nature structural biology

© 2003 Nature Publishing Group http://www.nature.com/naturestructuralbiology

Recognition of accessory protein motifs by the γ-adaptin ear domain of GGA3

Gregory J Miller[1,3], Rafael Mattera[2,3], Juan S Bonifacino[2] & James H Hurley[1]

Adaptor proteins load transmembrane protein cargo into transport vesicles and serve as nexuses for the formation of large multiprotein complexes on the nascent vesicles. The γ-adaptin ear (GAE) domains of the AP-1 adaptor protein complex and the GGA adaptor proteins recruit accessory proteins to these multiprotein complexes by binding to a hydrophobic motif. We determined the structure of the GAE domain of human GGA3 in complex with a peptide based on the DFGPLV sequence of the accessory protein Rabaptin-5 and refined it at a resolution of 2.2 Å. The leucine and valine residues of the peptide are partly buried in two contiguous shallow, hydrophobic depressions. The anchoring phenylalanine is buried in a deep pocket formed by the aliphatic portions of two conserved arginine residues, along with an alanine and a proline, illustrating the unusual function of a cluster of basic residues in binding a hydrophobic motif.

The sorting of transmembrane proteins at various stages of the endocytic and secretory pathways is mediated by interactions between signals contained within the cytosolic domains of the transmembrane proteins and adaptor proteins that are components of membrane coats. Four adaptor protein (AP) complexes named AP-1, AP-2, AP-3 and AP-4 have such a role for specific subsets of transmembrane proteins in the endosomal-lysosomal system[1,2]. These complexes are structurally related heterotetramers composed of two large subunits (γ-, α-, δ- or ε-adaptin and β1–4 adaptins), one medium subunit (μ1–4) and one small subunit (σ1–4). The large subunits are organized into an N-terminal 'trunk' domain, an extended 'hinge' domain and a C-terminal 'ear' domain. The trunk domains of the two large subunits assemble with the medium and small subunits to constitute the 'core' of each AP complex. The hinge domains of the two large subunits extend from the core such that the two ear domains are situated at relatively long distances from the core[1,2]. In addition to the AP complexes, several monomeric proteins have recently been shown to function as adaptors for other subsets of transmembrane proteins. Among these monomeric adaptors are the GGA proteins GGA1, GGA2 and GGA3 (refs. 3–6), which sort mannose-6-phosphate receptors and other receptors from the *trans*-Golgi network (TGN) to endosomes[7–10]. The GGAs have a modular organization consisting of VHS (Vps27, Hrs, Stam), GAT (GGA and TOM), hinge and GAE domains[4]. The hinge domains of the GGAs have the same overall properties as those of the large AP subunits, although they do not have significant sequence similarity with one another. The GAE domain of the GGAs is similar to the C-terminal ear of the γ-adaptin subunit of AP-1 (refs. 3–6).

The signal-recognition and membrane-recruitment activities of the AP complexes and the GGAs are mediated by their core and VHS-GAT domains, respectively. The hinge and ear domains, on the other hand, recruit scaffolding proteins (such as clathrin in the case of AP-1, AP-2, AP-3 and the GGAs)[1,2] and various accessory proteins that mediate coat assembly and disassembly, vesicle budding and interactions with the cytoskeleton and vesicle fusion machinery[11]. The AP-2 α- and β2-subunit ear domains bind DPF, DPW[12] or FXDXF peptide motifs[13] on the accessory proteins. X-ray crystallographic analyses of the AP-2 α ear domain revealed that it consists of an N-terminal nine-stranded β-sandwich subdomain reminiscent of an immunoglobulin fold, and a C-terminal 'platform' subdomain made up of a five-stranded β-sheet flanked on either side by one and two α-helices, respectively[13,14]. The sandwich subdomain binds DPW-containing sequences, whereas the platform subdomain binds DPF, DPW and FXDXF motifs[13]. The β2 ear domain has a similar fold[15], although it seems to bind only DPF and DPW motifs to its platform subdomain[13,15].

Recent studies have begun to provide insights into the ligand-binding properties and three-dimensional structure of the ear domains of the AP-1 γ subunit (of which there are two isoforms named γ1 and γ2) and of the related GGA proteins (all of these domains are herein referred to as GAE). These domains bind sequences that seem to conform to a [DE]FXXΦ motif (Φ represents leucine, phenylalanine, tryptophan or methionine) found in a distinct group of accessory proteins including Rabaptin-5 (ref. 16), enthoprotin (also known as epsinR or Clint, see ref. 17 and references therein), γ-synergin[17], p56 (ref. 18) and, in yeast, the enthoprotin homologs Ent3p and Ent5p[19]. Analysis of the crystal structures of the unliganded GAE domains of γ1-adaptin and of GGA1 indicates that they consist of an eight-stranded β-sandwich similar to that of the α-adaptin and β2-adaptin ears, but without the platform subdomain found in the latter two[18,20,21]. The three-dimensional structure of a liganded GAE

[1]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, and [2]Cell Biology and Metabolism Branch, National Institute of Child Health and Human Development, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland 20892, USA. [3]These authors contributed equally to this work. Correspondence should be addressed to J.H.H. (james.hurley@nih.gov).

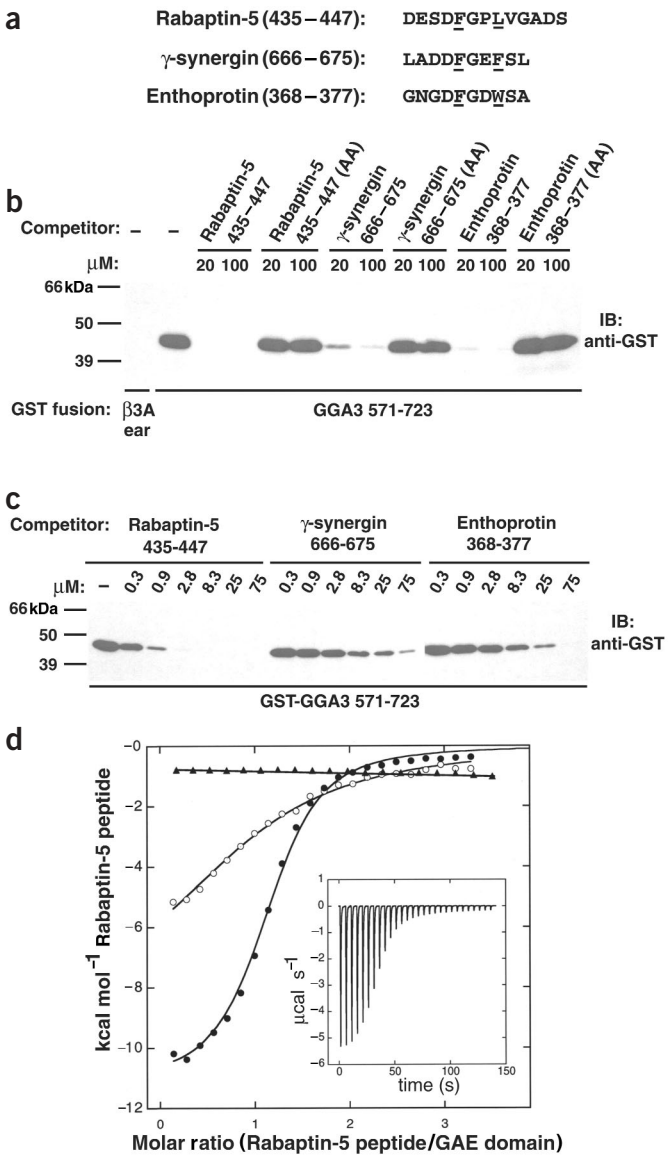

Figure 1 Binding of Rabaptin-5, γ-synergin and enthoprotin peptides to GGA3(571–723). (**a**) Binding assays were carried out with the indicated peptides, along with peptides with substitution of the underlined amino acids by alanine residues (designated as AA peptides in **b**). Each peptide begins with a cysteine residue (not shown) for biotinylation. (**b**) GST-GGA3(571–723) was preincubated in the presence or absence of the indicated concentrations of peptides and then incubated with biotinylated Rabaptin-5(435–447) immobilized on streptavidin-agarose (a GST-β3A ear construct was used as negative control). The GST-fusion proteins bound to the biotinylated Rabaptin-5 peptide were eluted and analyzed by SDS-PAGE and immunoblotting (IB) with an antiserum to GST. (**c**) Competition of binding of biotinylated Rabaptin-5(435–447) to GST-GGA3(571–723) by Rabaptin-5, γ-synergin and enthoprotin peptides at the indicated concentrations. Experiments were done as described in **b**. Positions of molecular-mass markers are indicated on left. (**d**) ITC analysis of Rabaptin-5 peptide binding to purified GGA3(571–723; closed circles) and to the T634M V648A (open circles) and R691Q (triangles) mutants. Data were fit to one-site binding models. Inset: differential heat released when 2 mM Rabaptin-5 peptide is injected into 100 μM GGA3 GAE domain. The trace is shown after subtraction of data from the injection of peptide into a buffer blank.

© 2003 Nature Publishing Group http://www.nature.com/naturestructuralbiology

domain, however, has not been resolved, and the mode of interaction of [DE]FXXΦ motifs with GAE domains remains to be determined.

Here we report the crystal structure at a resolution of 2.2 Å of the GAE domain of GGA3 in a complex with a [DE]FXXΦ-containing peptide from Rabaptin-5. The structure reveals a new mode of recognition of accessory proteins in which the aliphatic portions of basic residues in the GAE domain have a major role in binding hydrophobic side chains of the peptide motif.

RESULTS

Peptide binding to the GAE domain

We have previously demonstrated that the GAE domains of the GGAs and the γ1- and γ2-subunit isoforms of AP-1 bind a DFGPLV sequence in a central, unstructured region of Rabaptin-5 (ref. 16). GAE domains also bind sequences from enthoprotin and γ-synergin that seem to fit a tentative consensus motif, [DE]FXXΦ[17]. To assess whether sequences from all of these proteins bind to the same site on the GAE domain, we carried out competition experiments in which the binding of GST (glutathione *S*-transferase)-GGA3-GAE (residues 571–723) to a biotinylated DESDFGPLVGADS peptide derived from Rabaptin-5 was inhibited by different concentrations of Rabaptin-5, γ-synergin and enthoprotin peptides. The γ-synergin peptide we used, LADDFGEFSL, is contained within the C-terminal region of γ-synergin known to bind to γ-adaptin[22]. The enthoprotin peptide, GNGDFGDWSA, is similar to the P5 peptide used by Mills *et al.*[17] (**Fig. 1a**).

All three peptides inhibited binding, whereas variant peptides in which the phenylalanine and Φ residues were replaced by alanine residues were ineffective (**Fig. 1b**). The competition was concentration dependent and occurred with micromolar concentrations of peptides. The γ-synergin and enthoprotin peptides were less potent than the Rabaptin-5 peptide as competitors in this assay (**Fig. 1c**). Isothermal titration calorimetry (ITC) measurements showed that the Rabaptin-5 peptide binds to GGA3-GAE with $K_d = 5 \pm 1$ μM and a stoichiometry of virtually 1:1 (**Fig. 1d**).

Structure of the GAE–peptide complex

To elucidate the structural basis for the recognition of hydrophobic peptide motifs by GAE domains, we determined the structure of a T634M V648A mutant of the GGA3-GAE domain in complex with the DESDFGPLVGADS peptide from Rabaptin-5. The T634M V648A mutant was obtained fortuitously as a PCR artifact and cocrystallized with the peptide. These mutations decreased but did not abolish the interaction with Rabaptin-5 as judged by GST pull-down assays (**Fig. 2**) and ITC ($K_d = 29 \pm 4$ μM for the T634M V648A mutant as compared with 5 μM for the wild-type GGA3-GAE; **Fig. 1d**). The structure of the complex was resolved by multiwavelength anomalous dispersion (MAD) at the Se K-edge (**Fig. 3**; **Table 1**). The GGA3 (571–723) fragment comprises the C-terminal 28 residues of the hinge region and the entire GAE domain (**Fig. 4a**). All but the eight N-terminal residues of the GGA3(571–723) fragment are clearly apparent in the experimental electron density. The structure consists of an eight-stranded immunoglobulin fold β-sandwich, with one five-stranded sheet and one three-stranded sheet (**Fig. 4a**). The structure is similar to that of the GAE domain of the AP-1 γ1-adaptin subunit[20,21], as anticipated on the basis of the 34% sequence identity between the two. Of the 145 Cα positions in the GGA3-GAE structure, 109 can be superimposed on the γ1-adaptin GAE, with a r.m.s. deviation of 1.2 Å (**Fig. 4b**). The largest area of difference is in the N-terminal extension. Residues 579–595 of GGA3 form a long loop that wraps around the back of the five-stranded sheet. In contrast to the GGA1-GAE

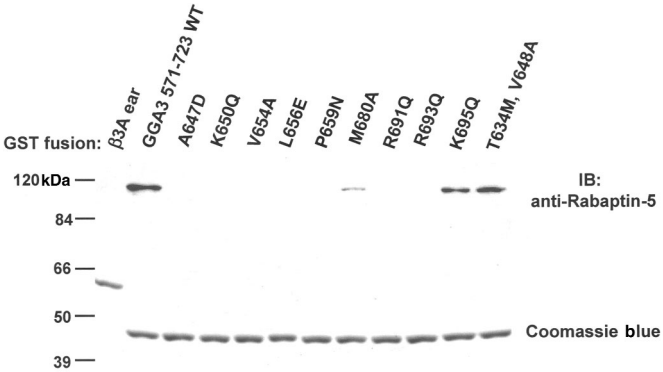

Figure 2 Binding of GGA3-GAE mutants to Rabaptin-5. The indicated GST fusion proteins were immobilized on glutathione-Sepharose and incubated with bovine brain cytosol (GST-β3A ear was used as negative control). Bound proteins were eluted and subjected to SDS-PAGE followed by immunoblotting (IB) with antibody to Rabaptin-5. Coomassie blue staining of GST fusion constructs is shown. Positions of molecular-mass markers are indicated on left.

domain[18], the N-terminal extension does not contain an α-helix. The other major differences between the structures of GAE domains in GGA3-GAE and γ1-adaptin are in the loops between strands β2 and β3, β5 and β6, and β7 and β8, and in the C-terminal portion of the protein from the end of β8 to the C terminus (**Fig. 4c**).

The Rabaptin-5 hydrophobic peptide binds to a surface formed by strands β4, β5 and β7 (**Fig. 4a**). The peptide is anchored by the burial of a phenylalanine residue that we have designated as position 0 (see **Supplementary Table 1** online). Of the 13 residues in the peptide, 8 are ordered and identifiable in electron density (the sequence DFG-PLVGA, numbered relative to the phenylalanine from –1 to +6). The peptide is in a β-strand conformation except for a kink at the Gly(+1)-Pro(+2) sequence. Given the high quality of the 2.2-Å electron density and the low average *B*-factor for the structure, we could characterize the water structure around the peptide in detail. Three β-type main chain–main chain interactions between the peptide and the domain, involving Pro(+2), Gly(+5) and Ala(+6), are mediated by water molecules. One interaction between the peptide main chain and the domain side chain is mediated by a chain of two water molecules between the carbonyl of Gly(+1) and the side chain of Arg693 (**Fig. 5**; (**Supplementary Table 1** online).

The N-terminal part of the peptide, comprising Asp(–1) and Phe(0), contacts β4 and β7 of the three-stranded sheet. The side chain of Asp(–1) makes a solvent-exposed salt bridge with the side chain of Lys650. The aromatic ring of Phe(0) is deeply wedged between the aliphatic portions of Arg691 and Arg693, Ala647 and the ring of Pro649 (**Fig. 5**). The middle of the peptide, consisting of Gly(+1) and Pro(+2), has limited interactions with the domain. Aside from water-mediated β-sheet pairing interactions, the only direct interaction is between the main chain NH of Gly(+1) and the carbonyl of Ala648. The C-terminal part, comprising Leu(+3), Val(+4) and Gly(+5), interacts with a surface formed at the edges of β4 and β5. The side chain of Leu(+3) is half-buried in a shallow, wide hydrophobic depression formed by the aliphatic parts of Gln645 and Lys695, together with Ala647 (**Fig. 5**). Val708 forms a back wall to this pocket but is not in direct contact with Leu(+3). The side chain of Val(+4) interacts with two adjacent side chains of β5, Val654 and the hydrophobic part of Lys655. The last side chain interaction in the peptide is between the methyl group of Ala(+6) and the rings of Pro658 and Pro659 (**Fig. 5**).

The Rabaptin-5 binding site is on the opposite side of the β-sandwich from the site where DPW-containing peptides bind to the β-sandwich portion of the α-appendage domain[13]. The counterpart of the DPW binding site on the GGA3-GAE domain is partly covered by the N-terminal extension. The Rabaptin-5 binding site adjoins one of the two residues mutated in the construct used for crystallization, V648A. Val648 is a core-packing residue that presumably stabilizes the conformation of β4 with respect to the rest of the GAE domain in the wild-type protein. The main chain of residue 648 directly interacts with the peptide backbone; it is also an immediate neighbor of Ala647, a peptide-binding residue. A combination of subtle effects on the conformation of several peptide-binding residues (644–650) in β4 probably accounts for the six-fold reduction in affinity observed for this mutant protein (**Fig. 1d**).

Peptide recognition mechanism

Consistent with the defining role of phenylalanine as the most conserved residue in reported GAE-binding sequences, the side chain of Phe(0) in the peptide buries the largest amount of solvent-accessible surface area. The recognition of the Phe(0) aromatic ring by a predominantly basic side chain cluster is the most notable aspect of this hydrophobic motif readout by the GGA3-GAE domain. Arg691 and Arg693 are absolutely conserved as basic residues (arginine or lysine) in all GAE domains (**Fig. 6**). The presence of this conserved basic cluster on the surface of the GAE domain of γ1-adaptin was noted by Wakatsuki and co-workers[21]. The obvious inference seemed to be that the basic cluster would be responsible for recognizing an acidic motif, but we have now found that the most critical contacts in the motif are hydrophobic. The interaction seems to be hydrophobic in nature, and favorable because the flatness of both the aromatic ring of the phenylalanine and the guanidino group of the arginine makes them complementary in shape. The interaction is reminiscent of the interaction of Arg905 of the platform portion of the α-appendage domain with the phenylalanine or tryptophan residues of DPF or DPW peptides[13]. This unusual role for a basic cluster in hydrophobic ligand binding is corroborated by mutational analysis, which shows that the replacement of Arg691 or Arg693 by glutamine abolishes Rabaptin-5 binding as judged by ITC (**Fig. 1d**) and GST pull-down assays (**Fig. 2**). The aliphatic moieties of the glutamine residues are smaller than those of arginine and would not be predicted to form equivalent hydrophobic interactions with the phenylalanine. Mutation to aspartic acid of

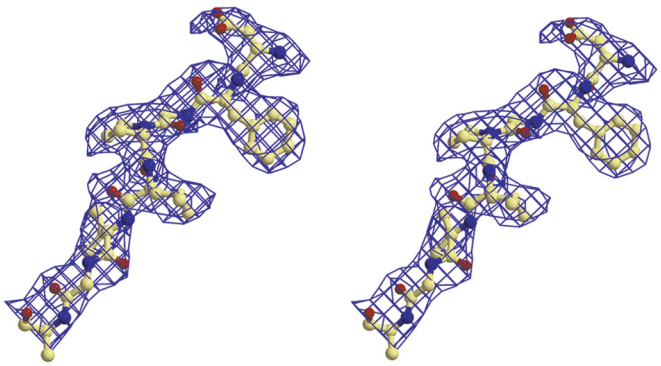

Figure 3 Structure determination of the GAE domain–peptide complex. Stereo view of electron density from the solvent-flattened MAD data set at 2.8-Å resolution and contoured at 1.0 σ is shown for the bound peptide. Electron-density map was drawn with Spock (http://mackerel.tamu.edu/spock), MolScript[31] and Raster3D[32].

© 2003 Nature Publishing Group http://www.nature.com/naturestructuralbiology

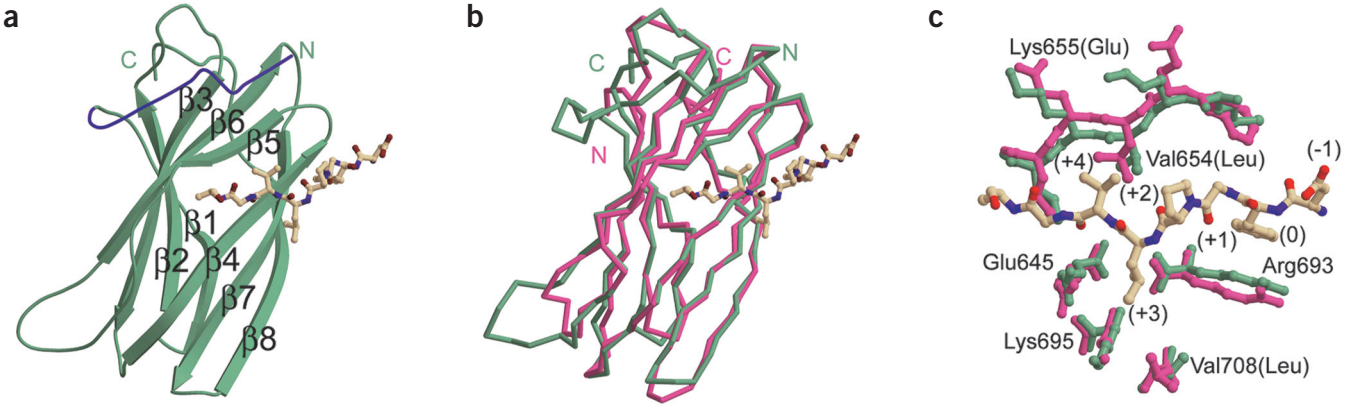

Figure 4 Overall structure of the GGA3-GAE domains from GGA3 and comparison to the γ1-adaptin ear. (**a**) Overall structure of the GGA3-GAE domain in ribbon representation, drawn with Spock and Molscript and rendered with Raster3D[32] (N-terminal extension of GGA3-GAE domain, blue). (**b**) Superposition of GGA3 (green) and γ-adaptin (pink) Cα backbones[20,21]. N and C termini are labeled in same color as respective Cα backbone traces. (**c**) Close-up view of superposed structures, showing differences in the binding sites (numbering of residues corresponds to the GGA3 sequence; substitutions in the γ1-adaptin GAE domain are shown in parentheses).

© 2003 Nature Publishing Group http://www.nature.com/naturestructuralbiology

Ala647, which also interacts hydrophobically with Phe(0), similarly abolishes binding (**Fig. 2**). Consistent with the presence of aspartic acid residues preceding the phenylalanine in the known binding sequences, Asp(–1) makes a salt bridge with the absolutely conserved Lys650. Mutation of this lysine to glutamine also abolishes Rabaptin-5 binding (**Fig. 2**), supporting an important role for the ε-amino group of this lysine in peptide recognition. This salt bridge is exposed to solvent, suggesting that additional electrostatic interactions could have a role in binding. Apart from short-range salt bridge interactions, longer-range electrostatic interactions between acidic residues in the N-terminal part of the motif and the GAE basic cluster could contribute further to binding or stabilization.

Leu(+3) and Val(+4) bind to relatively open and exposed hydrophobic surfaces. These surfaces had previously been predicted to bind accessory proteins by Owen, Wakatsuki and co-workers[20,21]. Modeling of the second phenylalanine in the γ-synergin peptide sequence (+3 position; **Fig. 1a**) indicates that it could bind in the same pocket as Leu(+3) of the Rabaptin-5 peptide. The Leu(+3) binding site is also made up largely of the aliphatic portions of hydrophilic residues (Lys695, conserved as lysine or arginine in most other GAE domains, and Gln645, which is conserved in most but not all GAE sequences, **Fig. 6**), although this hydrophilic cluster is separate from the basic cluster that binds to Phe(0). The back wall of the +3 site in the GAE domain of γ1-adaptin is less open than that of GGA3, owing to a V708L substitution in GGA3 (**Fig. 6**). The interaction with the leucine might provide more energy for the binding of a large hydrophobic side chain in position +3 to γ1-adaptin. In other GAE domains, this position is not well conserved, and the differences here could fine-tune the affinity for sequences with different residues at the +3 position. Val(+4) binds in a hydrophobic pocket composed of Val654 and the hydrophobic part of Lys655. The importance of this residue in the Rabaptin-5 binding sequence depends on the specific GAE domain being considered (it is more crucial for γ2-adaptin, GGA1 and GGA2 and less so for γ1-adaptin and GGA3)[16]. The +4 site is more open in the γ1-adaptin structure because a V654L substitution in GGA3 causes the entire middle part of β5 to be displaced from β4, resulting in a wider gap of ~1.4 Å between the two. The corresponding residue is always a valine in human GGAs, and is a leucine in human γ-adaptins and yeast GAE domains (**Fig. 6**). The role of the +4 site may be one of the more substantial differences in recognition between the GAE domains of the GGA3 and γ-adaptins.

The +6 position was not predicted to form part of the Rabaptin-5, enthoprotin or γ-synergin motifs, yet the α-methyl group of Ala(+6) in the Rabaptin-5 peptide buries substantial surface area against two proline residues at the end of β5. The importance of this interaction is underscored by the loss of binding observed with a P659N mutation (**Fig. 2**). There is little steric hindrance in this position, and this region is less conserved in the GAE family than other parts of the peptide binding site (Pro658 and Pro659 are present in about half of known GAE sequences; **Fig. 6**). The structure suggests that recognition requires two residues flanking the motif on the C-terminal side, but that the identity of the residues is not necessarily critical because main chain interactions and the Cβ atom of the second side chain are sequence-independent contributors to binding. The +6 position in the γ-synergin and enthoprotin peptides is occupied by a phenylalanine; the absence of this residue in the short enthoprotin and γ-synergin peptides used in our competition analysis (**Fig. 1a**) could explain their lower avidity for the GGA3-GAE domain (**Fig. 1c**).

The role of the glycine at position +1 in the Rabaptin-5 motif is intriguing because it does not directly interact with the GAE domain. The Gly(+1) in the peptide has conformational torsion angles of $(\phi,\psi) = -88°, -172°$. This places the glycine at the edge of the allowed region of the Ramachandran plot for a nonglycine residue, although it is within the core allowed region for a glycine residue. The conformation of the glycine facilitates an unusual register shift in the β-sheet pairing of the peptide with GAE domain strand β4. In a normal β-sheet, the side chains of all even-numbered residues would point to one side, and those of the odd-numbered residues would point to the other. In the bound Rabaptin-5 peptide, the side chains of Phe(0) and Leu(+3) are on the same side of the strand. The presence of the glycine makes it possible for the strand to kink such that the register of the strand is switched between positions 0 and +3. Certain combinations of nonglycine, nonproline residues may allow the strand to flex enough to accommodate the same register shift. However, the glycine is required in the context of the Rabaptin-5 sequence, in which the +2 position is occupied by the most conformationally rigid residue, proline.

DISCUSSION

Implications for peptide specificity

On the basis of the structural data presented here and the comparison of the GAE-binding sequences from Rabaptin-5, enthoprotin and

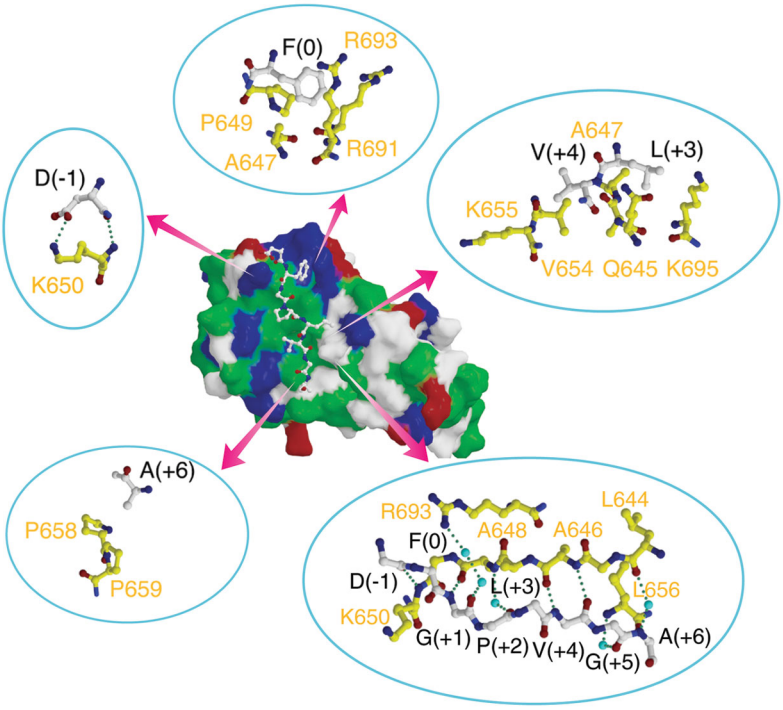

Figure 5 Molecular interactions between GGA3-GAE domain and Rabaptin-5 peptide (blue, basic residues; red, acidic; green, hydrophobic; white, remainder of amino acids).

γ-synergin (**Fig. 1a**), we propose that the probable consensus motif for peptide binding to GAE domains is DFGXΦ. Mutation of the F and Φ residues abrogates binding of the three peptides to GGA3-GAE (**Fig. 1b**). In addition, the three peptides compete for binding of the Rabaptin-5 peptide to the GGA3-GAE domain (**Fig. 1b,c**). Finally, mutations of residues in the peptide-binding site of the GGA3-GAE domain (**Fig. 2**) and of the analogous residues in the γ1-adaptin GAE domain[17,20,21] have largely similar effects on the binding of Rabaptin-5, γ-synergin and enthoprotin. Taken together, these observations indicate that the three accessory protein motifs interact with the same binding site defined in this study, and that this binding site is likely to be similar in the GGA and γ-adaptin GAE domains.

Differences in affinity between peptides from Rabaptin-5, enthoprotin and γ-synergin for binding to GGA3-GAE are nonetheless apparent (**Fig. 1c**). In addition, various accessory proteins bind with distinct avidities to the GAE domains of the GGAs and the γ-adaptins[5], and at least one mutation, substitution by alanine of the first aspartic acid in the DFGDW sequence in enthoprotin, has differential effects on the binding of this protein to GGA1 and γ1-adaptin[17]. These findings suggest that the identity of residues other than the anchoring phenylalanine confer fine specificity to the interactions of different accessory proteins with GAE domains.

Acidic residues may have a more prominent role in the enthoprotin and γ-synergin motifs than in the Rabaptin-5 motif. The structure of the peptide–GAE complex shows that ancillary contributions are made by an aspartic acid preceding the Phe(0). Substitution of acidic residues preceding the FGPLV sequence from Rabaptin-5 has shown that, although not critical for the binding of large Rabaptin-5 fragments, these amino acids contribute to the binding of smaller peptides containing the FGPLV sequence[16]. The +2 position is also occupied by aspartic acid and glutamic acid in the enthoprotin and γ-synergin sequences, respectively (**Fig. 1a**). A modeled Asp(+2) has its side chain near Lys650 of GGA3, which is conserved in γ1-adaptin. The interactions between the acidic peptide residues and basic GAE domain residues are solvent-exposed and therefore weak, but may be cumulatively stronger. The requirements at each position in the motif may depend to some extent on the surrounding context. For instance, because the role of Gly(+1) in the Rabaptin-5 sequence is to contribute flexibility, it may be more critical in the context of a proline at +2 than it is when a more flexible residue is present at this site. Therefore, there could be additional binding sequences in which the glycine residue is not present.

The γ1-adaptin GAE has a more tightly defined +3 site and a more open +4 site, suggesting that recognition differences could center on these positions. Consistent with this, phenylalanine and tryptophan are found at +3 in γ-synergin and enthoprotin, respectively. The well-defined wall of the large +3 pocket in the γ-adaptin GAE would accommodate a large residue and provide multiple interactions. Large side chains would probably fit into the +3 pocket of the GGA3-GAE because the γ-synergin and enthoprotin peptides, with their phenylalanine and tryptophan residues, also bind to this domain. The favorable hydrophobic contacts made between these large side chains and the valine-containing +3 site of GGA3-GAE are likely to be less extensive than the corresponding interactions with the leucine-containing +3 site of γ1-adaptin.

Comparison to the GGA1-GAE–p56 complex

After this article was accepted for publication, we compared our study with that of an accompanying article in this issue by Collins *et al.*[23], who independently solved the structure of the complex of the GAE domain of GGA1 with the peptide DDDFGGFEAAETFD from p56. The p56 peptide binds to the same site on GGA1 as that recognized by the Rabaptin-5 peptide on the GGA3-GAE domain. The backbone conformations of the two bound peptides are essentially identical. The interactions of the anchoring Phe(0) with the conserved arginine side chains, which form the linchpin of the binding site, are also identical. The close congruence of the two structures demonstrates the generality of the GAE-peptide binding model proposed here and in the accompanying paper[23].

Despite the close similarity in the two complex structures, the p56 peptide binds to the GGA1 GAE domain with an affinity ~14-fold lower than that of Rabaptin-5 peptide binding to the GGA3-GAE domain. The differences are not likely to be due to differences in surfaces of the GGA1 and GGA3 GAE domains, as these are nearly identically conserved between the GGA isoforms. As Collins *et al.*[23] have discussed, the Gly(+2) of the p56 peptide seems to destabilize binding through an entropic effect. The p56 peptide has a glutamic acid at +4 instead of a hydrophobic residue, and thus fails to make the favorable contacts observed at the +4 position in the Rabaptin-5 peptide. The combination of these two differences probably accounts for most of the difference in affinity.

© 2003 Nature Publishing Group http://www.nature.com/naturestructuralbiology

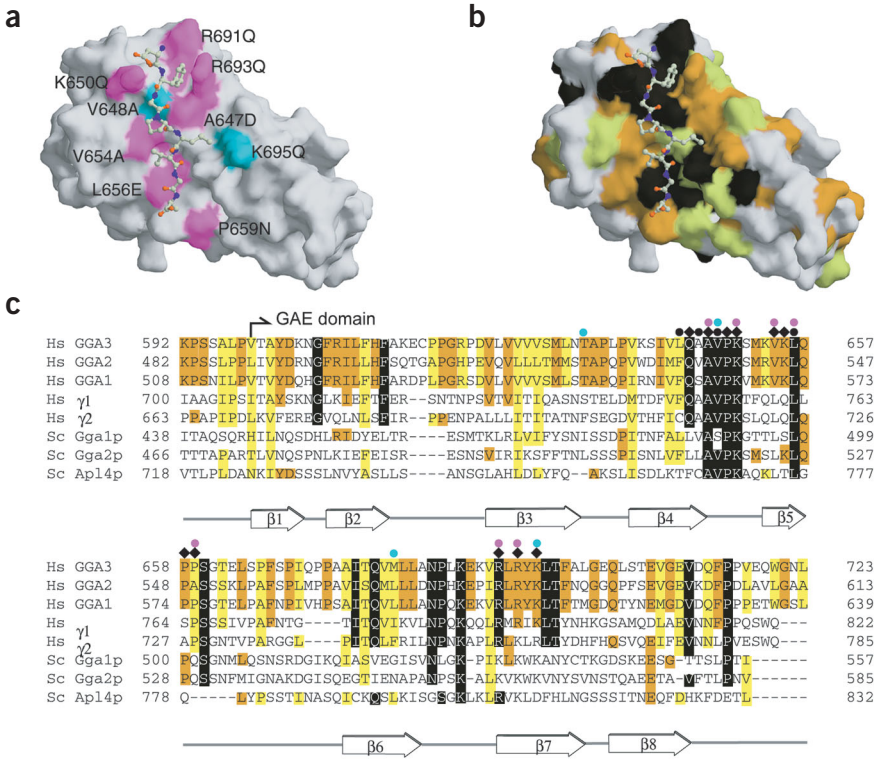

Figure 6 Conservation of GAE domains. (**a**) Mutational analysis of the GGA3-GAE peptide-binding site. Residues that abolish binding when mutated, purple; residues that either reduce or have no effect on binding when mutated, light blue. (**b**) Conservation of solvent-exposed residues of the GAE domain. Residues identical in all three human GGAs and two human γ-adaptins, black; residues identical in all three human GGAs but not in other GAE domains, orange; residues conserved but not identical in the human GGAs, yellow; other residues, white. (**c**) Annotated sequence alignment of GAE domains. Residues are boxed using the same color scheme as in **b**. Purple and light-blue circles above the alignment highlight residues mutated in the study, with colors as in **a**. Black diamonds, residues whose side chains interact with the peptide; black circles, residues whose main chains, but not side chains, interact with the peptide.

Concluding remarks

Our findings, together with previous observations and the results in the accompanying paper[23], demonstrate that the GGA and γ-adaptin and the α- and β2-adaptin ear domains recognize accessory proteins differently. The binding motifs on the accessory proteins, the binding sites on the ear domains and the modes of recognition of each are all different; this accounts for the distinct sets of accessory proteins bound by GGAs and AP-1 as compared with AP-2. The presence of GAE domains in the three human GGAs and the two γ-adaptin subunit isoforms of AP-1, and their binding to the same set of accessory proteins—notwithstanding differences in affinity—suggest that these proteins are regulated similarly and may therefore participate in common cellular processes.

METHODS

DNA constructs. A *Bam*HI-*Eco*RI fragment encoding human GGA3(571–723) (comprising the C-terminal 28 residues of the hinge region and the entire GAE domain) was generated by PCR amplification and subsequently subcloned into the corresponding sites of the pGST-parallel2 vector[24]. The cloned PCR fragment encoding human GGA3(571–723) contained two mutations (resulting in T634M and V648A substitutions) when compared with the previously reported sequence (GenBank entry AF219138). These two mutations were reverted using the QuikChange site-directed mutagenesis kit (Stratagene) and the pGST-parallel2 construct as template. Additional mutations in the wild-type GGA3(571–723) were introduced using the corresponding pGST-parallel2 construct as a template and the same mutagenesis system. The GST-human β3A-ear(799–1081) construct was described previously[25].

Peptides. Peptide DESDFGPLVGADS (human Rabaptin-5(435–447)) and additional peptides designed with an extra N-terminal cysteine for biotinylation purposes (CDESDFGPLVGADS, human Rabaptin-5(435–447); CDESDAGPAVGADS, termed Rabaptin-5(435–447) (AA); CLADDFGEFSL, human γ-synergin(666–675); CLADDAGEASL, termed human γ-synergin(666–675) (AA); CGNGDFGDWSA, human enthoprotin(368–377); CGNGDAGDASA, termed human enthoprotin(368–377) (AA) were obtained from New England Peptide.

Binding to biotinylated peptides. Peptide biotinylation, immobilization of biotinylated peptides (5 nmol peptide and 50 μl streptavidin-agarose in 1 ml PBS, pH 7.0), and washing of beads with immobilized peptides were carried out as described[16]. Aliquots of GST β3A-ear or GST-GGA3(571–723) (3 μg) were preincubated for 45 min at 4 °C in the presence of the indicated concentrations of free peptide (**Fig. 1b,c**) in a final volume of 500 μl of 15 mM HEPES, pH 7.0, 75 mM NaCl, 0.25% (v/v) Triton X-100 supplemented with 0.1% (w/v) BSA and protease inhibitors (EDTA-free Complete peptide binding buffer, Roche). At the end of this period, the mixtures were added to 50 μl of washed beads containing immobilized, biotinylated Rabaptin-5(435–447) peptide and incubated for an additional 90 min at 4 °C. Incubation was stopped by centrifugation for 2 min at 2,000*g* and 4 °C, and the beads washed twice by resuspension with 1 ml of binding buffer without BSA and centrifugation. Bead-bound proteins were eluted by incubation in 50 μl of 4× Laemmli buffer and subsequently subjected to SDS-PAGE and immunoblotting with a rabbit antiserum to GST[25].

GST pull-down assays. The immobilization of samples containing 15 μg of GST-human β3A-ear or GST-human GGA3(571–723) constructs (wild type or mutants) and the pull-down of interacting proteins from bovine brain cytosol were carried out as reported previously[16]. Bound proteins were subjected to SDS-PAGE and immunoblotting with mouse monoclonal antibody to Rabaptin-5 (Transduction Laboratories).

Isothermal titration calorimetry. GGA3 constructs were dialyzed into 20 mM MES buffer, pH 6.2, 100 mM NaCl and were used at concentrations of 50 μM and 100 μM. Rabaptin-5 peptide was dissolved in the same buffer and was used at either 1 mM or 2 mM. Before the binding experiments were done, TCEP was added to both GAE domain and peptide to a final concentration of 1 mM. Titration calorimetery measurements were taken using a MicroCal VP-ITC instrument at 30 °C. Rabaptin-5 peptide was injected into 1.4 ml GAE domain in 28 aliquots of 10 μl each at 180-s intervals. Data obtained from peptide injections into 1.4 ml buffer blanks were subtracted from the experimental data before analysis using Origin (Microcal). Affinities reported are the mean ± s.d. determined from three independent measurements.

Protein expression and crystallization. The native and SeMet GGA3-GAE domains were expressed as GST-tagged fusion proteins, the native GAE domain in *Escherichia coli* strain Rosetta (DE3) (Novagen) and the SeMet derivative in BL834 (Novagen), and all were purified by affinity chromatography using glutathione-Sepharose resin (Amersham Pharmacia). The column was washed with 10 column volumes of 1× PBS and 0.001% (v/v) Triton

© 2003 Nature Publishing Group http://www.nature.com/naturestructuralbiology

Table 1 Crystallographic data, phasing and refinement statistics

	SeMet	Native
Data collection and phasing		
Space group	$C222_1$	$C222_1$
Unit cell dimensions		
a (Å)	46.6	45.2
b (Å)	89.8	85.2
c (Å)	91.7	97.0
$\alpha = \beta = \gamma$ (°)	90	90
Wavelength (Å)	0.9791	1.54
Resolution (Å)[a]	50–2.8 (2.98–2.8)	50–2.2 (2.28–2.20)
Unique reflections	4,812 (459)	9,417
Completeness (%)	99.1 (99.4)	97.8 (98.5)
R_{merge} (%)[b]	0.057 (0.119)	0.044 (0.109)
Anomalous diff. (%)	0.041–0.052	
Dispersive diff. (%)	0.034–0.051	
F.O.M.[c]	0.57	
Refinement		
Resolution range (Å)	50.0–2.20	
Number of reflections	9,417 (833)	
R (%)[d]	0.2160 (0.239)	
R_{free} (%)[e]	0.2497 (0.322)	
Cross-validated Luzatti error	0.31	
R.m.s. deviations		
Bond length (Å)	0.00652	
Bond angle (°)	1.42	
Mean *B*-factor (Å^2) protein	22.9	
Mean *B*-factor (Å^2) peptide	25.8	
Number of protein atoms	1,165	
Number of solvent atoms	209	
Residues in core ϕ-ψ region	100%	

[a]Data in parentheses are for highest-resolution shells. [b]$R_{sym} = \Sigma_h\Sigma_i|I_i(h) - \langle I(h)\rangle| / \Sigma_h\Sigma_i I_i(h)$ where I is the observed intensity and $\langle I\rangle$ is the average intensity of multiple symmetry-related observations of that reflection. [c]Figure of merit of the MAD solution in SOLVE[27]. [d]$R = \Sigma(|F_{obs}| - k|F_{calc}|) / \Sigma|F_{obs}|$. [e]$R_{free}$ is the R value calculated for a test set of reflections, comprising a randomly selected 10% of the data not used during refinement.

X-100 and eluted using 20 mM glutathione. After dialysis, the GST tags were cleaved with recombinant TEV protease and the proteins were subjected to a second purification using glutathione-Sepharose. Following concentration, the proteins were further purified using a Superdex 200 (16/60) gel filtration column (Amersham-Pharmacia) equilibrated with 20 mM Tris-HCl, 150 mM NaCl, 10 mM DTT. The purified domain contains the N-terminal residues GAMGS followed by the human GGA3 residues 571–723. The proteins were dialyzed into 10 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM DTT and were concentrated to 35 mg ml^{-1} (native) and 28 mg ml^{-1} (SeMet) for use in crystallization trials. Both native and SeMet GAE domains crystallized in the presence of a two-fold molar concentration of peptide in 2 µl hanging drops over a 0.5 ml reservoir of mother liquor. The native GAE domain crystallized in 1.6 M sodium citrate, pH 6.5, whereas the SeMet GAE domain crystallized in 0.1 M sodium acetate, 30% (v/v) PEG 8000, 0.1 M sodium cacodylate, pH 6.0. The presence of peptide was confirmed in the native GAE crystals using mass spectrometry of crystals washed in mother liquor and dissolved in 0.1 M acetic acid.

X-ray data collection and structure determination. SeMet GAE crystals were cryoprotected in mother liquor supplemented with 5% (v/v) PEG 400, and native crystals were cryoprotected in mother liquor alone.

A three-wavelength multiwavelength anomalous dispersion data set was collected from frozen crystals of SeMet GGA3-GAE at beamline 8.2.1, Advanced Light Source (Berkeley, California). Data were collected in 1.5° increments for 10 s per exposure. A total of 90° direct and 90° inverse beam were collected in 15° wedges for each wavelength. Data were collected on an ADSC 4-panel CCD and processed with HKL2000 (ref. 26) keeping the processed intensities unmerged. Selenium positions were located with SOLVE[27] and the map improved with RESOLVE[28]. The resulting 2.8-Å map was interpreted with O[29], and the assignment of sequence initiated using the identified SeMet positions, together with the known structures of the homologous γ-adaptin GAE domains. The SeMet structure was partially refined with CNS[30]. Native data were collected on a RU-200 generator equipped with Osmic optics and an R-Axis IV++ detector, and processed to 2.2 Å with DENZO and SCALEPACK[26]. Native and SeMet crystals were nonisomorphous. After preliminary refinement of the SeMet structure, the model was placed in the unit cell of the native protein, and the structure was refined at a resolution of 2.2 Å with torsional dynamics against the maximum-likelihood-target function using CNS.

Coordinates. Coordinates have been deposited in the Protein Data Bank (accession code 1P4U).

Note: Supplementary information is available on the Nature Structural Biology website.

ACKNOWLEDGMENTS
We thank T. Earnest and the staff of the HHMI beamline 8.2.1, Advanced Light Source, Lawrence Berkeley Lab (Berkeley, California) for assistance with data collection, S. Sechi for assistance with mass spectrometry, and X. Zhu for expert technical assistance.

COMPETING INTERESTS STATEMENT
The authors declare that they have no competing financial interests.

Submitted 25 February; accepted 11 June 2003
Published online 13 July 2003; doi:10.1038/nsb953

1. Kirchhausen, T. Adaptors for clathrin-mediated traffic. *Annu. Rev. Cell Dev. Biol.* **15**, 705–732 (1999).
2. Robinson, M.S. & Bonifacino, J.S. Adaptor-related proteins. *Curr. Opin. Cell Biol.* **13**, 444–453 (2001).
3. Boman, A.L., Zhang, C.J., Zhu, X. & Kahn, R.A. A family of ADP-ribosylation factor effectors that can alter membrane transport through the trans-Golgi. *Mol. Biol. Cell* **11**, 1241–1255 (2000).
4. Dell'Angelica, E.C. *et al.* GGAs: a family of ADP ribosylation factor-binding proteins related to adaptors and associated with the Golgi complex. *J. Cell Biol.* **149**, 81–94 (2000).
5. Hirst, J. *et al.* A family of proteins with γ-adaptin and VHS domains that facilitate trafficking between the trans-Golgi network and the vacuole/lysosome. *J. Cell Biol.* **149**, 67–80 (2000).
6. Poussu, A., Lohi, O. & Lehto, V.P. Vear, a novel Golgi-associated protein with VHS and γ-adaptin 'ear' domains. *J. Biol. Chem.* **275**, 7176–7183 (2000).
7. Nielsen, M.S. *et al.* The sortilin cytoplasmic tail conveys Golgi-endosome transport and binds the VHS domain of the GGA2 sorting protein. *EMBO J.* **20**, 2180–2190 (2001).
8. Puertollano, R., Aguilar, R.C., Gorshkova, I., Crouch, R.J. & Bonifacino, J.S. Sorting of mannose 6-phosphate receptors mediated by the GGAs. *Science* **292**, 1712–1716 (2001).
9. Zhu, Y., Doray, B., Poussu, A., Lehto, V.P. & Kornfeld, S. Binding of GGA2 to the lysosomal enzyme sorting motif of the mannose 6-phosphate receptor. *Science* **292**, 1716–1718 (2001).
10. Takatsu, H., Katoh, Y., Shiba, Y. & Nakayama, K. Golgi-localizing, γ-adaptin ear homology domain, ADP-ribosylation factor-binding (GGA) proteins interact with acidic dileucine sequences within the cytoplasmic domains of sorting receptors through their Vps27p/Hrs/STAM (VHS) domains. *J. Biol. Chem.* **276**, 28541–28545 (2001).
11. Slepnev, V.I. & De Camilli, P. Accessory proteins in clathrin-dependent synaptic vesicle endocytosis. *Nat. Rev. Neurosci.* **1**, 161–172 (2000).
12. Benmerah, A., Begue, B., Dautry-Varsat, A. & Cerf-Bensussan, N. The ear of α-adaptin interacts with the COOH-terminal domain of the Eps 15 protein. *J. Biol. Chem.* **271**, 12111–12116 (1996).
13. Brett, T.J., Traub, L.M. & Fremont, D.H. Accessory protein recruitment motifs in clathrin-mediated endocytosis. *Structure* **10**, 797–809 (2002).
14. Owen, D.J. *et al.* A structural explanation for the binding of multiple ligands by the α-adaptin appendage domain. *Cell* **97**, 805–815 (1999).
15. Owen, D.J., Vallis, Y., Pearse, B.M., McMahon, H.T. & Evans, P.R. The structure and function of the β2-adaptin appendage domain. *EMBO J.* **19**, 4216–4227 (2000).

© 2003 Nature Publishing Group http://www.nature.com/naturestructuralbiology

16. Mattera, R., Arighi, C.N., Lodge, R., Zerial, M. & Bonifacino, J.S. Divalent interaction of the GGAs with the Rabaptin-5-Rabex-5 complex. *EMBO J.* **22**, 78–88 (2003).
17. Mills, I.G. *et al.* EpsinR: an AP1/clathrin interacting protein involved in vesicle trafficking. *J. Cell Biol.* **160**, 213–222 (2003).
18. Liu, W.W.Y. *et al.* Binding partners for the COOH-terminal appendage domains of the GGAs and γ-adaptin. *Mol. Biol. Cell* **14**, 2385–2398 (2003).
19. Duncan, M.C., Costaguta, G. & Payne, G.S. Yeast epsin-related proteins required for Golgi-endosome traffic define a γ-adaptin ear-binding motif. *Nat. Cell Biol.* **5**, 77–81 (2003).
20. Kent, H.M., McMahon, H.T., Evans, P.R., Benmerah, A. & Owen, D.J. γ-adaptin appendage domain: structure and binding site for Eps15 and γ-synergin. *Structure* **10**, 1139–1148 (2002).
21. Nogi, T. *et al.* Structural basis for the accessory protein recruitment by the γ-adaptin ear domain. *Nat. Struct. Biol.* **9**, 527–531 (2002).
22. Page, L.J., Sowerby, P.J., Lui, W.W. & Robinson, M.S. γ-synergin: an EH domain-containing protein that interacts with γ-adaptin. *J. Cell Biol.* **146**, 993–1004 (1999).
23. Collins, B.M., Praefcke, G.J.K., Robinson, M.S. & Owen, D.J. Structural basis for binding of accessory proteins by the appendage domain of GGAs. *Nat. Struct. Biol.* advance online publication, 13 July 2003 (doi:10.1038/nsb955).
24. Sheffield, P., Garrard, S. & Derewenda, Z. Overcoming expression and purification problems of RhoGDI using a family of 'parallel' expression vectors. *Protein Expr. Purif.* **15**, 34–39 (1999).
25. Dell'Angelica, E.C., Klumperman, J., Stoorvogel, W. & Bonifacino, J.S. Association of the AP-3 adaptor complex with clathrin. *Science* **280**, 431–434 (1998).
26. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
27. Terwilliger, T.C. & Berendzen, J. Correlated phasing of multiple isomorphous replacement data. *Acta Crystallogr. D* **52**, 749–757 (1996).
28. Terwilliger, T.C. Maximum-likelihood density modification. *Acta Crystallogr. D* **56**, 965–972 (2000).
29. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
30. Brunger, A.T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
31. Kraulis, P.J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
32. Merritt, E.A. & Bacon, D.J. Raster3D photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524 (1997).

© 2003 Nature Publishing Group http://www.nature.com/naturestructuralbiology