Crystal Structure of the Human GGA1 GAT Domain†

Guangyu Zhu,‡ Peng Zhai,‡ Xiangyuan He,§ Simon Terzyan,‡ Rongguang Zhang,|| Andrzej Joachimiak,|| Jordan Tang,§,⊥ and Xuejun C. Zhang*,‡

Crystallography Research Program and Protein Studies Program, Oklahoma Medical Research Foundation, 825 Northeast 13th Street, Oklahoma City, Oklahoma 73104, Structure Biology Center, Advanced Photon Source, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439, and Department of Biochemistry and Molecular Biology, Oklahoma University Medical Center, Oklahoma City, Oklahoma 73104

Received February 27, 2003; Revised Manuscript Received April 3, 2003

ABSTRACT: GGAs are a family of vesicle-coating regulatory proteins that function in intracellular protein transport. A GGA molecule contains four domains, each mediating interaction with other proteins in carrying out intracellular transport. The GAT domain of GGAs has been identified as the structural entity that binds membrane-bound ARF, a molecular switch regulating vesicle−coat assembly. It also directly interacts with rabaptin5, an essential component of endosome fusion. A 2.8 Å resolution crystal structure of the human GGA1 GAT domain is reported here. The GAT domain contains four helices and has an elongated shape with the longest dimension exceeding 80 Å. Its longest helix is involved in two structural motifs: an N-terminal helix−loop−helix motif and a C-terminal three-helix bundle. The N-terminal motif harbors the most conservative amino acid sequence in the GGA GAT domains. Within this conserved region, a cluster of residues previously implicated in ARF binding forms a hydrophobic surface patch, which is likely to be the ARF-binding site. In addition, a structure-based mutagenesis−biochemical analysis demonstrates that the C-terminal three-helix bundle of this GAT domain is responsible for the rabaptin5 binding. These structural characteristics are consistent with a model supporting multiple functional roles for the GAT domain.

Eukaryotic cells transport proteins and lipids, as well as nutrients, via exocytic and endocytic pathways, which are usually carried out by transporting vesicles. GGAs (i.e., Golgi-localizing, γ-adaptin ear homology domain, ARF-binding proteins) are a family of cytosolic proteins that mediate the association of vesicle−coat assembly with the *trans*-Golgi network (TGN)[1] membrane in the presence of another protein called ARF (i.e., ADP ribosylation factor) (*1*−*3*). The interaction of these two proteins is important for cargo sorting (*4*) and the recruitment of clathrin and accessory proteins (*5*) in the vesicle budding process. In addition, some GGAs have been found to interact with rabaptin5, a Rab4/Rab5 effector regulating endosome fusion, and are implicated in a vesicle fusion process (*3*, *6*, *7*). Three GGAs have been identified in humans (GGA1−3), two in yeast (Gga1p and Gga2p) (*1*−*3*, *8*−*11*), and one each in *Caenorhabditis elegans* (*12*) and *Drosophila melanogaster* (*5*). GGA1−3 are found ubiquitously expressed in adult human tissues (*1*). All GGAs share a common modular domain structure with distinct functions assigned to each of the four domains. At the N-terminus is the VHS (VPS27, Hrs, and STAM) domain, which binds the sorting signals of receptors, such as mannose 6-phosphate receptors (*13*) and memapsin2 (β-secretase) (*14*). This VHS-sorting signal interaction is thought to be the cargo recognition step in vesicle assembly. The second domain of a GGA protein, GAT (GGA and TOM1, ∼160 residues), is among the most conserved regions in the primary sequence of GGAs (*1*) (Figure 1) and comprises the ARF-binding activity of GGA (*15*). In some GGAs, the GAT domain directly interacts with rabaptin5 (*6*). The third (hinge) domain interacts with clathrin (*5*, *7*) and contains an autoinhibition acidic cluster−dileucine motif similar to that in the sorting signal of the cargo proteins (*16*). At the C-terminus is the GAE domain, which is homologous to the ear domain of the γ-adaptin subunit of the adaptor protein-1 (AP-1) complex. It interacts with, interestingly, the γ-adaptin ear domain of AP-1 as well as γ-synergin, rabaptin5, and other potential regulators of vesicle−coat assembly (*3*, *8*). Some of the cellular functions of GGAs may not be limited to its immediate protein family members. In this regard, the sequence making up the C-terminal two-thirds of the GGA GAT domains has a low but recognizable homology with the second domain of TOM1 (i.e., target of myb-1), which also contains an N-terminal VHS domain (*17*). The function of TOM1, however, has yet to be firmly established.

† This work was supported in part by NIH Grant AG-18933 and Alzheimer's Association Pioneer Award to J.T. X.C.Z. was supported by NIH Grant HL60626.

* To whom correspondence should be addressed. Tel: (405) 271-7402. Fax: (405) 271-7953. E-mail: zhangc@omrf.ouhsc.edu.

‡ Crystallography Research Program, Oklahoma Medical Research Foundation.

§ Protein Studies Program, Oklahoma Medical Research Foundation.

|| Argonne National Laboratory.

⊥ Oklahoma University Medical Center.

[1] Abbreviations: ARF, ADP ribosylation factor; GAP, GTPase activating proteins; GAT, GGA and TOM homologous domain; GEF, guanidine−nucleotide exchange factor; GGA, Golgi-localized, γ-ear-containing, ARF-binding protein; TOM1, target of myb1; TGN, *trans*-Golgi network; VHS, VPS27, Hrs, and STAM homologous domain.

10.1021/bi034334n CCC: $25.00
Published on Web 05/06/2003

FIGURE 1: Sequence alignment of GAT domains. Numbering and secondary structural elements (i.e., α-helices) are shown for GGA1 GAT at the top. GenBank sequences of AF233521 (human GGA1, i.e., hGGA1), AF233522 (hGGA2), AF219138 (hGGA3), Z68014 (*C. elegans* GGA, i.e., cGGA), NM_132343 (*D. melanogaster* GGA, i.e., dGGA), NP_010645 (yeast Gga1p, i.e., yGga1p), NP_011976 (yGga2p), and AJ006973 (hTOM1) were used. Sequence alignment was carried out with VectorNT (Informax Inc., Bethesda, MD) followed by manual adjustment in some gap regions at the termini of helices. Residues identical to those of GGA1 are in bold font. Filled downward triangles represent the positions of reported point mutations inhibitory to ARF binding; the open downward triangle represents a neutral point mutation; open upward triangles represent the C-terminal ends of some truncation constructs that (partially) retain the ARF-binding activity (*1*, *5*, *15*, *36*). GGA1 GAT residues that were omitted from the final refined model are underlined. This figure was produced with the program AlScript (*46*).

ARF proteins, a member of the Ras-like small GTPase superfamily (*18*), are the first group of interacting protein partners identified for GGAs (*1*). They were initially demonstrated to stimulate the ADP-ribosyltransferase activity of the cholera toxin A subunit (*19*). ARF proteins are now known to regulate many aspects of membrane traffic involving the Golgi and TGN (*20*, *21*). Their main function is to recruit multimeric coat complexes, including those containing GGAs, from the cytosol to the vesicle budding sites (*1*, *5*). In mammals, ARFs are divided into three classes on the basis of gene structure, sequence similarity, and cellular location (*22*). Members of all three classes have been shown to interact with GGA proteins (*15*). Like other Ras-like GTPases (*23*), ARF cycles between GTP-bound (active) and GDP-bound (inactive) forms, thus functioning as a molecular switch. In the GTP-bound form, which is promoted by the guanidine−nucleotide exchange factor (GEF), ARF associates with the membrane of specific cellular compartments (*24*, *25*) and interacts with its effectors such as a GGA protein. In contrast, GDP-bound ARFs are usually located in the cytosol. The intrinsic GTPase activity of ARF is very low; the 'turn-off' process is stimulated by the ARF-specific GTPase activating proteins (GAPs) (*26*−*29*) and synergistically enhanced by clathrin (*30*). While GGAs are devoid of GEF or GAP activity (*1*), the binding of GGA to ARF prevents the GAP from accelerating GTP hydrolysis and thus stabilizes the active form of ARF (*5*).

The well-conserved GAT domain is essential for the GTP-dependent ARF-binding activity of GGAs (*2*). This GAT−ARF interaction, which is universal between all GGA and ARF isoforms (*15*), appears to be sufficient for targeting a reporter protein to TGN in mammalian cells (*31*). Therefore, information on the three-dimensional structure of the GGA GAT domain is fundamental for understanding the structural basis of ARF binding. Here, we present a crystal structure of the GAT domain of human GGA1 at 2.8 Å resolution. The ARF-binding site can be traced to a conserved hydrophobic patch containing a cluster of residues previously found individually to be involved in ARF binding. Furthermore, our mutagenesis analysis demonstrates that a three-helix bundle in this GAT domain assumes full binding affinity between GAT and rabaptin5.

MATERIALS AND METHODS

Protein Expression. The cDNA of human GGA1 (*3*) was kindly provided by Dr. M. Robinson (University of Cambridge, Cambridge, U.K.). For the construct of the glutathione *S*-transferase (GST)−GAT fusion protein, the cDNA encoding the GGA1 GAT domain, plus short extensions at both ends of the peptide chain (i.e., residues 141−326), was amplified using polymerase chain reaction, cloned into plasmid pGEX2T (Amersham Biosciences, Piscataway, NJ), and expressed in *Escherichia coli* strain BL21. Three hours after induction with 0.1 mM isopropyl β-D-thiogalactoside at 0.6 OD_{600}, the bacterial cells were harvested and lysed with lysozyme. The recombinant protein was purified from the supernatant of the cell lysate by affinity chromatography using a glutathione−Sepharose 4B column (Amersham Biosciences). The GAT domain was separated from gel-immobilized GST after thrombin cleavage and was further

purified with Resource Q ion-exchange chromatography (Amersham Biosciences). The GGA1 GAT domain fractions were eluted at about 150 mM NaCl concentration, pooled, and dialyzed against a buffer of 10 mM Tris-HCl (pH 8.5) and 0.03% (v/v) β-mercaptoethanol. The sample was then concentrated to 7 mg/mL and stored on ice. Similarly, recombinant proteins of GGA1 GAT domain fragments containing residues 142−210 and 210−326 (i.e., $GAT^{142-210}$ and $GAT^{210-326}$) were expressed and purified as GST fusion proteins. Furthermore, the cDNA fragment encoding residues 551−862 of human rabaptin5 was subcloned into the pET15b vector (Novagen, Madison, WI) on the basis of the full-length cDNA kindly provided by Dr. G. Grenningloh (University of Lausanne, Lausanne, Switzerland). The recombinant His-tag fusion protein was purified with affinity chromatography using the Hi-Trap chelating column (Amersham Biosciences) charged with Ni^{2+}. The fusion protein was eluted with a buffer containing 0.2 M imidazole and further purified using Resource Q ion-exchange chromatography. The protein fractions eluted at about 250 mM NaCl concentration were dialyzed against a buffer of 10 mM Tris-HCl (pH 8.0) and 0.1% (v/v) β-mercaptoethanol, concentrated to 5 mg/mL, and stored on ice.

Crystallization. The crystals of the human GGA1 GAT domain (residues 141−326) were grown at 20 °C with the hanging drop method. The protein solution was mixed 2:1 (v/v) with the reservoir solution of 0.4 M Li_2SO_4, 0.1 M 2-(*N*-morpholino)ethanesulfonic acid (pH 5.5), 20% (v/v) ethylene glycol, and 0.1% (v/v) β-mercaptoethanol. Crystals appeared from a 6 μL drop over a 500 μL reservoir in 1 day and grew to ~1.0 × 0.3 × 0.3 mm in about 1 week. Similar crystal forms could be obtained without ethylene glycol. However, additional ethylene glycol allowed the crystal to be directly transferred to the 100 K nitrogen gas stream for data collection without a significant loss in data quality. The Se-Met-substituted GAT protein was prepared by expressing the protein in *E. coli* strain B834(DE3), grown in minimal media supplemented with Se-Met (Sigma, St. Louis, MO), purified, and crystallized using a protocol similar to that of the native protein. The crystal size (~0.1 × 0.1 × 0.2 mm), however, was smaller than the native one.

Data Collection and Processing. X-ray diffraction data from the native crystal of human GGA1 GAT domain were collected up to 2.8 Å resolution from a MAR345 image plate (Mar Research Inc., Norderstedt, Germany) and Rigaku X-ray generator (Molecular Structure Co., Woodlands, TX) equipped with an Osmic mirror system (Osmic Inc., Troy, NY). The native crystal produced a light ice ring in the diffraction pattern, which resulted in reduced data completeness (83%) in the resolution range of 3.4−4.0 Å. A heavy atom derivative was prepared by supplementing the crystal drop with 10% (v/v) 10 mM $K_2Pt(CN)_4$ and incubated for 16 h. The derivative crystal diffracted to 3.0 Å. In addition, a three-wavelength multiple anomalous dispersion (MAD) data set was collected to 3.5 Å resolution at the beam line SBC 19ID of the Advanced Photon Source at Argonne National Laboratory. Data were processed and scaled with HKL (*32*).

Structural Determination and Refinement. The Matthews coefficient V_M of the crystal form was estimated to be 3.3 $Å^3$/Da, assuming one protein molecule per crystal asymmetric unit. Most of the following crystallographic calculations were performed with the program suite CNS (*33*). A heavy atom search yielded a single site for the platinum derivative and six out of the seven selenium sites from the Se-Met crystal that were predicted on the basis of the amino acid sequence. Combination of the SIR phases of the platinum derivative and the MAD phases of the Se-Met derivative yielded an initial experimental electron density map at 3.0 Å resolution, which was further improved and extended to 2.8 Å resolution by solvent flattening and using the data set collected from the native crystal. The electron density map was interpretable after this density modification (see Figure 2), with a figure of merit of 0.89 for the combined phases. Manual model building was carried out with the program Turbo-Frodo (*34*). Structural refinement consisted of several rounds of torsion angle simulated annealing and restrained individual *B*-factor refinements, using a maximum likelihood target function against native data up to the highest resolution available and with experimental phases included at the early stages of refinement. Although the average *B*-factor of the final model appeared high, it was consistent with the value estimated from the Wilson plot (73 $Å^2$). Similar average *B*-factors are not uncommon for protein crystal structures at comparable resolutions reported to the Protein Data Bank. Coordinates and the experimental structural factors of this GAT crystal structure have been deposited in the Protein Data Bank under code 1oxz.

Rabaptin5−GAT Affinity Assay. Fifty microliters of glutathione−Sepharose 4B gel containing ~100 μg of the recombinant fusion protein of GST−GGA1 GAT variants was incubated, individually, with 10 μL of $rabaptin5^{551-826}$ (5 μg/μL) at room temperature for 2 h with gentle agitation. The gel was then washed three times with 500 μL of phosphate-buffered saline, followed by resuspension in 50 μL of the same buffer. A sample of 5 μL of gel slurries was subjected to (20%) SDS−polyacrylamide gel electrophoresis (PAGE) under the reducing condition. Because of unavailability of antibodies against the specific fragment of rabaptin5 used, the $rabaptin5^{551-826}$ content was examined with Coomassie Blue stain. Since $rabaptin5^{551-826}$ may contain free cysteines, this experiment was repeated in both the presence and absence of 0.2% (v/v) β-mercaptoethanol at the binding and washing steps. The results were essentially the same.

RESULTS

Structural Description. The crystal structure of the GGA1 GAT domain was determined at 2.8 Å resolution using SIR and MAD methods. Statistics of the data collection and refinement are summarized in Table 1. The refined model contained four α-helices, i.e., A1 of residues 171−183, A2 of residues 187−233, A3 of residues 244−269, and A4 of residues 273−301, connected by short loops (Figure 2). The electron density in the remaining regions of the peptide chain was weak and difficult to interpret even after convergence of the refinement. As a result, residues 141−170 and 303−326, including short extensions from both N- and C-termini of the GAT domain, were omitted from the final model. The GGA1 GAT domain is generally considered to consist of residues ~147−313 (*5, 15*). However, a recombinant protein of this shorter region did not produce a useful crystal in our hands. Therefore, we consider that the refined model represents the core of the GGA1 GAT domain. Our crystal structure agrees well with previous predictions on locations

Table 1: Statistics of Crystallographic Data, Phasing, and Refinement

Data Collection (Space Group $P6_3$; Cell Dimensions $a = b = 83.4$ Å, $c = 69.0$ Å, $\gamma = 120°$)					
	native	SeMet λ_1	SeMet λ_2	SeMet λ_3	Pt derivative
wavelength used (Å)	1.5418	0.94644	0.97938	0.97952	1.5418
resolution range (Å)	50 (2.90)[a]−2.80	50 (3.67)−3.50	50 (3.67)−3.50	50 (3.67)−3.50	50 (3.11)−3.00
no. of measurements	56805	23711	22953	23264	39461
no. of unique reflections	6561 (676)	3341 (328)	3328 (307)	3323 (307)	5121 (523)
completeness (%)	96.1 (100)	98.5 (99.4)	98.1 (93.0)	98.0 (93.6)	93.0 (98.9)
$I/\sigma(I)$	42.1 (4.2)	16.2 (2.3)	12.9 (1.8)	14.6 (2.0)	34.1 (3.8)
R_{merge}[b] (%)	4.4 (48.9)	9.6 (66.8)	11.6 (64.8)	9.9 (61.0)	5.0 (51.7)

Observed Diffraction Ratios[c]			
	λ_1	λ_2	λ_3
λ_1	0.087	0.092	0.099
λ_2		0.126	0.085
λ_3			0.100

Phase Power[d] and Figure of Merit (FOM)			
	MAD	SIR	combined
resolutions (Å)	50 (3.7)−3.5	50 (3.1)−3.0	50−3.0
phase power			
$\lambda_1-\lambda_1^-$	1.27 (0.43)		
$\lambda_1-\lambda_2^+$	0.76 (0.61)		
$\lambda_1-\lambda_2^-$	1.62 (0.83)		
$\lambda_1-\lambda_3^+$	1.04 (0.55)		
$\lambda_1-\lambda_3^-$	1.70 (0.93)		
Pt−native		1.70 (0.75)	
FOM	0.55 (0.27)	0.29 (0.15)	0.43
R_{iso}[e] (%)		14	
R_{Cullis}[f] (%)		63	

Refinement	
resolution (Å)	50−2.8
$R_{working}$[g] (%)/no. of reflections	24.1/5714
R_{free}[g] (%)/no. of reflections	28.5/699
no. of non-hydrogen atoms	1052
rms deviation from ideal values	
bond length (Å)	0.010
bond angle (deg)	1.4
improper angle (deg)	0.79
dihedral angle (deg)	19.2
average *B*-factor (Å^2)	93.1

[a] Values in parentheses are for the highest resolution bin. [b] $R_{merge} = \sum_h\sum_i|I_i(h) - \langle I(h)\rangle|/\sum_h\sum_i I_i(h)$, where $I_i(h)$ is the ith measurement and $\langle I(h)\rangle$ is the mean of all measurements of $I(h)$ for Miller indices h. [c] Values are $\langle(\Delta F)^2\rangle^{1/2}/\langle F^2\rangle^{1/2}$, where ΔF is the dispersive (off-diagonal elements) or Bijvoet difference (diagonal elements) computed between 50 and 3.5 Å resolution. [d] MAD phasing power is defined as $[\langle|F_D - F_N|^2\rangle/\int\varphi P(\varphi)(|F_N \exp(i\varphi) + \Delta F_H| - F_D)^2 d\varphi]^{1/2}$, where $P(\varphi)$ is the experimental phase probability distribution, F_N is the amplitude of the structure factor at the reference wavelength λ_1, F_D corresponds to the structure factor at wavelength λ_i (indicated by a superscript "+") or its Friedel mate (indicated by a superscript "−"), and ΔF_H is the difference in heavy atom structure factors between the two wavelengths. Similarly, the phasing power of a derivative is defined between the derivative and native data. [e] $R_{iso} = \sum_h|F_P - F_{PH}|/\sum_h F_P$, where F_P and F_{PH} are the observed structure factor amplitudes of the native and derivative data sets, respectively. [f] $R_{Cullis} = \sum_h|F_{PH} \pm F_P| - F_H|/\sum_h|F_{PH} \pm F_P|$, where F_H is amplitude of the calculated structural factor for the heavy atom. [g] $R = \sum(F_o - kF_c)/\sum F_o$; all reflections of $F > 0$ in the resolution range were included.

of helices in the primary sequence of the GAT domain (*2*, *5*), reflecting a dominant helical content. A structural homology search with the online program Dali (at http://www2.ebi.ac.uk/dali/), however, did not produce a useful guide for the model building of the current structure. The overall shape of the refined model was dominated by the 47 residue long helix A2, with the longest dimension of the protein molecule being over 80 Å. The helix A1 and N-terminal half of A2 participated in a helix−loop−helix structure. The C-terminal half of A2 and helices A3 and A4 formed a left-hand twisted three-helix bundle. The long helix A2 was slightly curved (∼17°), and the helix A3 had a 28° kink around the fairly conserved Pro 261. The angle between the two helix axes in the helix−loop−helix motif was 35°, and their distance was 8.6 Å. The left-hand twisted three-helix bundle had interhelix angles ranging from 25° to 32° and interhelix distances between 9.1 and 9.5 Å, consistent with previously reported statistics on interhelix geometry (*35*). In contrast, crystal packing of the GAT molecules resulted in less compact interhelix interactions between symmetry-related molecules.

The molecular surface of the GGA1 GAT domain exhibited a nonrandom electric charge distribution (Figure 3a). Along the long dimension, on one side of the molecule were basic residues Arg 191, Lys 195, Lys 198, Lys 206, Lys 250, Arg 255, Arg 258, Arg 265, Lys 295, and Arg 299, which collectively formed a long patch of positive charges. Clustered on the opposite side surface, including the loop region connecting helices A3 and A4, were acidic residues Glu 172, Glu 173, Glu 203, Asp 204, Glu 271, Asp 274, Glu 275, Glu 279, and Asp 285. The residues participating in the positively charged strip on the GGA1 GAT surface are not conserved among GGA proteins, thus suggesting that these charges are nonessential to common GGA functions. In contrast, the acidic cluster in the loop between helices A3 and A4 is well conserved, although its function remains

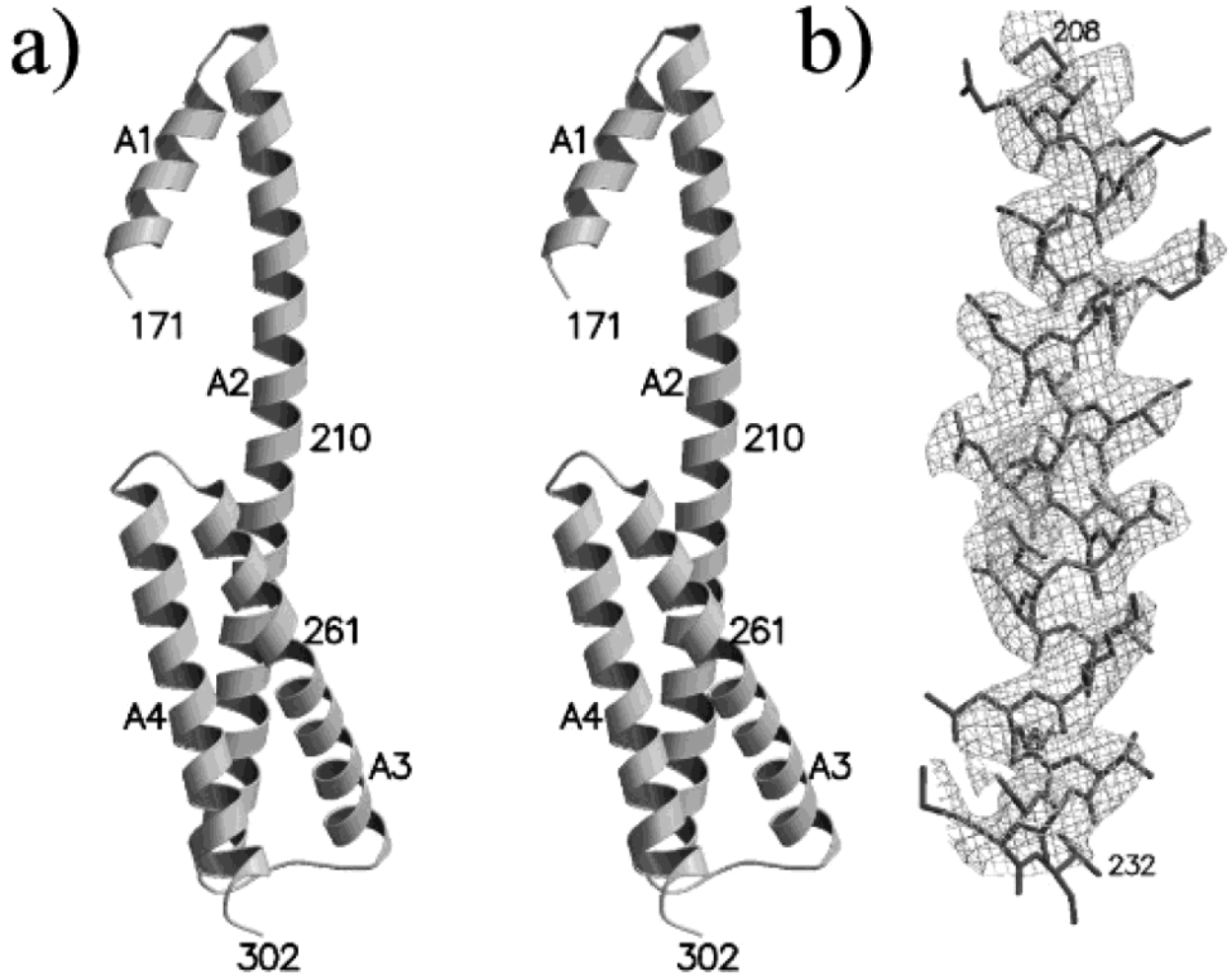

FIGURE 2: Three-dimensional structure of the GGA1 GAT domain. (a) Stereoview of a ribbon presentation of the overall structure. Helices are labeled as A1−A4. **(**b) Representative region of the experimental electron density map. A 2.8 Å map, initially phased with a combination of SIR and MAD methods and enhanced by solvent flattening, is contoured at 1.0 standard deviation and superimposed with the final refined model. These figures were produced with the program MolScript (*47*).

unclear at the present. Most of the hydrophobic residues were located inside the core of the N-terminal helix−loop−helix motif and the C-terminal three-helix bundle, consistent with the fact that GAT is an independently folded domain. There were, however, two large hydrophobic patches on the GAT surface that were involved in crystal contacts, thus excluded from solvent exposure. The first one consisted of the conserved residues Leu 182, Leu 190, Ile 197, and Val 201, which were located on one side of the helix−loop−helix motif (Figure 3b). Located within the C-terminal three-helix bundle, the second patch consisted of residues Pro 261, Phe 264, Leu 277, and Leu 281, most of which are conserved, at least among human GGA1−3. This latter region was packed around the crystallographic 3-fold axis, with the three phenylalanine side chains pointed toward each other. The functional significance of the first hydrophobic region will be discussed later.

The Three-Helix Bundle Motif of GGA1 GAT Is Sufficient To Bind with the Rabaptin5 C-Terminal Fragment. Because a number of mutagenesis experiments reported in the literature concur with the conclusion that the N-terminal half of the GAT domain is responsible for the ARF binding (*5*, *15*, *36*), we focused our effort on identification of the rabaptin5 binding motif in GGA1 GAT. Besides the full-length GAT domain, two truncation mutants, $GAT^{142-210}$ and $GAT^{210-326}$, were constructed, which contain the N-terminal helix−loop−helix and C-terminal three-helix bundle motifs, respectively. The C-terminal rabaptin5 fragment, consisting of residues 551−826, has previously been shown to be responsible for the GAT binding (*6*). Therefore, we carried out a GST-mediated pull-down experiment to identify the structural motif in GGA1 GAT that interacts with the recombinant rabaptin5 fragment, $rabaptin5^{551-826}$ (Figure 4). The result demonstrated that, under the experimental condition, the C-terminal three-helix bundle of GGA1 GAT binds to $rabaptin5^{551-826}$ (lane 4 in Figure 4) with an affinity comparable to that of the full-length GAT (lane 2). In contrast, the recombinant protein of the N-terminal helix−loop−helix motif alone lost binding ability toward rabaptin5 (lane 3).

DISCUSSION

The GGA GAT domain has been demonstrated to directly interact with ARFs (*15*) and rabaptin5 (*6*). The binding of ARF is necessary and sufficient to recruit GGAs to the TGN membranes in mammalian cells, while the binding of rabaptin5 has been taken as an implication of involvement in the membrane fusion. Our current structure lends some insight into the possible functional surface of the GAT domain that participates in these interactions. Overall, the

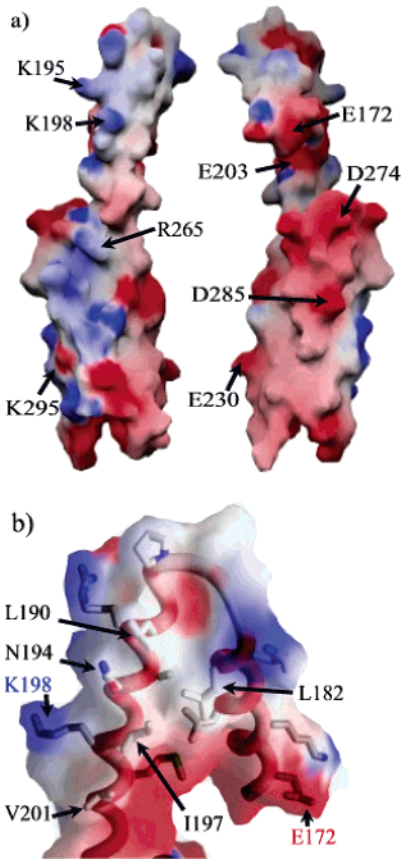

FIGURE 3: Molecular surface model of GGA1 GAT. (a) The overall structure. The color intensity corresponds to the calculated electrostatic potential from −8 kT/e (intense red) to +8 kT/e (intense blue). The orientation of the left figure is about the same as that of Figure 2a, and the right one is roughly a 180° rotation about a vertical axis from the left one. Some conservative charged residues are labeled. (b) A putative ARF-binding site. A worm model for the GAT backbone is superimposed on the semitransparent molecular surface model. Selected side chains are shown as stick models. Residues potentially interacting with ARF are labeled in black, together with Glu 172 and Lys 198 as positional references to the figures in panel a. These figures were produced with the program GRASP (*48*).

elongated shape of the current structure of the GGA1 GAT domain is consistent with its multifunctional roles. A number of biochemical and mutagenesis studies, including several truncation constructs of varied length in the GAT domain, have suggested that the ARF-binding affinity is confined to a region in the primary sequence that corresponds to helices A1 and A2 as well as their connecting loop (*1*, *5*, *15*). This region contains the most conserved amino acid sequence in the GAT domains among members of the GGA family (see Figure 1), consistent with their common function, i.e., ARF binding. Furthermore, single point mutations at positions Leu 178, Leu 182, Pro 187, Glu 188, Ala 193, Asn 194, and Val 201, which are located within this most conserved region, result in a disruption of ARF binding (*5*, *15*, *36*). They are positioned on the same side of the helix−loop−helix motif and form a flat, mostly hydrophobic surface. In contrast, mutation of Asp 189 (e.g., D189A), located outside of this hydrophobic patch, has no effect on activities related to ARF binding (*5*). When combining all the evidence, it seems reasonable to hypothesize that this hydrophobic patch constitutes the primary ARF-binding site. Although the current structure does not suggest the involvement of other

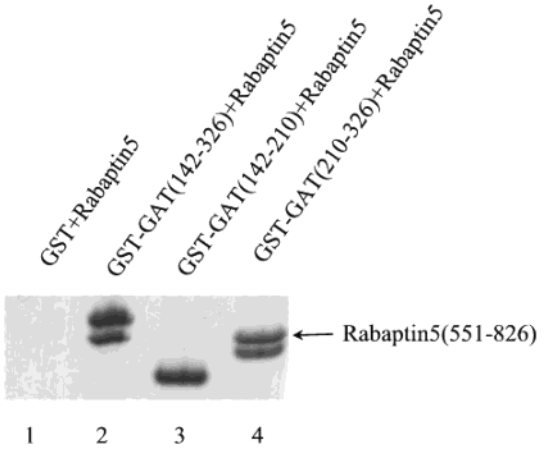

FIGURE 4: Analysis of the GAT−rabaptin5 interaction. GST-mediated pull-down assays were carried out between recombinant rabaptin5$^{551-826}$ and GST−GGA1 GAT variants. An aliquot of each pull-down sample was subjected to SDS−PAGE. The gel was subsequently stained with Coomassie Blue. Multiple repeats of this experiment yielded essentially identical results.

GAT regions in ARF binding, such a possibility cannot be ruled out. However, their effects, if they exist, are likely to be small because truncation mutagenesis experiments have demonstrated that GGA fragments of residues 1−226 and 170−232 individually retain ARF-binding ability (*5*, *15*); the overlap region of these two fragments roughly coincides with the N-terminal helix−loop−helix motif of our GAT crystal structure. In addition, the dimensions of an ARF molecule are ∼30 × 30 × 40 Å (*37*), while the distance from the center of the hydrophobic patch to the proximal end of the three-helix bundle in GAT is over 20 Å. Therefore, it is unlikely that the globular ARF molecule could have extensive interactions with both the N- and C-terminal motifs of the elongated GAT domain simultaneously, unless it bends significantly from the current structure, which has no supporting evidence at this time.

In the ARF molecule, the main GAT-binding region has been identified within the switch I region (i.e., residues 45−54) (*38*). From the crystal structure of GTP-bound ARF1, it is clear that the switch I and its surrounding region are predominantly hydrophobic in nature, lending the possibility of a direct interaction with the helix−loop−helix motif in GAT. This ARF region is conserved in sequence (*39*) but is known to assume different conformations, depending on the binding of either GTP or GDP (*40*), and to change the solvent accessibility of some residues including Phe 51, which is essential for GAT binding (*38*). It is also interesting to note that GGA proteins do not bind ARF homologues such as Arl1 and Arl2 (*41*), despite their 40−50% sequence identity with ARFs and the well-conserved sequence within the switch I region. Comparison of the GTP-bound Arl2 crystal structures (PDB code 1ksg) (*42*) with that of ARF shows that the main chain of residues 38−43 in ARF is significantly different from the corresponding part in Arf2, in addition to their amino acid sequence difference around this region (i.e., ^{37}LKLGE IVTTI PTIGF NV^{53} of ARF versus ^{36}FNGED VEIIS PTLGF NI^{52} of Arl2). Furthermore, the same region in ARF1 exhibits significant conformational changes between its GTP- and GDP-bound forms, supporting the notion that this region is an integral part of the GAT-binding site in ARF.

Rabaptin5, ~100 kDa in its monomeric form, has recently been shown to directly interact with the GAT domain of GGAs (*6*). In contrast with the monomeric GGAs (*2*, *3*), the active form of rabaptin5 is apparently a homodimer (*43*). The region of rabaptin5 primarily responsible for interaction with GAT is the coiled-coil region of residues ~551−661, with additional interaction supported by a C-terminal segment (*6*). In this study, we present evidence that a GAT domain interacts with rabaptin5 through the GAT C-terminal three-helix bundle. This observation confers a potentially important function to this previously vague region in the elongated GAT domain. The strength of the GAT−rabaptin5 interaction (see Figure 4) and the observation that rabaptin5 promotes conveyance of cargo from GGA-containing carriers into early endosomes argue that such interaction is biologically relevant (*6*). It has also been implicated that rabaptin5 and its associated proteins may interfere with the GGA−ARF interaction (*6*). Although direct evidence remains to be demonstrated, our crystal structure of the GAT domain suggests that such an interference is structurally feasible, especially considering that both GGA and rabaptin5 are multidomain proteins and involved in multimeric complexes. In view of the varied responses among GGAs to rabaptin5 interaction (*6*), their amino acid sequence variation in the three-helix bundle region is likely to provide clues to more details on the interaction and to the functional difference among GGA proteins in different intracellular trafficking pathways. More biochemical and structural studies along this direction would clearly be desirable.

The sequence alignment of GGA proteins (see Figure 1) suggests that their GAT domains are well conserved, as most of the insertions/deletions occur in loop regions between helices. A shorter A2 helix, by about one turn, is seen only in two yeast proteins. While the VHS domains in TOM1 and GGA are homologous in sequence (*44*), this similarity can be extended to only part of the GAT domain, corresponding to the three-helix bundle region but not in the N-terminal helix−loop−helix motif (Figure 1) (*2*). Furthermore, TOM1 does not seem to have a helical conformation at all in the region corresponding to the N-terminal half of the long helix A2, since six proline residues reside in this region. These observations are consistent with the notion that the helix−loop−helix motif in GGA GAT is responsible for the interaction with ARF; thus for regulating trafficking from the TGN, neither one is a function for TOM1 (*5*). Recently, two crystal structures of the human GGA1 GAT domain were deposited to the Protein Data Bank. One structure [PDB file 1naf (*45*)] contains essentially the same regions that we report here, and the other (1nwm) misses the N-terminal helix A1. Both crystal structures are consistent with our conclusions.

ACKNOWLEDGMENT

The authors thank N. Wakeham for excellent technical assistance and Drs. G. Li, G. Koelsch, L. Hong, and T. Mather for helpful discussion and critical reading of the manuscript.

REFERENCES

1. Boman, A. L., Zhang, C., Zhu, X., and Kahn, R. A. (2000) A family of ADP-ribosylation factor effectors that can alter membrane transport through the trans-Golgi, *Mol. Biol. Cell 11*, 1241−1255.
2. Dell'Angelica, E. C., Puertollano, R., Mullins, C., Aguilar, R. C., Vargas, J. D., Hartnell, L. M., and Bonifacino, J. S. (2000) GGAs: a family of ADP ribosylation factor-binding proteins related to adaptors and associated with the Golgi complex, *J. Cell Biol. 149*, 81−94.
3. Hirst, J., Lui, W. W., Bright, N. A., Totty, N., Seaman, M. N., and Robinson, M. S. (2000) A family of proteins with gamma-adaptin and VHS domains that facilitate trafficking between the trans-Golgi network and the vacuole/lysosome, *J. Cell Biol. 149*, 67−80.
4. Boman, A. L. (2001) GGA proteins: new players in the sorting game, *J. Cell Sci. 114*, 3413−3418.
5. Puertollano, R., Randazzo, P. A., Presley, J. F., Hartnell, L. M., and Bonifacino, J. S. (2001) The GGAs promote ARF-dependent recruitment of clathrin to the TGN, *Cell 105*, 93−102.
6. Mattera, R., Arighi, C. N., Lodge, R., Zerial, M., and Bonifacino, J. S. (2003) Divalent interaction of the GGAs with the Rabaptin-5-Rabex-5 complex, *EMBO J. 22*, 78−88.
7. Zhu, Y., Doray, B., Poussu, A., Lehto, V. P., and Kornfeld, S. (2001) Binding of GGA2 to the lysosomal enzyme sorting motif of the mannose 6-phosphate receptor, *Science 292*, 1716−1718.
8. Takatsu, H., Yoshino, K., and Nakayama, K. (2000) Adaptor gamma ear homology domain conserved in gamma-adaptin and GGA proteins that interact with gamma-synergin, *Biochem. Biophys. Res. Commun. 271*, 719−725.
9. Black, M. W., and Pelham, H. R. (2000) A selective transport route from Golgi to late endosomes that requires the yeast GGA proteins, *J. Cell Biol. 151*, 587−600.
10. Zhdankina, O., Strand, N. L., Redmond, J. M., and Boman, A. L. (2001) Yeast GGA proteins interact with GTP-bound Arf and facilitate transport through the Golgi, *Yeast 18*, 1−18.
11. Poussu, A., Lohi, O., and Lehto, V. P. (2000) Vear, a novel Golgi-associated protein with VHS and gamma-adaptin "ear" domains, *J. Biol. Chem. 275*, 7176−7183.
12. Washington University Genome Sequencing Center (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. The *C. elegans* Sequencing Consortium, *Science 282*, 2012−2018.
13. Puertollano, R., Aguilar, R. C., Gorshkova, I., Crouch, R. J., and Bonifacino, J. S. (2001) Sorting of mannose 6-phosphate receptors mediated by the GGAs, *Science 292*, 1712−1716.
14. He, X., Chang, W. P., Koelsch, G., and Tang, J. (2002) Memapsin 2 (beta-secretase) cytosolic domain binds to the VHS domains of GGA1 and GGA2: implications on the endocytosis mechanism of memapsin 2, *FEBS Lett. 524*, 183−187.
15. Takatsu, H., Yoshino, K., Toda, K., and Nakayama, K. (2002) GGA Proteins Associate with Golgi Membranes through Interaction between their GGAH Domains and ADP-ribosylation Factors, *Biochem. J. 365*, 369−378.
16. Doray, B., Bruns, K., Ghosh, P., and Kornfeld, S. A. (2002) Autoinhibition of the ligand-binding site of GGA1/3 VHS domains by an internal acidic cluster-dileucine motif, *Proc. Natl. Acad. Sci. U.S.A. 99*, 8072−8077.
17. Burk, O., and Klempnauer, K. H. (1999) Myb and Ets transcription factors cooperate at the myb-inducible promoter of the tom-1 gene, *Biochim. Biophys. Acta 1446*, 243−252.
18. Moss, J., and Vaughan, M. (1998) Molecules in the ARF orbit, *J. Biol. Chem. 273*, 21431−21434.
19. Kahn, R. A., and Gilman, A. G. (1986) The protein cofactor necessary for ADP-ribosylation of Gs by cholera toxin is itself a GTP binding protein, *J. Biol. Chem. 261*, 7906−7911.
20. Stamnes, M. A., and Rothman, J. E. (1993) The binding of AP-1 clathrin adaptor particles to Golgi membranes requires ADP-ribosylation factor, a small GTP-binding protein, *Cell 73*, 999−1005.
21. Tanigawa, G., Orci, L., Amherdt, M., Ravazzola, M., Helms, J. B., and Rothman, J. E. (1993) Hydrolysis of bound GTP by ARF protein triggers uncoating of Golgi-derived COP-coated vesicles, *J. Cell Biol. 123*, 1365−1371.
22. Moss, J., and Vaughan, M. (1995) Structure and function of ARF proteins: activators of cholera toxin and critical components of intracellular vesicular transport processes, *J. Biol. Chem. 270*, 12327−12330.
23. Vetter, I. R., and Wittinghofer, A. (2001) The guanine nucleotide-binding switch in three dimensions, *Science 294*, 1299−1304.
24. Donaldson, J. G., Finazzi, D., and Klausner, R. D. (1992) Brefeldin A inhibits Golgi membrane-catalysed exchange of guanine nucleotide onto ARF protein, *Nature 360*, 350−352.

25. Palmer, D. J., Helms, J. B., Beckers, C. J., Orci, L., and Rothman, J. E. (1993) Binding of coatomer to Golgi membranes requires ADP-ribosylation factor,. *J. Biol. Chem. 268*, 12083−12089.
26. Brown, M. T., Andrade, J., Radhakrishna, H., Donaldson, J. G., Cooper, J. A., and Randazzo, P. A. (1998) ASAP1, a phospholipid-dependent arf GTPase-activating protein that associates with and is phosphorylated by Src, *Mol. Cell. Biol. 18*, 7038−7051.
27. Cukierman, E., Huber, I., Rotman, M., and Cassel, D. (1995) The ARF1 GTPase-activating protein: zinc finger motif and Golgi complex localization, *Science 270*, 1999−2002.
28. Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., et al. (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*, *Nature 368*, 32−38.
29. Premont, R. T., Claing, A., Vitale, N., Freeman, J. L., Pitcher, J. A., Patton, W. A., Moss, J., Vaughan, M., and Lefkowitz, R. J. (1998) beta2-Adrenergic receptor regulation by GIT1, a G protein-coupled receptor kinase-associated ADP ribosylation factor GTPase-activating protein, *Proc. Natl. Acad. Sci. U.S.A. 95*, 14082−14087.
30. Goldberg, J. (1999) Structural and functional analysis of the ARF1-ARFGAP complex reveals a role for coatomer in GTP hydrolysis, *Cell 96*, 893−902.
31. Boman, A. L., Salo, P. D., Hauglund, M. J., Strand, N. L., Rensink, S. J., and Zhdankina, O. (2002) ADP−Ribosylation Factor (ARF) Interaction Is Not Sufficient for Yeast GGA Protein Function or Localization, *Mol. Biol. Cell 13*, 3078−3095.
32. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
33. 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) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
34. Roussel, A., and Cambillau, C. (1989) in *Silicon Graphics Geometry Partners Directory*, pp 77−79, Silicon Graphics, Mountain View, CA.
35. Senes, A., Ubarretxena-Belandia, I., and Engelman, D. M. (2001) The Calpha−H···O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions, *Proc. Natl. Acad. Sci. U.S.A. 98*, 9056−9061.
36. Boman, A. L., Salo, P. D., Hauglund, M. J., Strand, N. L., Rensink, S. J., and Zhdankina, O. (2002) ADP-Ribosylation Factor (ARF) Interaction Is Not Sufficient for Yeast GGA Protein Function or Localization, *Mol. Biol. Cell 13*, 3078−3095.
37. Goldberg, J. (1998) Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching, *Cell 95*, 237−248.
38. Kuai, J., Boman, A. L., Arnold, R. S., Zhu, X., and Kahn, R. A. (2000) Effects of activated ADP-ribosylation factors on Golgi morphology require neither activation of phospholipase D1 nor recruitment of coatomer, *J. Biol. Chem. 275*, 4022−4032.
39. Hosaka, M., Toda, K., Takatsu, H., Torii, S., Murakami, K., and Nakayama, K. (1996) Structure and intracellular localization of mouse ADP-ribosylation factors type 1 to type 6 (ARF1-ARF6), *J. Biochem. (Tokyo) 120*, 813−819.
40. Amor, J. C., Harrison, D. H., Kahn, R. A., and Ringe, D. (1994) Structure of the human ADP-ribosylation factor 1 complexed with GDP, *Nature 372*, 704−708.
41. Boman, A. L., Kuai, J., Zhu, X., Chen, J., Kuriyama, R., and Kahn, R. A. (1999) Arf proteins bind to mitotic kinesin-like protein 1 (MKLP1) in a GTP-dependent fashion, *Cell Motil. Cytoskel. 44*, 119−132.
42. Hanzal-Bayer, M., Renault, L., Roversi, P., Wittinghofer, A., and Hillig, R. C. (2002) The complex of Arl2-GTP and PDE delta: from structure to function, *EMBO J. 21*, 2095−2106.
43. Vitale, G., Rybin, V., Christoforidis, S., Thornqvist, P., McCaffrey, M., Stenmark, H., and Zerial, M. (1998) Distinct Rab-binding domains mediate the interaction of Rabaptin-5 with GTP-bound Rab4 and Rab5, *EMBO J. 17*, 1941−1951.
44. Misra, S., Beach, B. M., and Hurley, J. H. (2000) Structure of the VHS domain of human Tom1 (target of myb 1): insights into interactions with proteins and membranes, *Biochemistry 39*, 11282−11290.
45. Collins, B. M., Watson, P. J., and Owen, D. J. (2003) The Structure of the GGA1-GAT Domain Reveals the Molecular Basis for ARF Binding and Membrane Association of GGAs, *Dev. Cell 4*, 321−332.
46. Barton, G. J. (1993) ALSCRIPT a tool to format multiple sequence alignments, *Protein Eng. 6*, 37−40.
47. Kraulis, P. J. (1991) MOLSCRIPT-a program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−950.
48. Nicholls, A., Sharp, K. A., and Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins 11*, 281−296.

BI034334N