Crystal Structure of the Yeast Phox Homology (PX) Domain Protein Grd19p Complexed to Phosphatidylinositol-3-phosphate

Cong-Zhao Zhou, Ines Li de La Sierra-Gallay, Sophie Quevillon-Cheruel, Bruno Collinet, Philippe Minard, Karine Blondeau, Gilles Henckes, Robert Aufrère, Nicolas Leulliot, Marc Graille, et al.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 50, Issue of December 12, pp. 50371–50376, 2003

Printed in U.S.A.

Crystal Structure of the Yeast Phox Homology (PX) Domain Protein Grd19p Complexed to Phosphatidylinositol-3-phosphate*

Received for publication, April 28, 2003, and in revised form, September 25, 2003
Published, JBC Papers in Press, September 26, 2003, DOI 10.1074/jbc.M304392200

Cong-Zhao Zhou‡, Ines Li de La Sierra-Gallay§, Sophie Quevillon-Cheruel‡, Bruno Collinet‡, Philippe Minard‡, Karine Blondeau¶, Gilles Henckes¶, Robert Aufrère¶, Nicolas Leulliot‡, Marc Graille§, Isabelle Sorel‡, Philippe Savarin§, Françoise de la Torre§, Anne Poupon§, Joël Janin§, and Herman van Tilbeurgh‡‖

From the ‡Institut de Biochimie et de Biophysique Moléculaire et Cellulaire (CNRS-Unité Mixte de Recherche 8619), Université Paris-Sud, Bât. 430, 91405 Orsay, France, the §Laboratoire d'Enzymologie et Biochimie Structurales (CNRS-Unité Propre de Recherche 9063), Bât. 34, 1 Av. de la Terrasse, 91198 Gif sur Yvette, France, and the ¶Institut de Génétique et Microbiologie (CNRS-Unité Mixte de Recherche 8621), Université Paris-Sud, Bât. 360, 91405 Orsay, France

Phox homology (PX) domains have been recently identified in a number of different proteins and are involved in various cellular functions such as vacuolar targeting and membrane protein trafficking. It was shown that these modules of about 130 amino acids specifically binding to phosphoinositides and that this interaction is crucial for their cellular function. The yeast genome contains 17 PX domain proteins. One of these, Grd19p, is involved in the localization of the late Golgi membrane proteins DPAP A and Kex2p. Grd19p consists of the PX domain with 30 extra residues at the N-terminal and is homologous to the functionally characterized human sorting nexin protein SNX3. We determined the 2.0 Å crystal structure of Grd19p in the free form and in complex with D-myo-phosphatidylinositol 3-phosphate (diC_4PtdIns(3)P), representing the first case of both free and ligand-bound conformations of the same PX module. The ligand occupies a well defined positively charged binding pocket at the interface between the β-sheet and α-helical parts of the molecule. The structure of the free and bound protein are globally similar but show some significant differences in a region containing a polyproline peptide and a putative membrane attachment site.

Phosphatidylinositol (PI)[1] and its phosphorylated derivatives regulate many biological processes, including cell proliferation, cell survival, differentiation, signal transduction, cytoskeleton organization, and membrane trafficking (reviewed in Ref. 1). Various chemical species can be generated by single, double, or triple phosphorylations at the inositol hydroxy groups at positions 3, 4, and 5. Their synthesis and cellular concentrations are regulated by specific lipid kinases and phosphatases. One of the major mechanisms by which PIs regulate cellular processes is by their capacity to serve as membrane signals to affect intracellular localizations of effector proteins. A number of PI-binding motifs have now been characterized such as the FYVE domain (2), the pleckstrin homology domain (3), the epsin N-terminal homology domain (4), and the Phox homology (PX) domain (5). The PX domain was first identified as a motif of about 130 amino acids present in the $p40^{phox}$ and $p47^{phox}$ subunits of the neutrophil NADPH oxidase superoxide-generating complex. The PX domain was afterward recognized in a vast number of other proteins covering a wide palette of molecular functions: cell signaling pathways (phospholipases D, phosphatidylinositol 3-kinase), vesicular trafficking and yeast morphology (human sorting nexins; yeast Vps5p, Vps17p, Vam7p, and Mvp1p) and control of yeast budding and cell polarity (Bem1p and Bem3p).

The PX domain can be present alone or in combination with other functional domains (6). Three broad subclasses of PX containing proteins can be distinguished. The first subclass contains small proteins that essentially consist of the PX domain (sorting nexin 3 (SNX3), SNX10, SNX12, SNX23, SNX24, SNX26, and yeast Grd19p). PX domains in the second subclass contain extra sequences, such as coiled-coil regions that are involved in complex formation with other partners (SNX1, SNX2, Pps5, SNX4, SNX5, and SNX6). The third subclass encompasses PX proteins that contain additional well characterized domains: SH3 motifs are present in SNX9, SNX18, SNX28, $p40^{phox}$, $p47^{phox}$, FISH, Bem1p and Scd2p, but many more domain combinations exist (for a recent review see Ref. 7).

Sorting nexins are found in a variety of organisms as regulators of cellular protein trafficking that have the presence of the PX domain in common (7–9). The dynamic vessel transport processes at the late Golgi compartment of yeast (trans-Golgi network) require dedicated mechanisms for correct localization of resident membrane proteins. Recently a new gene, *GRD19*, was identified to be involved in the localization of late Golgi membrane proteins. It was concluded that Grd19p is a component of the retrieval machinery that functions by direct interaction with the cytosolic tails of certain TGN membrane proteins during the sorting/budding process at the prevacuolar compartment (10). The mammalian orthologue of Grd19p, SNX3, was shown to associate with early endosomes through a PX domain-mediated interaction with phosphatidylinosytol-3-phosphate (11). Overexpression of this protein alters/delays transport to the lysosome.

The phosphoinositode binding of all yeast PX domains was recently experimentally analyzed, and they all bind PtdIns(3)P with high affinity. The glutathione *S*-transferase SNX3/Grd19p fusion protein in particular was shown to bind with an apparent K_d between 0.15 and 0.5 μM (12).

* This work is supported by grants from the Ministère de la Recherche et de la Technologie (Programme Génopoles) and the Association pour la Recherche contre le Cancer (to M. G.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ To whom correspondence should be addressed. Tel.: 33-1-69-82-34-91; Fax: 33-1-69-82-31-29; E-mail: herman@lebs.cnrs-gif.fr.

[1] The abbreviations used are: PI, phosphatidylinositol; PtdIns(3)P, phosphatidylinositol 3-phosphate; DiC_4PtdIns(3)P, D-myo-phosphatidylinositol 3-phosphate; PX, Phox homology; SH, Src homology; SNX, sorting nexin; IPTG, isopropyl-1-thio-β-D-galactopyranoside; RMS, root mean square.

TABLE I
Data collection and refinement statistics

	Complex	Native	Edge	Peak	Remote
Wavelength (Å)	0.984	0.9797	0.9803	0.98	0.9825
f′, f′′			−21.09, 7.93	−12.88, 17.04	−9.74, 0.65
Unit-cell parameters/Spcgrp	P1	P 6_1 2 2			
a, b, c (Å)	31.43, 55.76, 64.75	55.73, 55.73, 187.51			
α, β, γ (°)	110.75, 97.35, 99.49	90, 90, 120			
Resolution (Å)	25.08–2.28	25–2.03	25–2.33	25–2.33	25–2.33
Total number of reflections	65105	60571	67536	67464	40121
Number of unique reflections	11970	17289	7654	7671	7503
Multiplicity	3.8	5.1	8.8	8.8	5.3
R_{merge}[a,b]	0.047 (0.156)	0.065 (0.296)	0.057 (0.254)	0.057 (0.2)	0.032 (0.155)
$I/\sigma(I)$[a]	25.1 (3.2)	13.6 (2.5)	21.2 (3.4)	26.6 (5.0)	27.0 (5.8)
Overall completeness (%)[a]	95.97 (77.31)	99.7 (98.1)	96.9 (96.1)	97.1 (96.3)	95.0 (87.4)
Refinement					
Resolution (Å)	20–2.3	20–2.03			
R_{cryst}/R_{free}[c]	0.222/0.256	0.217/0.247			
Non-hydrogen atoms	2231	1089			
Water molecules	78	63			
Mean B protein/ligand/wat (Å^2)	41/46/40	25.3/-/29			
Root mean square deviation					
Bonds (Å)	0.007	0.006			
Angles (°)	1.3	1.2			
Ramachandran analysis					
Most-favored	90.4	93.2			
Allowed	9.6	6.8			

[a] Numbers in parentheses are for highest resolution sell.
[b] $R_{merge} = \Sigma_h\Sigma_i|I_{hi} - \langle I_h\rangle|/\Sigma_h\Sigma_i I_{hi}$, where I_{hi} is the ith observation of the reflection h, while $\langle I_h\rangle$ is the mean intensity of reflection h.
[c] $|R_{factor} = \Sigma\ ||F_o| - |F_c||/|F_o|$. R_{free} was calculated with a small fraction (4.8%) of randomly selected reflections.

Structures of a few PX domains have now been determined by crystallography and NMR in solution. The overall structure of the PX domain, as observed in p40^{phox}, p47^{phox} and Vam7p, is a three-stranded β-sheet followed by three helices (13–15). Structures in complex with soluble PtdIns(3)P compounds and in the presence of sulfate ions have allowed to structurally define the lipid-binding pocket (13, 16). It was shown by NMR that the Vam7p PX domain binds specifically to PtdIns(3)P through conserved positive residues and that this binding induces conformational changes in the protein (17). We present here the structure of Grd19p, the first of an intact PX domain protein, free and in complex with a soluble diC$_4$PtdIns(3)P compound. Our results allow us to answer the pending question on how the structure of a PX domain adapts to PtdIns(3)P binding. The overall structure is analogous to those determined for other PX domains. The ligand induces some conformational changes in the loops surrounding the lipid-binding site.

EXPERIMENTAL PROCEDURES

Cloning and Purification—The *YOR357c* open reading frame DNA was amplified by PCR using oligodeoxynucleotides synthesized by MWG-Biotech and cloned between the NdeI and NotI sites of pET29 vector from Novagen. The expression strain BL21(DE3) pLysS was obtained from Novagen. 2xYT medium was purchased from BIO 101, Inc. and isopropyl β-D-thiogalactopyranoside (IPTG) from Sigma. BL21 (DE3) pLysS cells transformed with the construct were grown at 37 °C up to an $A_{600\ nm}$ of 1. Expression was induced with 0.3 mM IPTG, and the cells were grown for a further 4 h at the same temperature. Cells were collected by centrifugation, washed, and resuspended in lysis buffer (20 mM Tris-HCl, pH 8, 200 mM NaCl, 5 mM β-mercaptoethanol) at 4 °C. Cells were lysed by 3 cycles of freeze/thawing and 30 seconds of sonication and were then centrifuged. The His-tagged proteins were purified using a Ni^{2+} affinity column (nickel-nitrilotriacetic acid, Qiagen Inc.) and standard protocols. Eluted protein was further purified by gel filtration using a Superdex™ 75 (Amersham Biosciences) equilibrated against a buffer containing 20 mM Tris-HCl, pH 6.8, 200 mM NaCl, 10 mM β-mercaptoethanol. The purity of the pooled fractions was checked by SDS-PAGE and the integrity of the protein samples by mass spectrometry.

Se-Met-labeled proteins were prepared as follows. Transformed cells were grown in M63 medium at 37 °C up to an $A_{600\ nm}$ of 1. The culture was then complemented with a mixture of amino acids (L-Lys, L-Phe, and L-Thr at 100 mg/l; L-Ile, L-Leu, L-Val and L-Se-Met at 50 mg/l) during 30 min, and the recombinant protein was overexpressed by addition of 0.3 mM IPTG during 4 h. The protein was purified as described before. Seleno-methionine incorporation into the protein was >95% as assayed by mass spectrometry.

Crystallization and Data Collection—The protein (concentrated at 3 mg/ml) was crystallized at 293 K by the hanging-drop vapor diffusion method using the sparce matrix crystallization screening kits from Hampton Research (18). Crystals of the native protein were grown from 1:1 drops of protein and 16% polyethylene glycol 4000, 0.1 M Na-HEPES, pH 7.5 (reservoir solution). Rod-shaped crystals of a maximum size of 200 micron appeared within 1 week. Hexagonal-shaped crystals from the Se-Met-labeled protein were grown under similar conditions. Crystals of the complex were obtained by mixing 1 μl of complex solution (protein at 7 mg/ml with 2 mM diC$_4$PtdIns(3)P in 0.1 M NaCl, 10 mM β-mercaptoethanol, 20 mM Tris-HCl, pH 7) with 1 μl of 10% polyethylene glycol 8000, 8% ethylene glycol, 0.1 M Hepes, pH7.5 (reservoir solution).

X-ray data were collected from cryo-cooled crystals using mother liquor containing 30% glycerol as a cryoprotectant for the apo-protein and 30% ethylene glycol for the complex. Native and multi-wavelength anomalous diffraction data on Se-Met-substituted crystals were collected at beam line BM30A of the European Synchrotron Radiation Facility. Crystals of the apo-protein belong to space group P$6_1$22 and have one molecule in the asymmetric unit (solvent content is 44%, see Ref. 19). Crystals of the complex belong to space group P1 with two molecules per asymmetric unit (solvent content = 54%). Diffraction data were processed using the programs DENZO and SCALEPACK (20). Relevant cell parameters and data collection statistics are reported in Table I.

Structure Determination—Resolution of the crystal structure of Grd19p was performed by the multiple anomalous diffraction method on Se-Met-labeled protein crystals (Table I). The position of four Se atoms could be located and the structure was phased with the program SOLVE, using the data to 3.5 Å resolution (21). Solvent flattening using RESOLVE brought the overall figure-of-merit from 0.55 to 0.74. The phases were subsequently extended to 2.3 Å with RESOLVE. The resulting map allowed us to place 63% of the residues automatically (22). These starting phases were used for the automatic procedure of model building and refinement implemented in warpNtrace, against the native 2 Å resolution data (23). The ARP model was completed and manually adjusted with the program turbo-FRODO (afmb.cnrs-mrs.fr/TURBO_FRODO/turbo.html). Refinement was performed using the native data set up to 2.03 Å resolution following standards protocols as implemented in the program REFMAC (24). All refinement statistics can be found in Table I.

The current model includes residues 29–162 and 69 water molecules. No electron density was detected for the 28 N-terminal residues. Mass

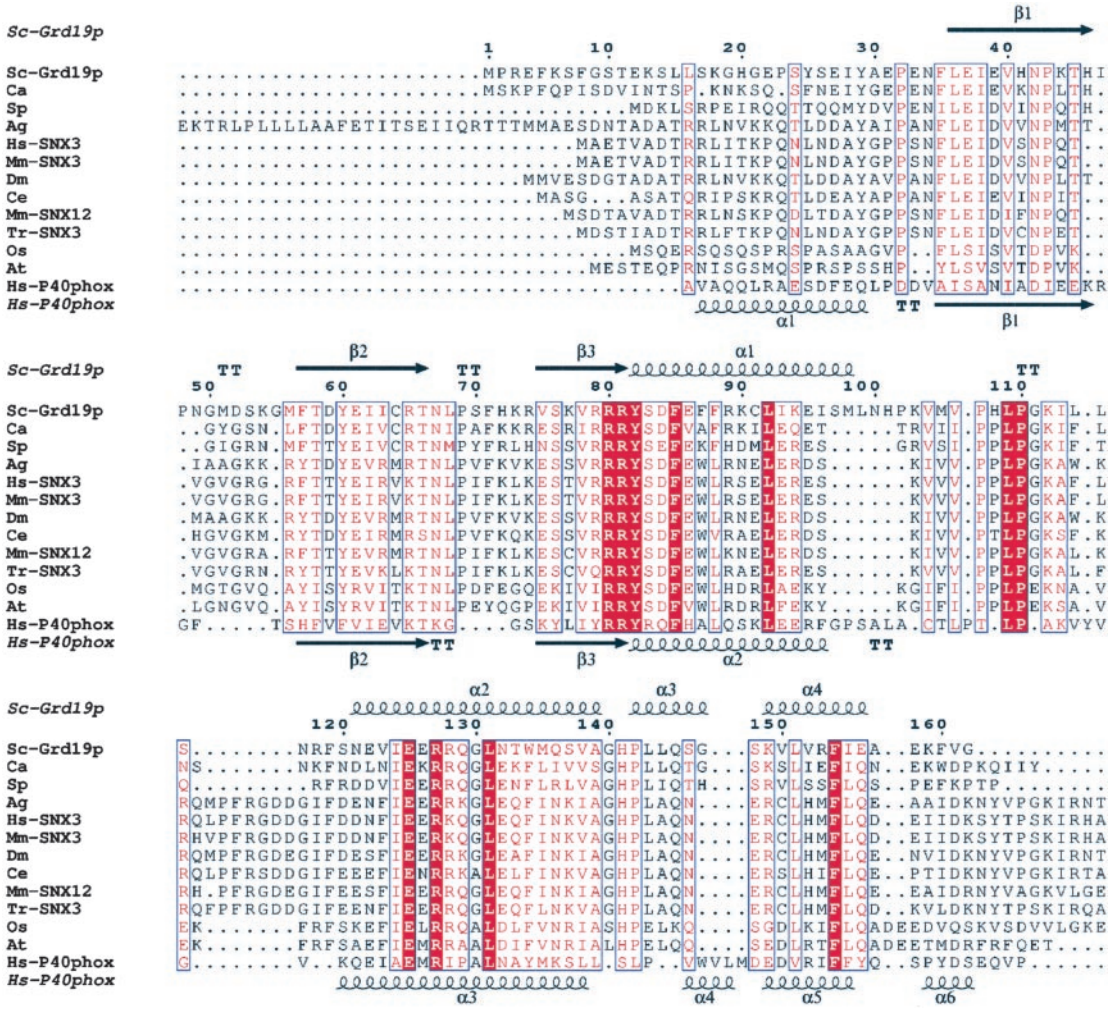

FIG. 1. **Multiple sequence alignment of the "SNX3 branch" of the PX domain superfamily.** The alignment was obtained using CLUSTALW (32). Residue numbering refers to the sequence of Grd19p (*YOR357C*) from *Saccharomyces cerevisiae*. Secondary structure elements were obtained from the structure of Grd19p without ligand using the ESPript server (33). A $p40^{phox}$ structure-based sequence alignment together with secondary structure assignment, based on its crystal structure, is shown at the *bottom*. *Ca*, *Candida albicans* (emb_CAA21987.1); *Sp*, *Schizosaccharomyces pombe* (NP_596480.1); *Ag*, *Anopheles gambiae* (EAA07353.1); *Hs*, SNX3 *Homo sapiens* (CAB75627.1); *Mm-SNX3*, *Mus musculus* (NP_059500.1); *Dm*, *Drosophila melanogaster* (NP_650214.1); *Ce*, *Caenorhabditis elegans* (NP_492437.1); *Mm-SNX12*, *Mus musculus* (NP_061363.1); *Tr-SNX3*, *Takifugu rubripes* (AAM61764.1); *Os*, *Oryza sativa* (BAB91776.1); *At*, *Arabidopsis thaliana* (NP_196232.1).

spectrometric analysis of a dissolved crystal indicated that it contained partially proteolyzed protein (the majority of the crystal sample missed the 28 N-terminal residues; data not shown). The final model has all residues in allowed regions of the Ramachandran plot and 93.2% in the most favorable regions according to the program PROCHECK (25).

The structure of the Grd19p in complex with diC_4PtdIns(3)P was solved by molecular replacement with the apo-protein structure as search model and using the CNS package (R_{factor} 39.2%, and R_{free}-factor 37.7%) (26). The cross-validated, sigma-A weighted map calculated from this model showed very clearly the presence of the diC_4PtdIns(3)P ligand in the binding site. The coordinates and structure factors have been deposited in the Protein Data Bank (accession number 1OCS for the native and 1OCU for the complex).

RESULTS AND DISCUSSION

Overall Structure of Grd19p—The structure of Grd19p, the first for an intact PX domain protein, was solved by multiple anomalous diffraction using Se-Met-substituted crystals at a resolution of 2.0 Å. Analysis of the amino acid sequence of Grd19p reveals a well identified PX domain (residues 30–162, Fig. 1), but the twenty-nine N-terminal residues have no sequence similarity with other motifs or with any other PX protein. In both Grd19p crystal structures (with and without ligand), clear electron density was only observed for the PX domain from residue Tyr-29. Mass spectrometry and SDS-PAGE analysis of a dissolved crystal showed that protein sample present in the crystal was partially proteolysed at this position. The most likely hypothesis is that the N-terminal region of Grd19p is unstructured and hence susceptible to proteolysis. It is not known whether this region is important for function. The Grd19p PX domain has the overall structure as observed for other PX domains: a fairly flat molecule consisting of a three-stranded anti-parallel meander β-sheet, β1 (residues 36–47), β2 (56–67), and β3 (74–81) at the N-terminal part and a four helical bundle at the C-terminal part (α1 (83–99), α2 (122–139), α3 (142–147), and α4 (149–157)). Fig. 2*A* shows the superposition of the structure of the Grd19p-PX domain to those of the PX domains from $p40^{phox}$, $p47^{phox}$, and Vam7p (13–16). The sequence similarity between these PX domains is rather low (only 11 strictly conserved residues between Grd19p and $p40^{phox}$, Fig. 1) explaining the rather high mutual RMS deviations between the superposed structures: $p47^{phox}$ (RMS: 1.5 Å for 86 Cα positions), $p40^{phox}$ (RMS: 2.2 Å for 117 Cα positions), and Vam7p (RMS: 2.1 Å for 86 Cα positions). The structured core is very well conserved, and the most pronounced differences are observed in the loop connections that are of unequal length between the various PX domains. Especially the poly-proline loop (containing residues 100–116 in Grd19p) whose C-terminal end comes close to the PtdIns(3)P-

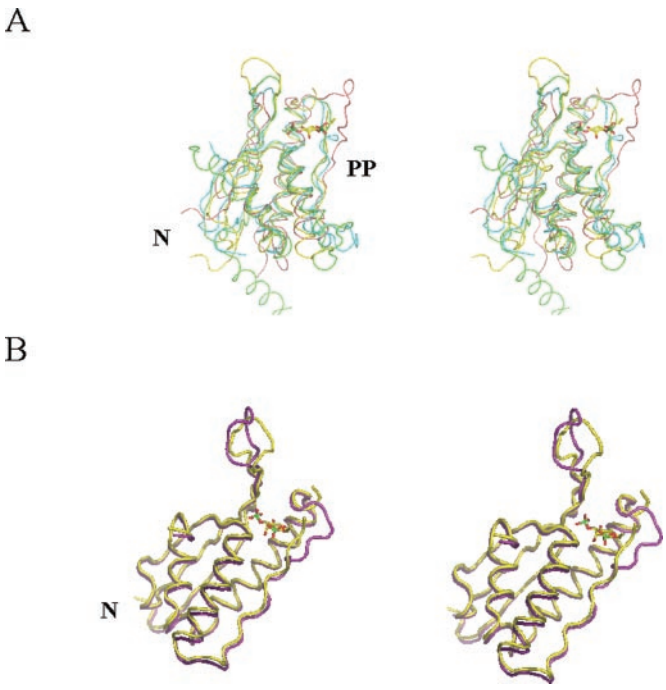

FIG. 2. **Overall comparison between PX domain structures.** *A*, ribbon presentation of the superposition between Grd19p-PX (*yellow*) and the other available PX domain structures. $p47^{phox}$ PX is in *blue* (16), Vam7p is in *pink* (15), and p40phox PX (13) is in *green*. $diC_2PtdIns(3)P$ bound to Grd19p is represented in *stick*. The proline-rich region is also indicated (*PP*). *B*, global superposition between the free (*pink*) and ligand-bound (*yellow*) structures of Grd19p. The N-terminal is marked. Structures were superposed with the molecular graphics program TURBO, all figures representing structures are prepared using PyMol (www.pymol.org).

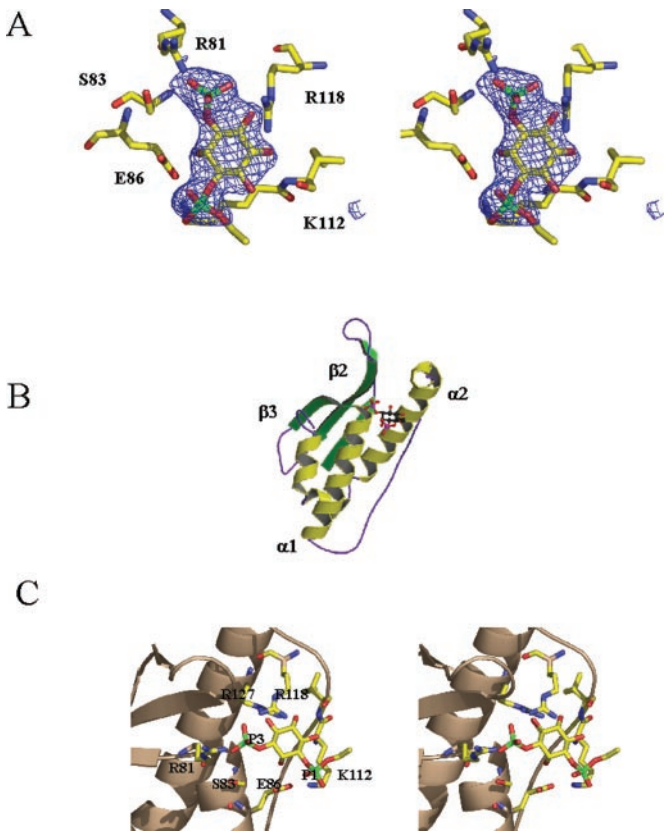

FIG. 3. *A*, stereo view of an omitted $F_o - F_c$ electron density map covering the bound diC4PtdIns(3)P, contoured at 3σ. *B*, overall view of binding of diC4PtdIns(3)P to Grd19p. Helices are colored *yellow*, strands *green*, and loop connections *purple*. The ligand is represented in *sticks* (*gray*, carbon; *red*, oxygen; and *pink*, phosphate). Some of the secondary structure elements (Fig. 1) are labeled. *C*, detailed stereo view of diC4PtdIns(3)P binding to Grd19p (*yellow*, carbon atoms; *green*, phosphate; *red*, oxygen; and *blue*, nitrogen). Positions of the phosphate groups of the ligand on the 1- and 3-OH position are labeled as *P1* and *P3*, respectively.

binding site adopts very different conformations in the various PX domains (Fig. 2*A*). This could be related to different functions of these loops in the various proteins. Because it is rich in exposed hydrophobic residues, the C-terminal part of this loop has been coined the membrane attachment site. In the case of $p47^{phox}$ this loop is documented to interact with a SH3 domain within the same protein. PX domains of the SNXs represent a distinct subgroup of the PX superfamily and may have evolved independently from other PX domains (27). Our structure of the Grd19p-PX shows that the various branches of the PX superfamily probably all share a common scaffold, despite the fact that they cannot be retrieved by mutual BLAST searches.

Structure of Grd19p in Complex with $diC_4PtdIns(3)P$ and Comparison with the Structure of $p40^{phox}$ PX Bound to PtdIns(3)P—Plasma resonance experiments showed that PtdIns(3)P binds to Grd19p with a high apparent affinity constant K_d between 0.15 and 0.5 μM (12). We therefore used the soluble $diC_4PtdIns(3)P$ compound in our co-crystallization trials. Because the lipid-binding site in the crystal of the Grdp19p-PX apo form is blocked through interaction with a symmetry related molecule, crystals of the complex were obtained by co-crystallization of the protein with $diC_4PtdIns(3)P$. The complex crystallized with two molecules in the asymmetric unit and in a different space group from the apo-protein (Table I). Gel filtration of the purified protein showed that it elutes as a monomer in absence of ligand (the apo-protein structure has one molecule bound in the asymmetric unit). We suppose that dimerization in the presence of the ligand is probably induced by the crystallization conditions. The structure of the complex could be solved in a straightforward way by molecular replacement using the structure of the apo-protein at a resolution of 4 Å and the structure was refined against 2.3 Å data (refinement statistics are in Table I). The structures of both molecules in the asymmetric unit are identical (RMS of 0.47Å for all Cα positions), but with one noticeable difference in the lipid-binding site. During refinement of the protein the residual electron density revealed very clearly the presence of the bound PtdIns(3)P head group in the two Grd19p copies of the asymmetric unit (Fig. 3*A*). The phosphate head group has well defined electron density in both molecules of the asymmetric unit, but only some guesses can be made on the positions of the hydrophobic tails. The interactions of the ligand with Grd19p also are very similar in both copies of the molecule. The ligand refined with B factors comparable with those from the surrounding protein residues.

The structures of Grd19p with and without ligand represent the first case where both forms of the same PX domain could be structurally characterized and offers the possibility to evaluate conformational changes induced by ligand binding. The overall structures of the ligand-bound and free Grd19p are very similar (Fig. 2*B*). Superposition of both structures yields a RMS of 0.58 Å for 120 Cα positions. The core region that provides the majority of specific interactions to the $diC_4PtdIns(3)P$ is identical in both structures. Two regions, however, adopt different conformations in apo and complexed Grd19p: the loop connecting β1 to β2, comprising amino acids 45–53, and the region connecting α1 and α2 (residues 111–117). The first region, situated close to the $diC_4PtdIns(3)P$-binding site, is involved in dimer formation. The second region, which adopts a different conformation, provides one rim of the ligand-binding site. This region is situated at the exit of the hydrophobic moiety of the

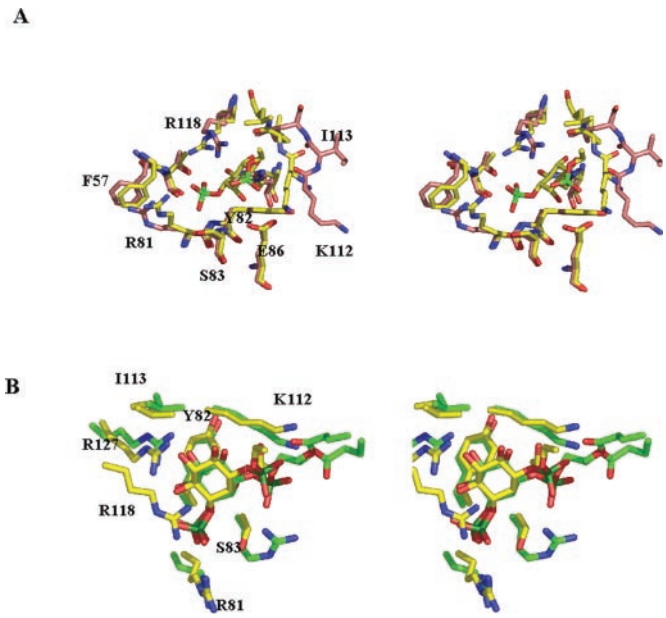

FIG. 4. *A*, stereo view of the superposition of the ligand-bound (C atoms in *yellow*) and free form (C atoms in *wheat*) of Grd19p. Residues in a radius of 5 Å around the bound ligand are represented. *B*, stereo view of the superposition of the ligand-binding sites of Grd19p and $p40^{phox}$. Grd19p, same color code as Fig. 4*A*. $p40^{phox}$ has carbon atoms in *green*. Only the Grd19p residues are labeled.

ligand. In the complex, the peptide is moving in toward the bound diC_4PtdIns(3)P, and some residues are changing conformation (illustrated in Fig. 4*A*). Probably the key factor in this change is the 3.5 Å movement of Lys-112, approaching its Nz position toward a non-bridging oxygen of the 1-phosphate group of the bound ligand (Figs. 3*B* and 4*A*).

The electron density for residues 114 and 115 is unclear in both copies of the asymmetric unit, suggesting mobility. These residues have well defined electron density in the structure of the apo form, where they form part of a region that is involved in crystal packing. In the complex the loop has no neighboring contacts. The bottom of the binding pocket and the protein regions in contact with the specificity-determining 3-phosphate group are not undergoing any conformational changes upon ligand binding.

The phosphoinositide head group binds in a pocket contained between the N-terminal regions of helix α1, strand β2, the poly-proline loop, and the N-terminal part of helix α2 (Fig. 3*B*). Prominent interactions are formed between the 3-phosphate, 1-phosphate, and 4/5-hydroxyl groups of the ligand and well conserved residues in the binding pocket (Fig. 3*C*). The bottom of the pocket is occupied by a few well ordered water molecules. The well conserved Arg-81 provides the main specific interaction with the 3-phosphate group of the bound diC_4PtdIns(3)P: its NH_2 group is at hydrogen bond distance from the 3-phosphate oxygen. In other PX domains, this residue was shown to be critical in ligand binding and endosomal localization *in vivo*, and a basic residue is found in all PX domains that bind 3-phosphoinositides (11, 13). The 3-phosphate group is further stabilized by a hydrogen bond with the Ser-83 side chain OH group. Serine is present at this position in 70% of the yeast PX domain sequences, or is replaced by residues with analogous H-bonding possibilities. It is totally conserved in the SNX3 family members (Fig. 1). The conformations of the side chains interacting with the ligand are very similar for the two molecules of the asymmetric unit, except for Arg-118. In the B molecule Arg-118 forms a second salt bridge with the 3-phosphate group, but in the A molecule its side chain is pointing away from it (not shown).

Superposition of the PtdIns(3)P-bound complex of the $p40^{phox}$ PX domain with the Grd19p complex demonstrates that the ligands bind in very similar orientations and that their binding is stabilized by similar interactions (Fig. 4*B*) (13). Superposition of the x-ray structure of the $p47^{phox}$ domain with bound sulfate ions and the Grd19p complex, shows that the sulfate ion occupies exactly the same position as the 3-phosphate group of the diC_4PtdIns(3)P ligand in Grd19p. Taken together these results suggest that PtdIns(3)P-specific PX domains are interacting in very similar ways with their phosphatidylinositol-ligands.

Filter binding assays showed that Grd19p is very specific for the phosphate at the 3-OH position: no significant binding could be observed for any other combination of phosphate positions on the PtdIns head (4 or 5 position alone or in combination with the 3 position) (12). Modeling suggests that a phosphate on the 5-OH position would create a steric clash with the main chain of residues 112 and 113 and with the side chain of Arg-127. The weak binding of compounds with phosphates on the 4 position is more difficult to rationalize. A torsional reorientation of the side chain of Arg-118 would be sufficient to accommodate the phosphate group at the 4 position, a conformation that is in fact observed in one of the two Grd19p molecules in the asymmetric unit. A 4-phosphate group could even engage in a favorable interaction with Arg-127. Bravo *et al.* suggest that the environment around the 4-OH group is too crowded to accommodate an extra phosphate (13). PX domains that show affinity for PtdIns(4)P compounds however all have equivalent residues around the 4-OH position. There probably exist subtle adjustments to optimize binding of appropriate PX domains to these compounds. The 1-phosphate group is very close to the Nz atom of the Lys-112 side chain, which is further stabilized by a salt bridge with Glu-86. Lys-112 is conserved in all SNX3 family PX domains (Fig. 1).

Not only the 3-phosphate group but also the inositol hydroxyl groups are engaged in specific interactions with the protein. The 4- and 5-OH groups form hydrogen bonds with the NH_1 and NH_2 of Arg-127 (Fig. 3*C*). Mutation of the equivalent Arg-105 in the $p40^{phox}$ PX and Arg-90 in the $p47^{phox}$ PX domains abolishes binding (28), and mutation of an equivalent residue eliminates endosomal localization in cytokine-independent survival kinase (29). This arginine is conserved in almost all PX sequences. The inositol 2-OH group forms a charged hydrogen bond with the carboxylate group of Glu-86. A carboxylate group (Glu or Asp) is present in all the yeast PX sequences at this position suggesting that this interaction could be important for inositol recognition. The 2-OH group of the PtdIns(3)P bound to the $p40^{phox}$ PX domain does not form any hydrogen bonds. Grd19p-Glu-86 is replaced by His-63 in $p40^{phox}$ PX and this residue has its side chain turned away from the ligand.

One face of the inositol ring rests on the hydrophobic part of the Tyr-82, a very well conserved residue in all yeast PX domains (Figs. 3*C* and 4*A*). This type of hydrophobic stacking is a frequently observed theme in binding sites of carbohydrates on proteins. The hydrophobic tail of the ligand is close to the mobile region of the proline-rich region loop, rich in hydrophobic residues (Ile-113, Leu-114, and Leu-115). The position and hydrophobic character of this loop suggests that this region could participate in interaction with the membrane lipid bilayer. Large chemical shifts were observed in this region upon binding of the Vam7p PX domain to micelles (17). Considerable sequence divergence and large insertions among the various PX sequences occur at this region.

One non-bridging oxygen of the 1-phosphate group of the ligand is close to the side chain of Lys-112. The corresponding non-bridging oxygens of the ligand in the $p40^{phox}$ complex are

stabilized by charge interactions with the basic residues Arg-60 (NH_1) and Lys-92 (NZ) (the equivalent of the Grd19p-Lys112). Grd19p has Ser-83 at the equivalent position for $p40^{phox}$-Arg-60, but its side chain is too far for interaction with the 1-phosphate group. Mutation of $p40^{phox}$-Lys-92 greatly decreases affinity for PtdIns(3)P, and mutation of Arg-60 to Ala decreases affinity 3-fold (13). Mutagenesis studies have indicated that Lys-92 is important for affinity but that the mutation of Arg-60 to Ala causes only a 3-fold loss (13).

The crystal structure of the PX domain of the $p47^{phox}$ revealed the presence of two well bound sulfate ions in two distinct basic pockets (16). The first sulfate ion in the $p47^{phox}$ structure occupies an identical position to the 3-phosphate group in the Grd19p complex, and both groups are stabilized by analogous interactions with the protein. The second sulfate of the $p47^{phox}$ PX domain is bound at the irregular exit of the first helix and the start of the proline-rich region. $p47^{phox}$ provides a positively charged pocket that favorably interacts with the sulfate ion (His-51, Lys-55, Arg-70, and His-74). It was suggested and biochemically verified that this pocket provides a binding site for phospholipids with a small head group, such as phosphatidic acid, cooperatively binding with the PtdIns(3)P binding region. The corresponding region in Grd19p adopts a rather different structural conformation (not shown). Grd19p provides residues that could interact with an anion group of a phospholipid: His-108 occupies about the same position as His-74 in $p47^{phox}$, and Arg-89 is superposable to His-51, however it has no equivalent residues for $p47^{phox}$ Arg-70 and Lys-55.

In the resting state of $p47^{phox}$, its PX domain is masked by an interaction with a SH3 domain within the same protein. This SH3 domain binds to the proline-rich region, and this interaction is regulated by the phosphorylation state of its C-terminal part (indicated in Fig. 2*A*). The proline-rich region is partially present in Grd19p (containing Pro-107 and Pro-110), but this protein is not documented to interact with SH3 domains. However, Grd19p is specifically involved in the retrieval of late Golgi membrane proteins, DPAP A and Kex2p, from the prevacuolar compartment (30, 31). Direct physical interaction was demonstrated between the cytosolic domain of DPAP A and Grd19p, but not with other PX-containing domains from yeast (10). The presence of a F*X*F*X*D retrieval motif in the membrane proteins contributed to the stability of the interaction. This motif consists of a cluster of basic amino acids and is followed by a few acidic residues (30, 31). The interaction depends on the presence of Mg^{2+} and is resistant to high salt concentrations. Nothing is known about the structural determinants of this interaction regarding Grd19p. In absence of further biochemical data it is difficult to speculate about potential membrane protein interaction sites on the surface of Grd19p. We are in the stage of characterizing these interactions using peptides derived from the tails of DPAP and Kex2p.

CONCLUSION

This paper describes the first comparison between a PX protein in ligand-bound and free form. The overall structure of Grd19p is very similar to other PX domain structures, despite very low sequence similarity. The position of the diC_4PtdIns(3)P in the recognition site of Grd19p shows that the mode of interaction is very similar to what was observed for the complex of the $p40^{phox}$ PX domain. The Grd19p inositol-binding site undergoes small but significant structural changes upon ligand binding. The structure allows us now to go further into the characterization of the late Golgi membrane protein retrieval process through the study of interactions with peptides from the cytosolic regions and the functional characterization of surface mutants.

Acknowledgment—We thank the staff from the European Synchrotron Radiation Facility beam lines for help with data collection.

REFERENCES

1. Martin, T. F. (1998) *Annu. Rev. Cell Dev. Biol.* **14,** 231–264
2. Stenmark, H., Aasland, R., Toh, B. H., and D'Arrigo, A. (1996) *J. Biol. Chem.* **271,** 24048–24054
3. Shaw, G. (1996) *Bioessays* **18,** 35–46
4. Itoh, T., Koshiba, S., Kigawa, T., Kikuchi, A., Yokoyama, S., and Takenawa, T. (2001) *Science* **291,** 1047–1051
5. Ponting, C. P. (1996) *Protein Sci.* **5,** 2353–2357
6. Xu, Y., Seet, L. F., Hanson, B., and Hong, W. (2001) *Biochem. J.* **360,** 513–530
7. Worby, C. A., and Dixon, J. E. (2002) *Nat. Rev. Mol. Cell. Biol.* **3,** 919–931
8. Kurten, R. C., Cadena, D. L., and Gill, G. N. (1996) *Science* **272,** 1008–1010
9. Haft, C. R., de la Luz Sierra, M., Barr, V. A., Haft, D. H., and Taylor, S. I. (1998) *Mol. Cell. Biol.* **18,** 7278–7287
10. Voos, W., and Stevens, T. H. (1998) *J. Cell Biol.* **140,** 577–590
11. Xu, Y., Hortsman, H., Seet, L., Wong, S. H., and Hong, W. (2001) *Nat. Cell Biol.* **3,** 658–666
12. Yu, J. W., and Lemmon, M. A. (2001) *J. Biol. Chem.* **276,** 44179–44184
13. Bravo, J., Karathanassis, D., Pacold, C. M., Pacold, M. E., Ellson, C. D., Anderson, K. E., Butler, P. J., Lavenir, I., Perisic, O., Hawkins, P. T., Stephens, L., and Williams, R. L. (2001) *Mol. Cell* **8,** 829–839
14. Hiroaki, H., Ago, T., Ito, T., Sumimoto, H., and Kohda, D. (2001) *Nat. Struct. Biol.* **8,** 526–530
15. Lu, J., Garcia, J., Dulubova, I., Sudhof, T. C., and Rizo, J. (2002) *Biochemistry* **41,** 5956–5962
16. Karathanassis, D., Stahelin, R. V., Bravo, J., Perisic, O., Pacold, C. M., Cho, W., and Williams, R. L. (2002) *EMBO J.* **21,** 5057–5068
17. Cheever, M. L., Sato, T. K., de Beer, T., Kutateladze, T. G., Emr, S. D., and Overduin, M. (2001) *Nat. Cell Biol.* **3,** 613–618
18. Jancarik, J., and Kim, S.-H. (1991) *J. Appl. Cryst.* **24,** 409–411
19. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491–497
20. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
21. Terwilliger, T. C. (1999) *Acta Cryst. D.* **55,** 849–861
22. Terwilliger, T. C. (2003) *Acta Cryst. D.* **59,** 45–49
23. Perrakis, A., Morris, R. J., and Lamzin, V. S. (1999) *Nature Struct. Biol.* **6,** 458–463
24. Murshudov, G., Vagin, A., and Dodson, E. (1997) *Acta Cryst. D.* **53,** 240–255
25. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
26. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D.* **54,** 905–921
27. Teasdale, R. D., Loci, D., Houghton, F., Karlsson, L., and Gleeson, P. A. (2001) *Biochem. J.* **358,** 7–16
28. Kanai, F., Liu, H., Field, S. J., Akbary, H., Matsuo, T., Brown, G. E., Cantley, L. C., and Yaffe, M. B. (2001) *Nat. Cell Biol.* **3,** 675–678
29. Xu, J., Liu, D., Gill, G., and Songyang, Z. (2001) *J. Cell Biol.* **154,** 699–705
30. Nothwehr, S. F., Roberts, C. J., and Stevens, T. H. (1993) *J. Cell Biol.* **121,** 1197–1209
31. Wilcox, C. A., Redding, K., Wright, R., and Fuller, R. S. (1992) *Mol. Biol. Cell* **3,** 1353–1371
32. Higgins, D., Thompson, J., Gibson, T., Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
33. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) *Bioinformatics* **15,** 305–308