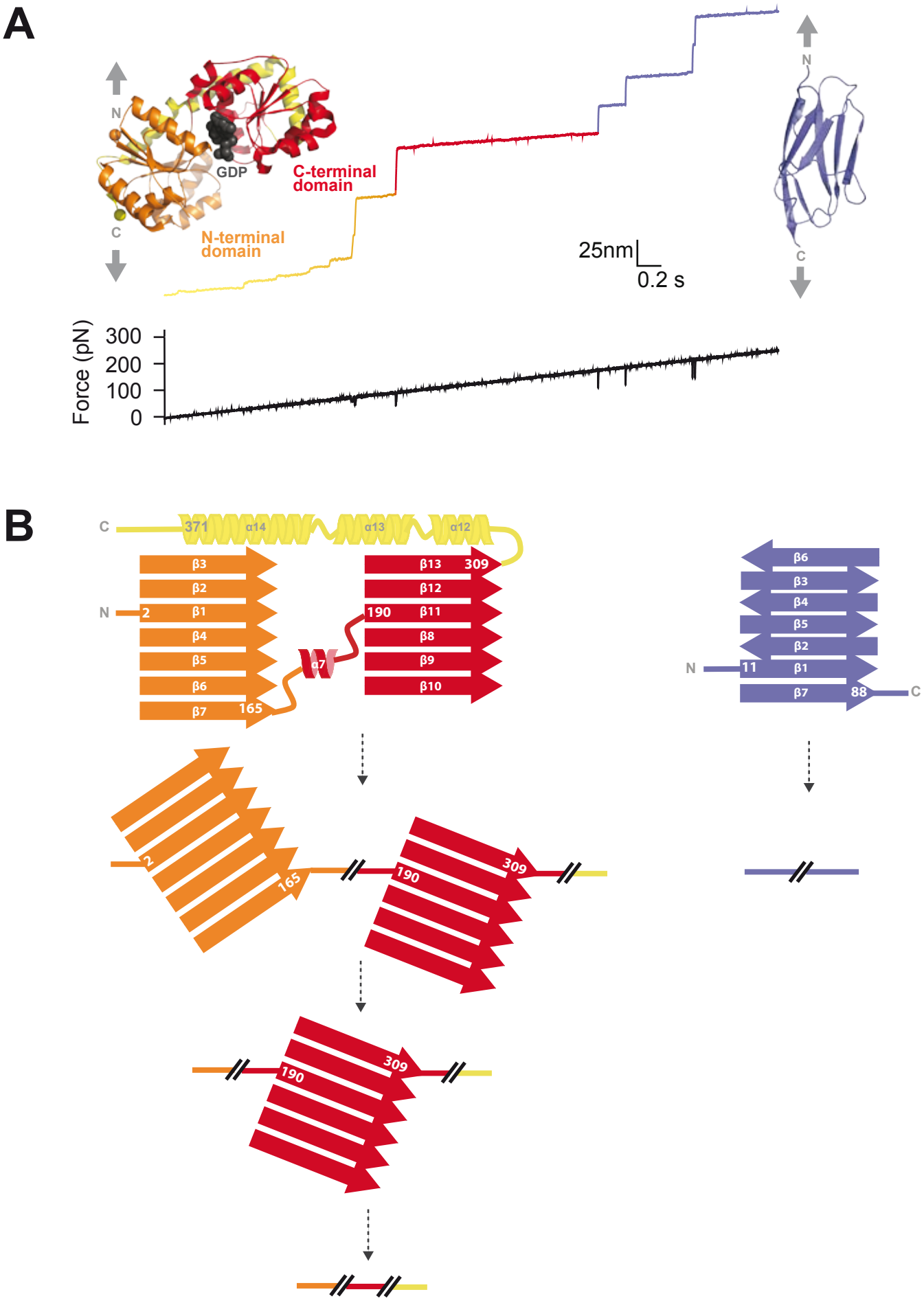


Figure 4

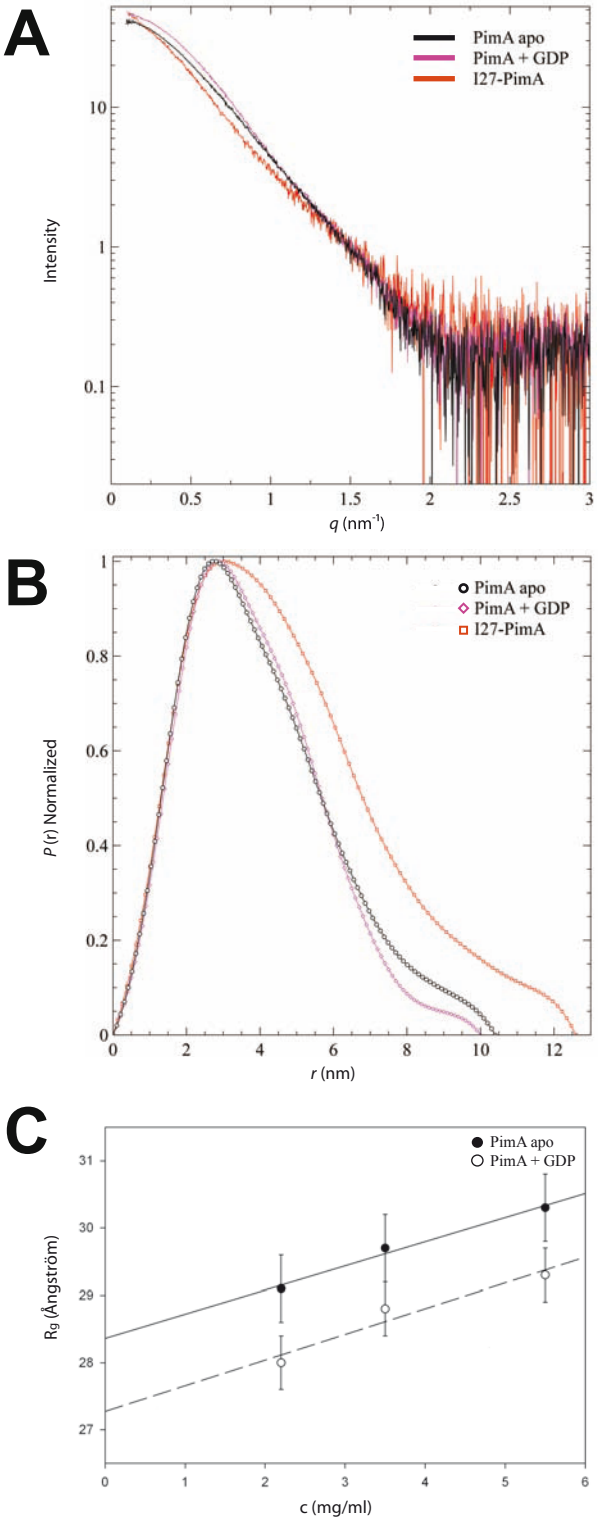

A
PimA apo
PimA + GDP
I27-PimA
Intensity
10
1
0.1
0
0.5
1
1.5
2
2.5
3
q (nm-1)
B
PimA apo
PimA + GDP
I27-PimA
P (r) Normalized
1
0.8
0.6
0.4
0.2
0
0
2
4
6
8
10
12
r (nm)
C
PimA apo
PimA + GDP
Rg (Ångström)
31
30
29
28
27
0
1
2
3
4
5
6
c (mg/ml)

Figure 5

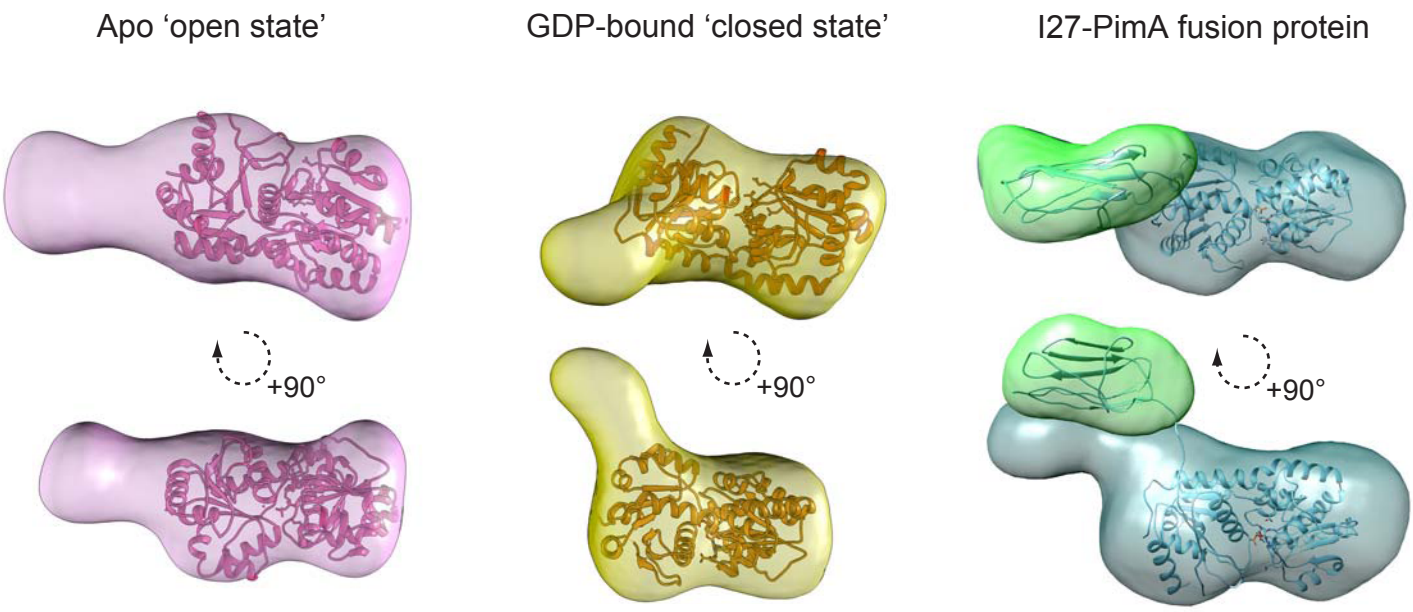

Apo ‘open state’
GDP-bound ‘closed state’
I27-PimA fusion protein
+90°
+90°
+90°

Figure 6

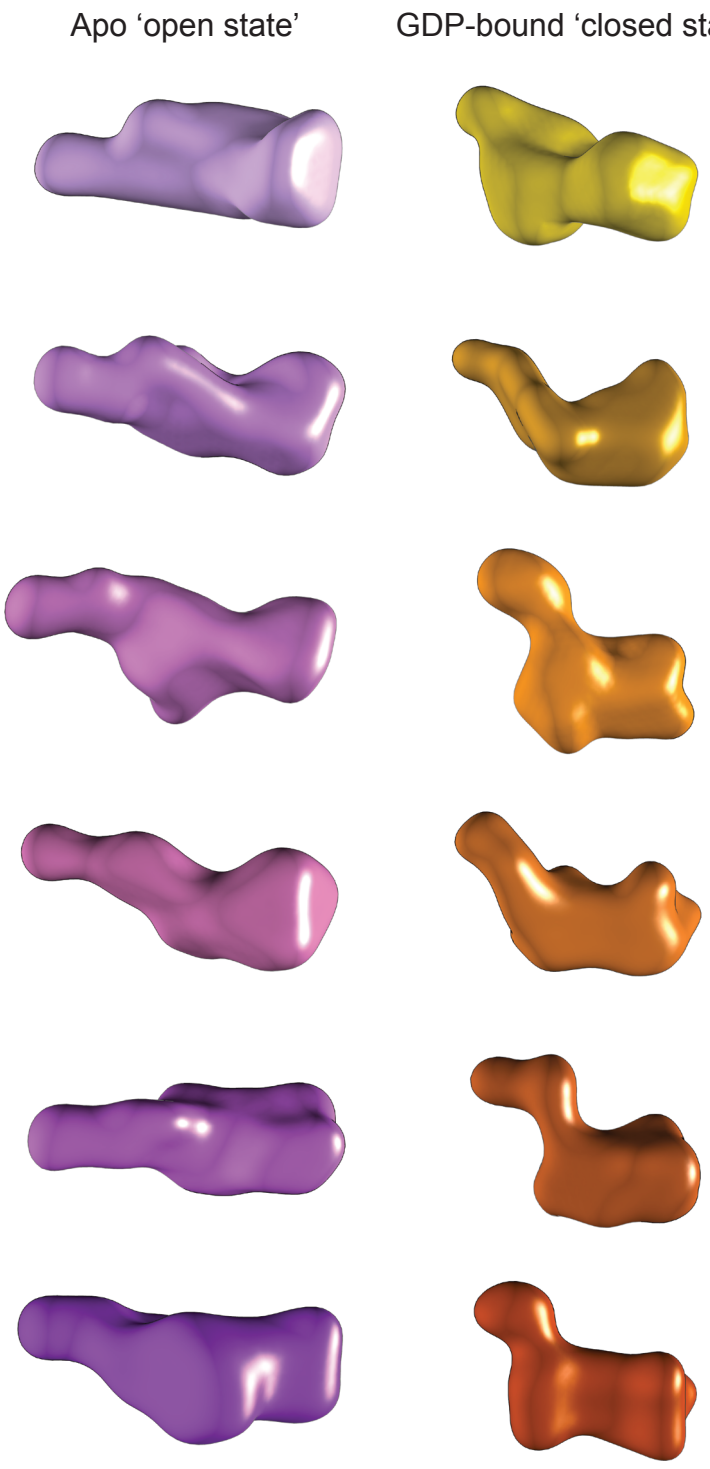
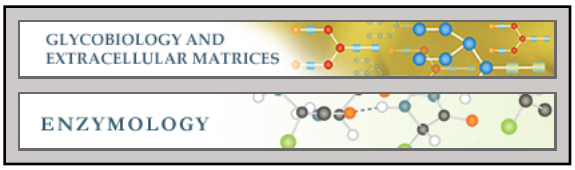


Figure 7

Glycobiology and Extracellular Matrices:

Conformational Plasticity of the Essential Membrane-Associated Mannosyltransferase PimA from Mycobacteria



David O. Giganti, Jorge Alegre-Cebollada, Saioa Urresti, David Albesa-Jove, Ane Rodrigo-Unzueta, Natalia Comino, Michael Kachala, Sonia Lopez-Fernandez, Dmitri I. Svergun, Julio M. Fernandez and Marcelo E. Guerin

J. Biol. Chem. *published online August 20, 2013*