THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 279, No. 9, Issue of February 27, pp. 8018–8028, 2004
Printed in U.S.A.

Definition of the Consensus Motif Recognized by γ-Adaptin Ear Domains*

Received for publication, October 29, 2003, and in revised form, December 2, 2003
Published, JBC Papers in Press, December 9, 2003, DOI 10.1074/jbc.M311873200

Rafael Mattera‡, Brigitte Ritter§¶, Sachdev S. Sidhu‖, Peter S. McPherson, and Juan S. Bonifacino‡ ‡‡**

From the ‡Cell Biology and Metabolism Branch, NICHD, National Institutes of Health, Bethesda, Maryland 20892, the §Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, Quebec H3A 2B4, Canada, and the ‖Department of Protein Engineering, Genentech Inc., South San Francisco, California 94080

The heterotetrameric adaptor complex 1 (AP-1) and the monomeric Golgi-localized, γ ear-containing, Arf-binding (GGA) proteins are components of clathrin coats associated with the *trans*-Golgi network and endosomes. The carboxyl-terminal ear domains (or γ-adaptin ear (GAE) domains) of two γ-adaptin subunit isoforms of AP-1 and of the GGAs are structurally similar and bind to a common set of accessory proteins. In this study, we have systematically defined a core tetrapeptide motif ΨG(P/D/E)(Ψ/L/M) (where Ψ is an aromatic residue), which is responsible for the interactions of accessory proteins with GAE domains. The definition of this motif has allowed us to identify novel GAE-binding partners named NECAP and aftiphilin, which also contain clathrin-binding motifs. These findings shed light on the mechanism of accessory protein recruitment to *trans*-Golgi network and endosomal clathrin coats.

Clathrin-coated vesicles sort biosynthetic and endocytic cargo within the interconnected system of membrane-bound organelles formed by the *trans*-Golgi network, endosomes, lysosomes, and plasma membrane (1, 2). The coats on these vesicles contain, in addition to clathrin, an array of adaptor proteins involved in vesicle formation and cargo selection (2, 3). Among these are the heterotetrameric adaptor protein (AP)[1] complexes AP-1, AP-2, and AP-3. These complexes are composed of two large (90–130-kDa), one medium (~50-kDa), and one small (~20-kDa) "adaptin" subunit, namely γ·β1·μ1·σ1 for AP-1, α·β2·μ2·σ2 for AP-2, and δ·β3·μ3·σ3 for AP-3 (subunits listed in order of decreasing molecular mass). Each of the large subunits (γ/α/σ and β1–β3, respectively) is organized into three domains: an amino-terminal "trunk," a connecting "hinge," and a carboxyl-terminal "ear." The trunk domains of the large subunits together with the medium (μ1–μ3) and small (σ1–σ3) subunits constitute the "core" of the AP complexes, from which the two hinge and ear domains project outward (4). Recent studies have identified a family of monomeric proteins named GGAs (*i.e.* GGA1, GGA2, and GGA3 in humans) that also function as clathrin adaptors (4). The amino-terminal half of the GGAs is unrelated to the AP subunits, but the carboxyl-terminal half of the GGAs is structured as hinge and ear domains analogous to those of the large AP subunits. Moreover, the earlike domain of the GGAs exhibits significant sequence similarity to that of the γ-adaptin subunit of AP-1 (of which there are two isoforms termed γ1- and γ2-adaptin) (5, 6). This conserved domain, which is present in the three GGAs and the two γ-adaptins, is herein referred to as the γ-adaptin ear (GAE) domain.

The ear domains of the AP large subunits function as platforms for the recruitment of accessory proteins via recognition of specific sequence motifs. The AP-2 α and β2 ear domains are the most extensively characterized biochemically and structurally. They bind a large number of accessory proteins (including Eps15, epsin 1, amphiphysin I, AP180, auxilin, and dynamin 1 and 2) that participate in various aspects of clathrin-mediated endocytosis (7). These accessory proteins contain DPF, DPW, or F*X*D*X*F motifs that are responsible for interactions with the α and β2 ear domains (8, 9). The α ear comprises two subdomains: an amino-terminal immunoglobulin-like β-sandwich that binds the DPW sequence and a carboxyl-terminal platform that binds DPF, DPW, and F*X*D*X*F motifs (9–11). The β2 ear domain has a similar structure but only binds DPF and DPW sequences on its carboxyl-terminal platform subdomain (9, 12). Recent crystallographic analyses revealed that the GAE domain of the γ-adaptins and the GGAs consists of a β-sandwich similar to that of the α and β2 ear domains but lacking the platform subdomain (13–15). The GAE domain, however, does not interact with the same accessory proteins and motifs that bind to the α and β2 ear domains. Instead, it interacts with a distinct set of accessory proteins that includes Rabaptin-5 (16, 17), γ-synergin (18, 19), enthoprotin/Clint/epsinR (20–23), and p56 (15) in mammalian cells and Ent3p and Ent5p in yeast (24).

We have recently determined that Rabaptin-5 binds to the GAE domains of the γ-adaptins and the GGAs via the sequence FGPLV (residues 439–443 of Rabaptin-5) (25). Substitution of Phe (residue 439), Gly (residue 440), Leu (residue 442), and, to some extent, Val (residue 443) by Ala impairs the recognition of Rabaptin-5 by the GAE domains, whereas substitution of Pro (residue 441) is less critical (25). Thus, hydrophobic contacts between the FGPLV sequence and GAE domains are the main determinants of binding; however, acidic residues preceding

* 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AY367088.

¶ Supported by a Jean Timmons Costello Fellowship from the Montreal Neurological Institute.

** A Canadian Institutes of Health Research (CIHR) investigator. Work in this author's laboratory was supported by CIHR Grant MT-13461.

‡‡ To whom correspondence should be addressed: Cell Biology and Metabolism Branch, NICHD, Bldg. 18T/Rm. 101, National Institutes of Health, Bethesda, MD 20892. Tel.: 301-496-6368; Fax: 301-402-0078; E-mail: juan@helix.nih.gov.

[1] The abbreviations used are: AP, adaptor protein complex; AD, Gal4 transcription activation domain; BD, Gal4 DNA-binding domain; GAE, γ-adaptin ear domain; GGA, Golgi-localized, γ ear-containing, Arf-binding protein; NECAP, adaptin-binding endocytic coat-associated protein; GST, glutathione *S*-transferase; GFP, green fluorescent protein.

this sequence provide a stabilizing effect that is noticeable when assaying relatively short (406–476 or 428–455) but not longer fragments of Rabaptin-5 (25). In line with these observations, crystallographic studies demonstrated that the Phe and Leu residues constitute the main points of attachment of the FGPLV sequence to the GAE domain of GGA3 (26), with the Gly and Pro being important for the correct spacing of the anchoring residues. The preceding acidic amino acids and other neighboring residues contribute secondary, less critical, interactions.

Analyses of GAE-binding determinants from enthoprotin, Ent3p, Ent5p, and p56 have also identified short sequences that contain a critical Phe separated by two residues from another Phe or Trp. In the case of p56, these two residues anchor the sequence to the GAE domain of GGA1 in a manner similar to that of the key Phe and Leu residues of Rabaptin-5 binding to the GGA3-GAE domain (27). Thus, all of these proteins seem to bind to the GAE domains by virtue of a conserved peptide motif. Comparison of known or putative GAE-binding sequences has led to the proposal of various consensus motifs, including (D/E)F*X*D(F/W) (23), $(D/E)_{2-3}F\mathit{XX}\emptyset$ (24), and $(D/E)(G/A)_{0-1}F(G/A)(D/E)\emptyset$ (28) (where *X* represents any amino acid, and Ø is a hydrophobic residue). However, the requirement for specific residues at different positions within the motif has not been systematically analyzed. The experimental analysis of these requirements is important not only to define a canonical motif that describes all known GAE-binding sequences but also to predict novel GAE-binding partners. We have undertaken such an analysis using *in vitro* binding assays, the yeast two-hybrid system, and the screening of combinatorial phage display peptide libraries. This combined experimental approach has allowed us to define a general consensus GAE-binding motif and to identify novel ligands for the GAE domain.

EXPERIMENTAL PROCEDURES

Recombinant DNA

pGAD Constructs—Mutations in pGAD424-Rabaptin-5-(406–476) or pGAD10-*myc*-Rabaptin-5-(5–476) templates (25) were introduced using the QuikChange mutagenesis kit (Stratagene, La Jolla, CA) and 35-mer primers from Sigma. cDNA fragments encoding mouse NECAP 1-(171–275) and NECAP 2-(170–266) were amplified from full-length constructs in pCMV-Tag-2B (Invitrogen) and subcloned into the EcoRI/BamHI sites of pGAD424. The NECAP 1-(171–275) construct contained K175R and V273F substitutions, whereas the NECAP 2-(170–266) construct contained T174S and A225T substitutions when compared with the wild type sequences. The cDNAs encoding full-length aftiphilin as well as a 1–523 fragment of this protein were cloned by PCR amplification from a human brain cDNA library (Clontech, Palo Alto, CA) and subcloned into the BamHI site of pGAD424.

pGBT9 Constructs—The pGBT9 constructs expressing BD fusions to the γ1- and γ2-adaptin ears and GAE domains of GGA1, GGA2, and GGA3 were described in Ref. 25.

pEGFP Constructs—The cDNA encoding full-length aftiphilin was subcloned into the BamHI site of pEGFP-C2 (Clontech).

GST Fusion Constructs—Vectors directing the expression of GST fusions to the mouse γ1-adaptin ear (residues 707–822; Fig. 2), human γ2-adaptin ear (residues 670–785), human GGA1-GAE (residues 515–639), human GGA1-hinge + GAE (residues 315–639), human GGA2-GAE (residues 489–613), human GGA2-hinge + GAE (residues 331–613), and human GGA3-GAE (residues 494–723, comprising also a fragment of the hinge region; Fig. 1) were described in Ref. 25. The pGST-parallel 2-GGA3-GAE (residues 571–723, wild type, comprising a smaller portion of the hinge and used in the experiments shown in Figs. 2 and 5*C*) was described in Ref. 26. The GST fusion to NECAP 1 residues 250–256 (DLWGDFS) was generated by annealing oligonucleotides 5′-AATTCGACTTGTGGGGAGACTTCAGCTGAC-3′ (sense) and 5′-TCGAGTCAGCTGAAGTCTCCCCACAAGTCG-3′ (antisense) and subcloning into the EcoRI/XhoI sites of pGEX-4T1 (Amersham Biosciences). The GST fusion to NECAP 2 residues 241–247 (DIWGDFT) was obtained by a similar procedure using oligonucleotides 5′-AATTCGATATCTGGGGAGACTTCACCTGAC-3′ (sense) and 5′-TCGAGTCAGGTGAAGTCTCCCCAGATATCG-3′ (antisense). The pGST fusion to the mouse γ1-adaptin ear (residues 704–822) shown in Fig. 5*B* was generated by PCR amplification of I.M.A.G.E. consortium clone ID 1265666 and subcloning into the BamHI/SmaI sites of pGEX-2T.

His_6-tagged Proteins—cDNA clones encoding His_6-tagged mouse NECAP 1 and NECAP 2 were described in Ref. 32.

Antibodies

The rabbit anti-GST antiserum was described in Ref. 29. Anti-α-adaptin and anti-γ-adaptin monoclonal antibodies (Fig. 5*B*) were purchased from BD Transduction Laboratories (Lexington, KY), and the anti-His monoclonal antibody was from Qiagen (Valencia, CA). The monoclonal anti-γ1-adaptin used for immunofluorescence (clone 100/3) was obtained from Sigma, whereas the rabbit anti-GFP was from Molecular Probes, Inc. (Eugene, OR). The secondary antisera used for immunofluorescence were purchased from Jackson Immunoresearch Laboratories (West Grove, PA).

Binding Assays with Biotinylated Peptides

Peptides were obtained from New England Peptide (Fitchburg, MA). Biotinylation mixtures contained 0.2 μmol of peptide and 0.33 μmol of EZ-link™ PEO-maleimide-activated biotin (Pierce) in 200 μl of phosphate-buffered saline, pH 7.0. Mixtures were incubated overnight at room temperature and quenched by the addition of 10 mM β-mercaptoethanol. Immobilization of biotinylated peptides (5 nmol), incubation with GST fusions (3 μg), washing of beads, and analysis of bound proteins was performed as described in Ref. 25.

Yeast Two-hybrid Analysis

Transformation of AH109 cells, isolation of double transformants, and evaluation of the interaction of fusion proteins were performed as previously described (25).

Isolation of GAE Domain Binding Peptides from a Phage-displayed Library

Previously described methods (30, 31) were used to construct a library of random octapeptides fused to the N terminus of the M13 bacteriophage major coat protein (2×10^{10} unique members). Phage pools displaying the library were cycled through rounds of binding selection with GST fusion proteins coated on 96-well MaxiSorp™ immunoplates (Nalge Nunc, Rochester, NY) as capture targets. The library was sorted separately against each target, and phage were propagated in *Escherichia coli* XL1-blue cells (Stratagene) supplemented with both M13-KO7 helper phage (New England Biolabs, Beverly, MA) to facilitate phage production and 10 μM isopropyl-1-thio-β-D-galactopyranoside to induce expression of the library. After three rounds of selection, phage from individual clones were analyzed in a phage enzyme-linked immunosorbent assay (30). Phage that bound to the target GST fusion protein but not to a control GST fusion were subjected to DNA sequence analysis.

GST Fusion Pull-down Assays

Preparation of brain extracts and pull-down assays were performed as previously described (32). In the experiments shown in Fig. 5*A*, aliquots of the brain extracts (2 mg) were incubated for 2 h at 4 °C with GST fusion proteins (5 μg) precoupled to glutathione-Sepharose (Amersham Biosciences). In the experiments shown in Fig. 5*B*, the immobilized GST fusion proteins were incubated for 2 h at 4 °C with purified His_6-NECAP 1 or His_6-NECAP 2 (1 μg).

Immunofluorescence Microscopy

HeLa cells (plated on 12-mm coverslips and grown to ~30% confluence) were transfected with pGFP-C2-human aftiphilin (2.5 μg/well) using the Fugene™ reagent (Roche Applied Sciences) and incubated for an additional 18–24 h. Coverslips were quickly washed in phosphate-buffered saline and immediately fixed for 10 min in methanol/acetone (1:1) at −20 °C. Incubation with primary and secondary antibodies was performed as previously described (25). Images were obtained in an inverted confocal laser-scanning microscope (LSM410; Carl Zeiss Inc., Thornwood, NY).

RESULTS

Selectivity of Binding of Rabaptin-5, γ-Synergin, and Enthoprotin Peptides to GAE Domains—As a first step to define a consensus motif for binding to GAE domains, we analyzed various peptides containing candidate sequences from γ-syner-

TABLE I
Binding of GAE domains to biotinylated peptides from Rabaptin-5, γ-synergin, enthoprotin, and NECAP 1

GAE partner	Residues[a]	Peptide sequence[b] 0 +3	Binding to GST-GAE fusions[c]				
			γ1	γ2	GGA1	GGA2	GGA3
Rabaptin-5	435-447	DESDFGPLVGADS	+	+	+	Weak	+
γ-Synergin	666-675	LADDFGEFSL	ND	+	−	−	−
γ-Synergin	666-678	LADDFGEFSLFGE	+	+	+	−	−
γ-Synergin	772-781	SDDDFADFHS	ND	−	−	ND	ND
Enthoprotin	340-352	SADLFGGFADFGS	+	+	−	+	−
Enthoprotin	368-377	GNGDFGDWSA	ND	+	−	−	−
Enthoprotin	368-380	GNGDFGDWSAFNQ	+	+	−	−	−
Enthoprotin	420-429	SSDLFDLMGS	ND	−	−	ND	ND
NECAP 1	248-260	SNDLWGDFSTASS	+	+	Weak	−	−

[a] The numbering corresponds to sequences of the human proteins (the 248-260 sequence in human NECAP 1 is identical to the 248-260 sequence in the mouse protein).
[b] The 0 and +3 positions in the known or putative GAE-binding sites are underlined.
[c] The columns with information on the binding to GST-GAE fusions summarize the results shown in Figs. 1 and 2 and in Fig. 5*C* (NECAP 1 peptide) as well as additional experiments with the SDDDFADFHS and SSDLFDLMGS peptides. ND, not determined.

gin and enthoprotin. The candidate sequences were selected based on (*a*) the identification of the GAE-binding sequence FGPLV in Rabaptin-5 (25), (*b*) the presence of GAE-binding sites in the 518–786 segment of γ-synergin (18) and the 370–451 segment of enthoprotin (20), and (*c*) the analysis of GAE-binding sequences from enthoprotin (23) and Ent3p/Ent5p (24). The initial set of peptides tested (Table I) included LADDFGEFSL and SDDDFADFHS (residues 666–675 and 772–781, respectively, in human γ-synergin) and GNGDFGDWSA and SSDLFDLMGS (residues 368–377 and 420–429, respectively, in human enthoprotin) (residues at positions 0 and +3 in known or putative GAE-binding sequences have been underlined throughout). The enthoprotin-(368–377) peptide is similar to the P5 peptide used by Mills *et al.* (23). We used the DESDFGPLVGADS peptide (residues 435–447 in human Rabaptin-5) for comparison and also designed variant peptides containing alanine substitutions in place of the hydrophobic residues present at positions 0 and +3 (Fig. 1). All peptides included an amino-terminal Cys for conjugation to biotin.

These biotinylated peptides were assayed for binding to GST fusion proteins comprising the GAE domains of γ2-adaptin, GGA1, GGA2, and GGA3. The Rabaptin-5-(435–447) peptide bound to all four GST-GAE domains but not to the GST-β3B ear used as a negative control (Fig. 1). The order of reactivity of the Rabaptin-5-(435–447) peptide for the GGA-GAE domains (GGA1 > GGA3, with minimal binding to GGA2) was similar to that observed in pull-downs of full-length Rabaptin-5 from bovine brain cytosol by various GST-GGA-GAE fusion proteins, as was the lack of binding to the Rabaptin-5 peptide containing the F439A and L442A substitutions (positions 0 and +3) (25). Importantly, only one peptide from each pair of γ-synergin and enthoprotin peptides (LADDFGEFSL and GNGDFGDWSA, respectively) bound specifically to the GST-γ2-adaptin ear fusion protein, although not to the other GST-GAE fusion proteins (Fig. 1). The other two peptides (SDDDFADFHS and SSDLFDLMGS) did not bind to any of the GAE domains tested (data not shown; results summarized in Table I). Variants of the active γ-synergin and enthoprotin peptides containing Ala substitutions at 0 and +3 did not bind to the GST-γ2-adaptin ear (Fig. 1).

The lack of binding of the LADDFGEFSL and GNGDFGDWSA peptides to the GGA-GAE domains could have been due to a low affinity of the corresponding interactions. Indeed, the peptide-binding assays were performed using 4–10 μM biotinylated peptides together with a 0.05–0.1 μM concentration of the GST fusion proteins and, thus, were designed to detect interactions in the low micromolar (~0.5–10 μM) range of affinities (use of higher concentrations of biotinylated peptides and GST fusions resulted in high nonspecific binding). Therefore, interactions in the high micromolar range of affinities could not be detected in these direct binding assays. In this context, we have recently demonstrated using competition-binding assays that both the LADDFGEFSL and GNGDFGDWSA peptides bind to GST-GGA3-GAE, albeit with affinities that are 25–50 times lower than that of the Rabaptin-5 DESDFGPLVGADS peptide (26).

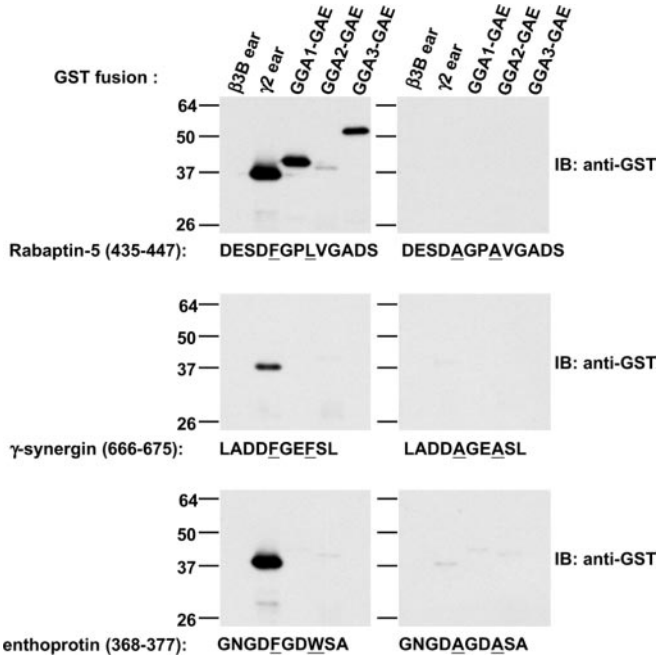

FIG. 1. **Binding of GAE domains to biotinylated peptides (10-mers) derived from γ-synergin and enthoprotin.** Peptides (10-mers) containing candidate GAE-binding sequences from human γ-synergin (LADDFGEFSL and SDDDFADFHS; residues 666–675 and 772–781, respectively) and human enthoprotin (GNGDFGDWSA and SSDLFDLMGS, residues 368–377 and 420–429) were designed to contain an additional amino-terminal Cys for biotinylation. The 0 and +3-positions in the GAE domain-binding sequences are *underlined* in the *panel labels* (the amino-terminal Cys residues are not shown). Additional biotinylated peptides containing Ala substitutions at the corresponding 0 and +3-positions (*right panels*) were used as specificity controls. A biotinylated peptide containing the GAE-binding sequence from Rabaptin-5 (residues 435–447) (25) and the corresponding control peptide with Ala substitutions at 0 and +3 were used for comparison. The biotinylated peptides were immobilized on streptavidin-agarose and incubated with the indicated GST fusions. Bound GST fusions were eluted and identified by SDS-PAGE and immunoblotting (*IB*) with anti-GST serum. The biotinylated peptides SDDDFADFHS (γ-synergin) and SSDLFDLMGS (enthoprotin) did not show detectable binding to any of the GAE domains (results not shown, but summarized in Table I). *Numbers* on the *left* of each *panel* indicate the positions of molecular mass markers (in kDa).

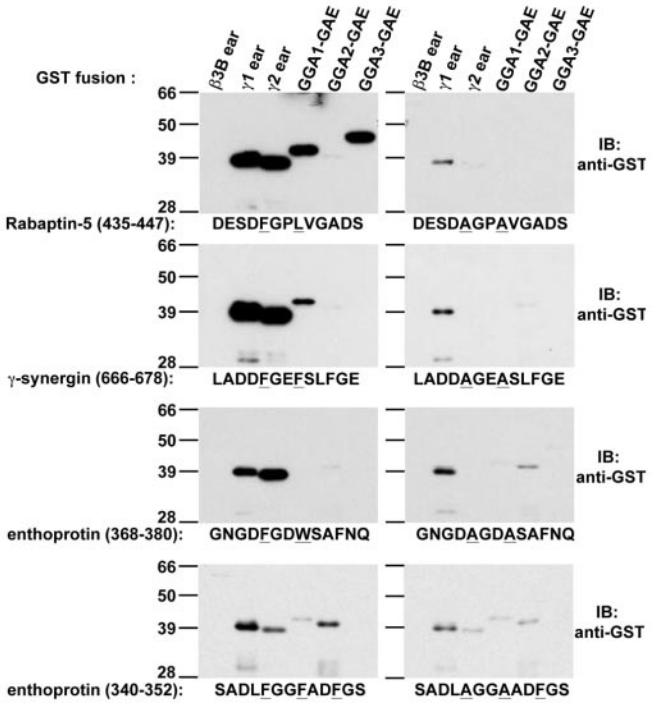

FIG. 2. **Binding of GAE domains to biotinylated peptides (13-mers) derived from Rabaptin-5, γ-synergin, and enthoprotin.** We designed additional peptides (13-mers) corresponding to the candidate GAE-binding sequences in human γ-synergin and human enthoprotin as well as their corresponding controls with Ala substitutions at the 0 and +3-positions (sequences shown under panels; residues at motif positions 0 and +3 are *underlined*). All peptides contained amino-terminal Cys residues for biotinylation. The sequence corresponding to the enthoprotin (residues 340–352) peptide contains two putative motifs in tandem (position +3 of the first motif corresponds to position 0 in the second). The experiments were performed as indicated in the legend to Fig. 1.

The γ-synergin and enthoprotin peptides (10-mers) shown in Fig. 1 were shorter than the Rabaptin-5 peptide (13-mer). Differences in peptide length could account for their preferential recognition by specific GAE domains. This issue was addressed by assaying 13-mer peptides corresponding to the positive sequences in γ-synergin and enthoprotin and including the same number of carboxyl-terminal residues present in the Rabaptin-5 peptide (Table I). The longer γ-synergin peptide (residues 666–678) bound not only to γ-adaptins (both γ1- and γ2-adaptins were used in this experiment) but also weakly to GGA1-GAE, whereas the longer enthoprotin peptide (residues 368–380) only interacted with the γ-adaptins (Fig. 2; summary of results in Table I). An additional enthoprotin 13-mer (residues 340–352), containing two overlapping candidate motifs, showed specific interactions with the GAE domains of γ1- and γ2-adaptins and GGA2 (Fig. 2; Table I).

Two related conclusions were drawn from these experiments. First, the Gly at position +1 is important for GAE binding, as demonstrated by the presence of this residue in all active peptides from γ-synergin and enthoprotin (Table I). This is consistent with the observation that substitution of Ala for Gly in the DFGPLV sequence from Rabaptin-5 abrogated binding to GAE domains (25) (see also Fig. 5*A* in this study), arguing against the equivalency of these residues at position +1. Second, the proposed GAE-binding motifs (D/E)F*X*D(F/W) (23), $(D/E)_{2-3}FXXØ$ (24), and $(D/E)(G/A)_{0-1}F(G/A)(D/E)Ø$ (28) required further elaboration, given that the peptides SDDDFADFHS (residues 772–781 in γ-synergin) and SSDLFDLMGS (residues 420–429 in enthoprotin) match the above motifs but did not detectably bind to any of the GAE domains tested (Table I). This was further evidenced by the discrepancy between the Pro (position +2) in the Rabaptin-5 GAE-binding sequence (DFGPLV) and the proposed motifs.

Yeast Two-hybrid Analysis of Allowable Substitutions in the GAE-binding Motif—To further define the consensus GAE-binding motif, we studied the effect of systematic substitutions in Rabaptin-5 residues 439–442 on the binding to GAE domains. This was analyzed by the yeast two-hybrid system using Gal4 transcription activation domain (AD) fusions to the Rabaptin-5 sequences and Gal4 DNA-binding domain (BD) fusions to the GAE domains of γ1-adaptin, γ2-adaptin, GGA1, GGA2, and GGA3.

Residues Allowed at Positions 0 and +3—We substituted hydrophobic amino acids for either the Phe (position 439) or Leu (position 442) residues of Rabaptin-5 (residues 0 and +3 of the GAE-binding motif, respectively). The mutations were introduced in two Rabaptin-5 fragments, namely 406–476 and 5–476, in order to assess whether their effect was dependent on the fragment size. We used as controls the F439A and L442A substitutions, which do no interact with GAE domains when introduced in the context of the Rabaptin-5-(5–476) fragment (25). The results shown in Fig. 3*A* demonstrate that substitution of Phe^{439} in the Rabaptin-5-(406–476) fragment by Trp resulted in stronger interactions with all GAE domains. Substitution of Phe^{439} by Tyr in the same Rabaptin-5 fragment maintained interactions with the GAE domains of γ1-adaptin, GGA2, and GGA3 but weakened those with the γ2-adaptin ear and GGA1-GAE. All other hydrophobic amino acid substitutions for Phe^{439} abolished binding, except for a weak interaction between the F439M mutant and GGA3-GAE. On the other hand, the substitution of Leu^{442} was more permissive; replacement with either Met, Trp, Tyr, or Phe allowed interactions with all five GAE domains, albeit with varying avidities. Similar results were obtained when the 0 and +3 substitutions were made in the context of the Rabaptin-5-(5–476) fragment (Fig. 3*B*), except that the F439Y mutant showed broader interactions with GAE domains than those seen when introduced in the smaller template, and the F439M mutant interacted with both GGA1-GAE and GGA3-GAE (Fig. 3*B*). The results obtained with all GAE domains shown in Fig. 3 thus indicate that position 0 of the motif (Rabaptin-5 residue 439) can be occupied by either Trp, Phe, or Tyr and that the position +3 (Rabaptin-5 residue 442) can be occupied by these same residues as well as Leu or Met (Table II).

Residues Allowed at Position +1—Given the similar results obtained with the Phe^{439} or Leu^{442} substitutions in the two Rabaptin-5 templates, we only introduced the Gly^{440} substitutions (position +1 of the motif) in the smaller fragment. We replaced Gly^{440} by a more limited set of residues, based on our previous results and on those reported by others. We examined the effect of its substitution by acidic residues given that the presence of an Asp at position +1 of the enthoprotin sequence SSDLFDLMGS results in a peptide that does not bind GAE domains, despite the presence of appropriate hydrophobic residues at the 0- and +3-positions (Table I). We also substituted Gly^{440} by Ser, given the presence of this residue at the +1-position in one of the two acidic Phe motifs in Ent3p (DDDEFSEFQ, residues 268–276) (24) and also by Val, to mimic another related sequence in this protein (EPEDFVDFF, residues 235–243). We observed that substitution of Gly^{440} in Rabaptin-5-(406–476) by either Glu, Val, or Ser abrogated the binding of this fragment to the various GAE domains (Fig. 4*A*). The G440A substitution was also inactive, save for a weak interaction with GGA3-GAE, as previously reported for the same mutation introduced in the Rabaptin-5-(5–476) fragment (Fig. 6 in Ref. 25). These observations indicate that the Gly at

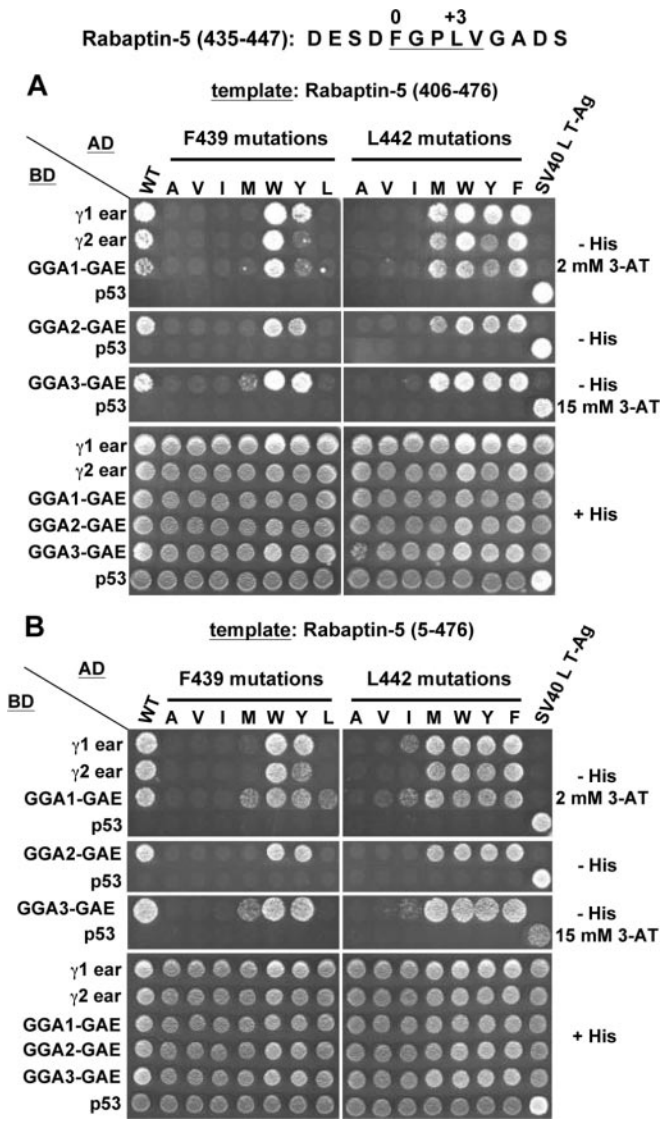

FIG. 3. **Yeast two-hybrid analysis of hydrophobic residue substitutions at the 0 or +3-positions (Phe^{439} and Leu^{442}, respectively) in the GAE-binding sequence of Rabaptin-5.** The indicated substitutions were introduced in the Gal4 AD constructs pGAD-Rabaptin-5-(406–476) (*A*) or pGAD-Rabaptin-5-(5–476) (*B*), and the interactions of the resulting mutants were assayed following co-transformation of the AH109 strain with Gal4 BD pGBT9-GAE constructs. Co-transformants were plated on medium without histidine (−*His*), to detect *HIS3* reporter gene activation upon interaction of constructs, and plated on medium with histidine (+*His*) as a control for loading and growth of the co-transformants. Controls for nonspecific interactions included co-transformation of pGAD-Rabaptin-5-(406–476) mutants with a BD-p53 construct as well as co-transformation of the pGBT9-GAE vectors with an AD-SV40 large T-antigen construct (*T-Ag*). Co-transformation with vectors encoding the BD-p53 and AD-SV40 large T-antigen fusions provided a positive control for interactions. Different concentrations of 3-aminotriazole (*3-AT*; a competitive inhibitor of the His3 protein) were required to prevent nonspecific interactions and background growth. Interactions with GGA2-GAE were studied in −His plates without 3-AT; interactions with the GAE domains of the γ1- and γ2-adaptins and GGA1 required −His plates with 2 mM 3-AT, whereas those with GGA3-GAE required −His plates containing 15 mM 3-AT.

position +1 is critical and cannot be replaced by other residues without a drastic loss of binding.

Residues Allowed at Position +2—We substituted acidic residues for Pro^{441} in Rabaptin-5 (+2-position in the motif), given the presence of acidic residues at the equivalent position of the γ-synergin (residues 666–678) and enthoprotin (residues 368–380) sequences (Table I). Additional substitutions of Lys, Phe, Gly, or Val for Pro^{441} in Rabaptin-5-(406–476) were introduced to test whether other residues would be effective at this position. We had previously concluded that the P441A substitution in Rabaptin-5-(5–476) did not significantly alter the binding to GAE domains, although a partial reduction in some interactions (mainly with GGA2-GAE and, to a lesser extent, with the γ2-adaptin and GGA1-GAE domains) was nonetheless noticeable when compared with the wild-type construct (Figs. 6 and 7 in Ref. 25). When the same mutation was introduced in the smaller template Rabaptin-5-(406–476), the effects were more marked; this mutant only interacted with the GGA3-GAE (and, minimally, with γ1-adaptin GAE) (Fig. 4*A*). A total loss of interactions was observed when Pro^{441} was replaced by either Lys, Phe, Gly, or Val (Fig. 4*A*). In contrast, the P441D and, particularly, P441E mutants interacted with the GAE domains (again significantly stronger with GGA3 and γ1-adaptin GAE and weaker with the other constructs) (Fig. 4*A*). These results indicate that the +2-position in the motif can be occupied by Pro, and, less effectively, by Glu or Asp (Table II).

We also examined the effect of introducing an acidic amino acid at +2 in combination with a Phe at +3, to mimic the pair present in the γ-synergin-(666–678) sequence. Given that the −1, 0, and +1-positions in γ-synergin-(666–678) are occupied by Asp, Phe, and Gly, respectively (as also found in the Rabaptin-5 GGA-GAE-binding sequence), the P441E/L442F substitution resulted in a Rabaptin-5-(406–476) construct containing a sequence from −1 to +3 that was identical to that present in the γ-synergin-(666–678) GAE-binding sequence. The constructs with DF or EF combinations at the +2/+3-positions interacted with the various GAE domains in a manner similar to Rabaptin-5-(406–476) (Fig. 4*A*). This supported the previous conclusion that constructs with acidic residues, particularly Glu, at +2 display interactions similar to that observed with Pro at this position. These observations also indicate that the EF combination at +2/+3 in the γ-synergin-(666–678) peptide (resulting in a DFGEF sequence from −1 to +3, as shown in the *last lane* of Fig. 4*A*) is not responsible for the lack of binding of this peptide to GGA3-GAE *in vitro* (Fig. 2).

En Bloc Replacement of the FGPL Sequence from Rabaptin-5 by Sequences in Other Proteins—We substituted *en bloc* the residues at the 0 to +3-positions of the Rabaptin-5 sequence (in the 405–476 fragment) to obtain the cognate sequences from other GAE-binding proteins. The substitutions generated the following sequences (Table I): FADF (residues 776–779 in γ-synergin), FGDW (residues 372–375 in enthoprotin), FDLM (residues 424–427 in enthoprotin), and FGEM (residues 379–382 in Ent3p) (24). Interestingly, two recently identified proteins that are enriched in clathrin-coated vesicles (gi 27229051 ref NP_080543.2 and gi 13384758 ref NP_079659.1) (20), named NECAP 1 and NECAP 2 (32), contain WGDF sequences (residues 252–255 and 243–246 in mouse NECAP 1 and NECAP 2, respectively; identical sequences are also present in the human orthologs of these proteins). Given that the WGFD sequence satisfied the emerging criteria for the 0 to +3-positions in a GAE-binding motif (Table II), we also introduced these substitutions in Rabaptin-5-(406–476). The sequences FGDW (from enthoprotin) and WGDF (from NECAP 1 and NECAP 2) interacted with all GAE domains, whereas the construct containing the FDLM substitution did not interact with any of them (Fig. 4*B*). These results support the conclusions summarized in Table II, particularly the importance of the Gly at position +1 and the presence of Trp at 0, Asp at +2 and Trp and Phe at +3 in active GAE-binding sequences. In contrast, the FADF and FGEM-containing constructs only interacted with GGA3-GAE. The lack of binding of the FADF-containing construct to the GAE domains

TABLE II
Active residues at different positions of the GAE domain-binding motif

Position	Active residues[a]	Results from
−1, −2 or −3	Asp, Glu	Table III
0	Trp, Phe, Tyr	Figs. 1–5 and 7; Table III
+1	Gly	Figs. 1, 2, 4, 5, and 7; Table III
+2	Pro, Asp, Glu	Figs. 1, 2, 4, 5, and 7; Table III
+3	Trp, Phe, Leu, Met, Tyr	Figs. 1–5 and 7; Table III
+4	Val, Ala, Gly, Ser, Thr, Gln	Figs. 1, 2, 5, and 7, Table III; sequence comparisons
Summary[b]	(D/E)(L/I/A/G)$_{0-2}$(W/F/Y)G(P/D/E)(W/F/Y/L/M)(V/A/G/S/T/Q)	
	0 +1 +2 +3	
Tetrapeptide consensus (0 to +3)[c]	(W/F/Y)G(P/D/E)(W/F/Y/L/M)	
	or	
	ψ G(P/D/E)(ψ/L/M)	

[a] The active residues were defined based on the information indicated in the last column.
[b] The (L/I/A/G)$_{0-2}$ separating the amino-terminal acidic residues from the 0-position was defined based on the information presented in Table III and on the sequences in NECAP 1 and NECAP 2 (Figs. 5 and 7) and other candidate GAE partners (see "Discussion").
[c] ψ symbolizes residues with aromatic side chains (Trp, Phe, or Tyr).

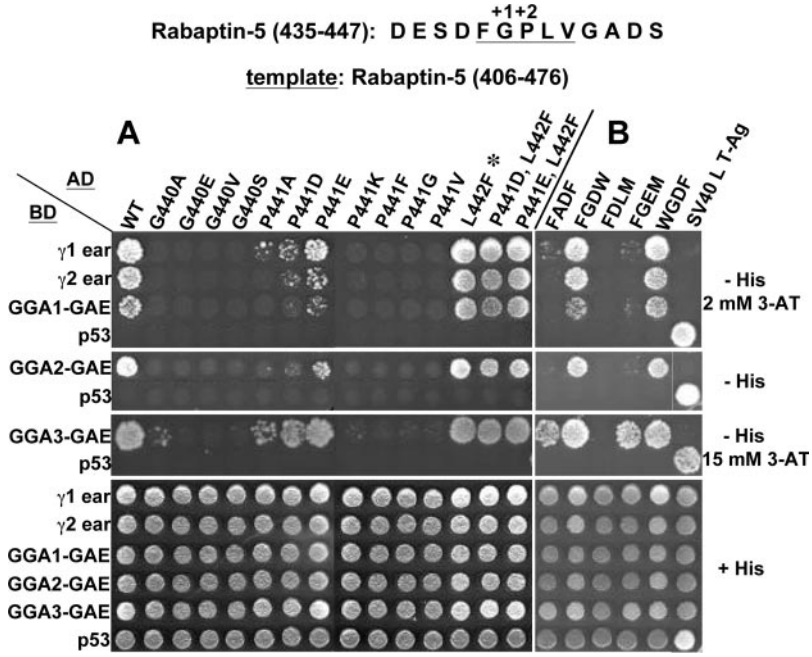

FIG. 4. **Effect of single or multiple substitutions in the GAE-binding sequence of Rabaptin-5.** *A*, the indicated single substitutions at positions +1 or +2 of the GAE-binding sequence (residues Gly[440] and Pro[441], respectively) were introduced in the Rabaptin-5-(406–476) template. Additional P441D and P441E substitutions were also introduced in the Rabaptin-5-(406–476) L442F mutant analyzed in Fig. 3*A* (also included as control in this experiment; see *asterisk* label). *B*, *en bloc* substitutions in the Rabaptin-5 GAE-binding sequence (406–476 fragment). These substitutions were introduced to mimic the sequences present at positions 0 to +3 in putative GAE-binding motifs from γ-synergin (FADF, residues 776–779), enthoprotin (FGDW, residues 372–375; FDLM, residues 424–427), Ent3p (FGEM, residues 379–382), and NECAP 1/2 (WGDF, residues 252–255 and 243–246 in mouse NECAP 1 and mouse NECAP 2, respectively).

of γ1-adaptin, γ2-adaptin, GGA1, and GGA2 is consistent not only with the results obtained with the γ-synergin-(772–781) peptide containing this sequence (Table I) but also with the absence of the critical Gly at position +1 in the motif. The recognition of the FADF construct by GGA3-GAE indicates that this yeast two-hybrid construct is more permissive than the others and may still recognize, albeit less effectively, a sequence with Ala instead of Gly at +1 (see also results with the single G440A mutant in Fig. 4*A* and the effect of the same substitution in the Rabaptin-5-(5–476) reported in Ref. 25). The very weak binding of the FGEM-containing Rabaptin-5-(406–476) construct to all but the GGA3-GAE domain may indicate that, although both the FGPM and FGEF combinations from 0 to +3 are active (Figs. 3 and 4*A*, *last lane*, respectively), the presence of Glu at +2 restricts the ability of most of the GAE domains (with the exception of the more permissive GGA3-GAE) to recognize a sequence containing Met at +3.

Taken together, the yeast two-hybrid analysis of mutants in Figs. 3 and 4 pointed to a (W/F/Y)G(P/D/E)(W/F/Y/L/M) consensus or more concisely ΨG(P/D/E)(Ψ/L/M) (where Ψ symbolizes an aromatic residue), for positions 0 to +3 (Table II).

Combinatorial Analysis of the GAE-binding Motif by Phage Display Peptide Library Screening—We also identified GAE domain-binding peptides by screening a phage display library using GST fusions to the GAE domains of γ1-adaptin, γ2-adaptin, GGA1, GGA2, and GGA3. The results of the screening (Table III) can be summarized as follows: (*a*) all GAE domains bound similar peptide sequences; (*b*) acidic residues were frequently observed preceding the 0-position; (*c*) with one exception (Phe at position 0 in one of the clones selected by GGA3-GAE), all peptides contained Trp at 0 and +3, consistent with the strong binding observed when this residue was introduced at the corresponding positions of the Rabaptin-5-(406–476) and -(5–476) fragments (Fig. 3, *A* and *B*, respectively); (*d*) almost all peptides contained Gly at +1, in agreement with the sequences of the GAE-binding peptides derived from Rabaptin-5, γ-synergin, and enthoprotin (Table I) and with the yeast two-hybrid analysis (Fig. 4*A*); (*e*) Glu or Pro occupied the +2-position in most peptides (and in all of the frequent binders, with the exception of two that were selected with GGA3-GAE), again consistent with the yeast two-hybrid analyses in Fig. 4*A* (isolation of a GGA3-GAE frequent binder containing Ala at +2 also supported our conclusion that this is the most permissive of the GAE domains we have compared (Fig. 4*A*, *second lane* from *left*)); (*f*) with the exception of Cys, which increases the apparent affinity of peptides and the likelihood of selection through disulfide bridge dimerization, Val was the residue most frequently found at position +4 (coincident with its pres-

TABLE III
Phage-displayed peptides selected with GAE domains

GST fusion target	Selected peptide[a] (0 … +3)	n[b]
γ1-Adaptin ear	D L K W G E W **C**	12
	T W G P W M R W	2
	D L K W G E W F	1
γ2-Adaptin ear	G E E W G P W V	6
	D L K W G E W **C**	4
	A L E W G A W V	1
	S D W A P W V G	1
	E T W **C** W S T W	1
	F G E W V G W R	1
	D L K W G Q W F	1
GGA1-GAE	D L K W G E W C	3
GGA2-GAE	D L K W G E W C	10
	E W G P W V A A	4
	T R W T E W **C** G	1
GGA3-GAE	E T I W G A W V	6
	L D W G M W V Q	2
	G E W G P W I V	2
	G S W G P W V G	2
	G V D W G R W V	1
	L H Q W G A W V	1
	D E F G H W V V	1

[a] Residues at the 0 and +3 positions are underlined. The presence of cysteine residues (boldface type) in some of the selected peptides increases their likelihood of selection through disulfide bridge dimerization.
[b] n indicates the number of times that the indicated peptide was isolated.

ence at the cognate position of Rabaptin-5). The peptides selected in the screening thus conform to a (D/E)WG(P/D/E)WV motif, which probably corresponds to a subset of peptides with the highest affinity for GAE domains, characterized by the presence of Trp at 0 and +3.

Identification of Novel GAE-binding Partners on the Basis of the Consensus Motif—The definition of the GAE-binding motif allowed us to identify novel interaction partners for the GAE domain. As shown in Fig. 4*B*, substituting the WGDF sequence from NECAP 1 and NECAP 2 for the FGPL residues at 0 to +3 in Rabaptin-5-(406–476) generated a construct that interacted with the different GAE domains. We tested *in vitro* for the binding of rat brain adaptors to GST fusion proteins containing the GAE-binding sequences from NECAP 1 (DLWGDFS) and NECAP 2 (DIWGDFT). The resulting immunoblots revealed an interaction of both NECAP peptides with AP-1 but not AP-2 (Fig. 5*A*). We also measured the interaction of GST fusion proteins containing the γ1-adaptin GAE and the hinge + GAE domains of GGA1 and GGA2 with His_6-tagged NECAP proteins. We observed a strong interaction of the GST-γ1-adaptin ear with both His_6-NECAP 1 and His_6-NECAP 2 constructs, and the latter was also able to bind to the GST-GGA-hinge + GAE constructs (Fig. 5*B*). Last, we tested the interaction of the panel of GST-GAE constructs with a biotinylated 13-mer peptide representing the 248–260 region of NECAP 1. The results confirmed a strong *in vitro* interaction of NECAP 1 with the γ-adaptin GAE domains (Fig. 5*C*). Thus, the presence of a GAE-binding motif in the NECAPs and their localization to clathrin coats indicate that they are relevant binding partners for the γ-adaptins and the GGAs.

Mills *et al.* (23) recently reported that the hypothetical human protein FLJ20080 (gi 31377758 ref NP_060127.2) contains several candidate GAE-binding sequences. Using PCR amplification, we cloned from a human brain library a cDNA encoding a protein closely related to FLJ20080. This hypothetical protein differs from the reported FLJ20080 in (*a*) an additional stretch of 28 amino acids between residues 818 and 819, (*b*) a P822L substitution, and (*c*) the deletion of S860 (Fig. 6*A*). The resulting hypothetical protein has 936 amino acids, as compared with 909 in FLJ20080, and contains multiple putative GAE-binding sequences, some of which are overlapping, an EF-hand domain, and an LLNLD clathrin-binding motif (29) (Fig. 6, *A* and *B*). Because of its affinity for earlike domains (see below), we have named this protein aftiphilin (from the Greek afti (ear) and philos (friend)). The putative GAE-binding sequences in aftiphilin (all located in predicted random coils) are EFGEFG (residues 27–32); DFGDFGDFG (overlapping sequences corresponding to residues 431–439); and DFGEFG (residues 477–482). Transient expression of GFP-tagged full-length aftiphilin in HeLa cells showed a predominantly cytosolic localization, with some discrete areas of co-localization with γ1-adaptin (Fig. 6, *C–E*) and clathrin (data not shown). Using the yeast two-hybrid system, we tested the possible interaction of GAE domains with a fragment of aftiphilin (residues 1–523) including all of the above listed candidate GAE-binding sequences. This experiment included the NECAP 1-(171–275) and NECAP 2-(170–266) constructs containing DLWGDFS and DIWGDFT sequences, respectively (acidic residues at −2 instead of −1) to also analyze the interaction of these proteins with GAE domains using the yeast two-hybrid system. The results in Fig. 7 demonstrate a strong interaction of aftiphilin (1–523) with the γ1- and γ2-adaptin ears and a weak interaction with GGA1-GAE and GGA3-GAE. In addition, we observed interactions of the NECAP 1 and NECAP 2 yeast two-hybrid constructs with the γ1-adaptin ear and GGA3-GAE (stronger for NECAP 2 than for NECAP 1).

With regard to the motif definition, the results with aftiphilin-(1–523) suggest that the GAE domains (especially the γ-adaptin ears) may also accommodate the presence of Gly at position +4 (present in all of the aftiphilin sequences). This observation, together with the identity of the residues present at +4 in NECAP 1 (Ser), NECAP 2 (Thr), Rabaptin-5 (Val), and in the active γ-synergin and enthoprotin peptides (Ser in both cases) (Table I) led us to conclude that this position in the motif is quite permissive and may accommodate either small hydrophobic residues (Ala or Val), polar but uncharged residues (such as Ser or Thr), or Gly. Finally, the results obtained with the NECAP constructs containing Asp at −2 indicate that the amino-terminal acidic residues in the motif may also be located at this position.

DISCUSSION

Taken together, the data obtained from our peptide binding, yeast two-hybrid, phage display, and GST pull-down assays, along with the analysis of sequences in known GAE interaction partners, indicate that the minimal GAE-binding consensus motif is (W/F/Y)G(P/D/E)(W/F/Y/L/M) or, more succinctly, ΨG(P/D/E)(Ψ/L/M) (where Ψ represents residues with aromatic side chains (Table II). A more comprehensive summary of sequences including the findings regarding positions outside the 0 to +3 motif is presented in Table II (see discussion below). The definition of this motif revises and extends the previously proposed GAE domain-binding sequences. The most salient findings of our study are that (*a*) position 0 of the motif is not restricted to Phe, as previously assumed, but can be occupied by other aromatic residues such as Trp and Tyr, (*b*) Gly is strictly required at position +1 for high affinity binding, (*c*) Pro is permitted, in addition to Asp and Glu, at position +2, and (*d*) position +3 can be occupied by any of several bulky hydrophobic residues (*i.e.* Leu, Phe, Trp, Met, and Tyr). Other important features revealed by this study are the frequent occurrence of acidic residues at positions −3 to −1 and the preference for certain residues at position +4. The sections below discuss the specific features of the motif.

Importance of Acidic Residues at Positions −1 to −3—Most of the GAE-binding sequences characterized to date contain

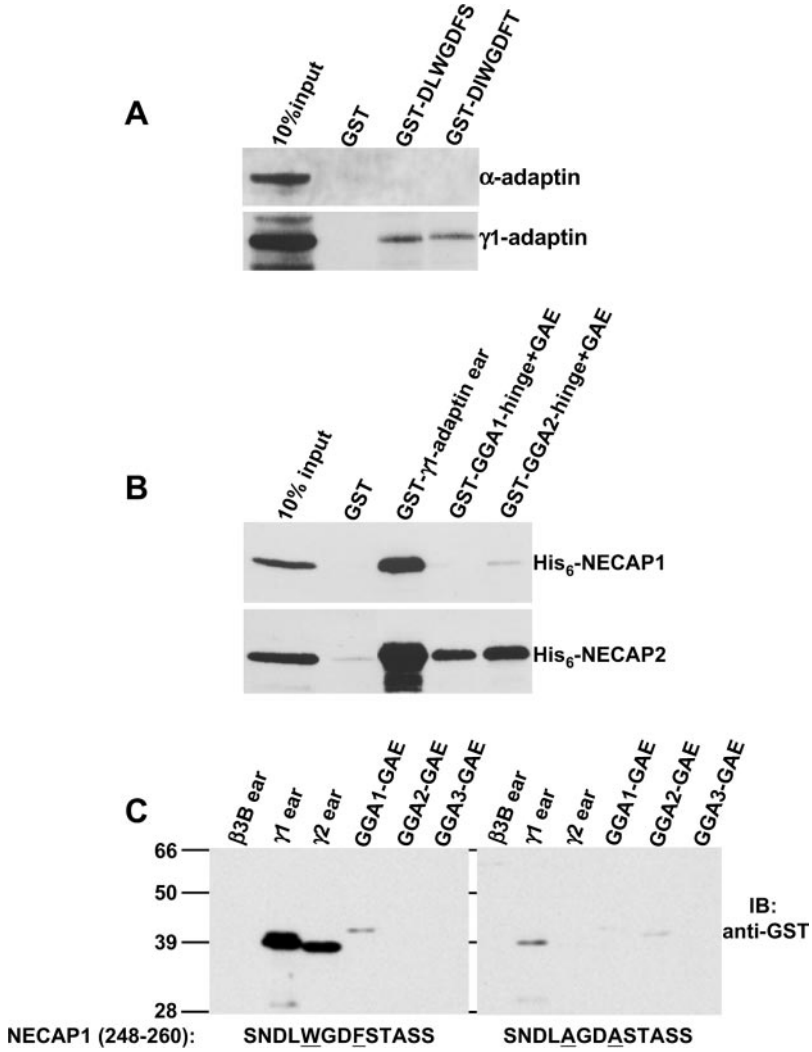

FIG. 5. ***In vitro* assays of the interaction between GAE domains and NECAP proteins.** *A*, GST fusions to the −2 to +4 sequences in the NECAP 1 and NECAP 2 motifs (DLWGDFS and DIWGDFT, respectively) were immobilized on glutathione-Sepharose and incubated with brain extracts. Bound proteins were analyzed by immunoblotting with anti-α-adaptin or anti-γ1-adaptin antisera. Control lanes show the pull-down by GST and the signal in 10% of the brain extracts (10% input). *B*, GST fusions to the hinge + GAE domains of GGA1 and GGA2 were immobilized and incubated with His_6–tagged NECAP proteins. Bound proteins were eluted and detected by immunoblotting with anti-His_6 antiserum. Control lanes are as indicated for *A*. *C*, binding of GST-GAE fusions to the biotinylated peptide SNDLWGDFSTASS (residues 248–260 in NECAP 1). Experiments were performed as described in the legend to Fig. 1.

acidic residues at positions −1 to −3 relative to Ψ (position 0). In a previous study, we showed that acidic residues preceding the Phe at 0 in the Rabaptin-5 motif are not critical for the interaction of a large fragment of this protein (residues 5–476) with the GAE domains but are quantitatively important for the interaction of smaller Rabaptin-5 fragments (residues 406–476 or 428–455) (25). The relatively minor contribution of these acidic residues is in agreement with crystallographic analyses, which showed only a solvent-exposed salt bridge and a main-chain interaction between the Asp at −1 in the Rabaptin-5 peptide and Lys^{650} of the GGA3-GAE domain (26) and a single main-chain hydrogen bond between the Asp at −1 in the p56 peptide and the GGA1-GAE domain (27). The limited nature of these interactions led Collins *et al.* (27) to propose that the upstream acidic residues may act mainly through nonspecific electrostatic interactions with basic residues surrounding the recognition site on the GAE domains and by destabilizing potential secondary structures in the accessory proteins and increasing their accessibility.

The GAE-binding peptides selected from the phage display library (Table III) contained acidic residues at one or two positions (in 80 and 11% of the total peptides isolated, respectively) preceding Trp at 0 or Phe at 0. The exact position of these residues within the peptide varied, however, ranging from −3 to −1. In this context, the GAE-binding sequences from NECAP 1, NECAP 2, enthoprotin-(340–352), and epsin-related proteins from *Caenorhabditis elegans* and *Schizosaccharomyces pombe* all have Asp at −2 followed by either Leu, Ile, Gly, or Ala at −1. This indicates that the acidic amino acid residues can be separated from the Ψ (position 0) residue by at least one other residue and that this may be sufficient for the stabilizing effect on the other critical interactions with GAE domains.

The Anchoring Residues at Positions 0 and +3—The two hydrophobic anchoring residues of the GAE-binding sequences from Rabaptin-5 and p56 bind to hydrophobic pockets on the GAE domain (26, 27). We have found that the requirements for binding to each pocket, however, are different. The aromatic residues Phe, Trp, and Tyr are all functional at position 0, whereas the aliphatic hydrophobic residues Val, Ile, Leu, and Met are not (Fig. 3, *A* and *B*). In contrast, Leu, Met, Trp, Tyr, or Phe, although not Ile, are all active at position +3. These differences in the residues allowed at the 0 and +3-positions can be explained by the structure of the pockets in the GGA-GAE domains. The Phe at 0 in the Rabaptin-5 and p56 peptides participates in a stacking interaction with the guanidino group of Arg^{693} in GGA3 (26) and Arg^{609} in GGA1 (27), respectively. The nature of this interaction probably restricts this position to aromatic residues. On the other hand, the binding pocket for the residue at position +3 is more open and does not contain stacking residues, which makes it capable of accommodating a wider range of hydrophobic residues.

Requirement for Gly at Position +1—One of the most remarkable features of the GAE-binding motif is the critical importance of Gly at position +1. All of the γ-synergin and enthoprotin peptides that bind to GAE domains in our *in vitro*

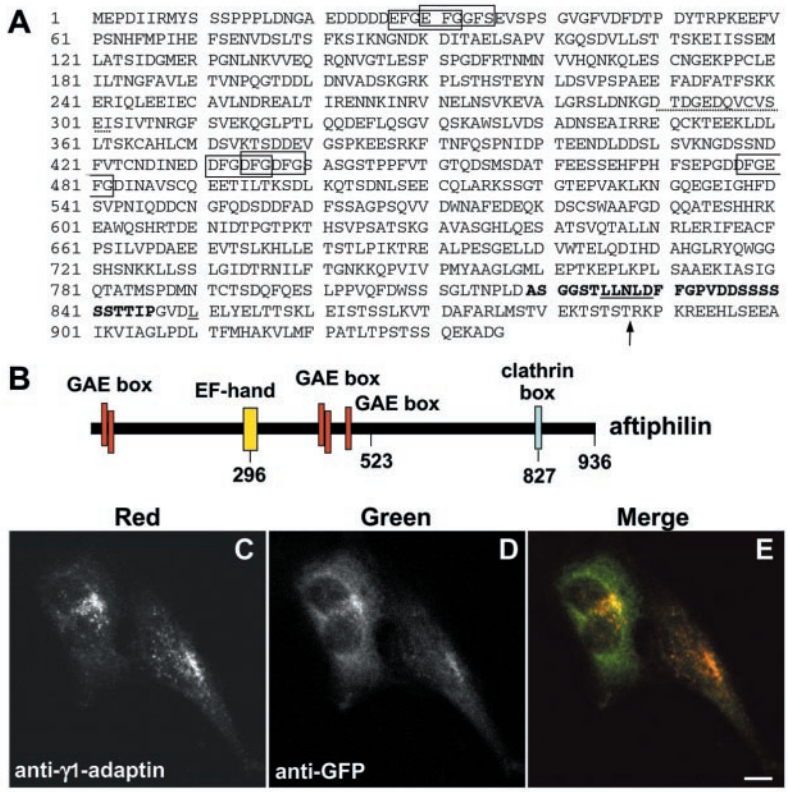

FIG. 6. **Identification of a GAE domain partner.** *A*, sequence of human aftiphilin highlighting the GAE-binding motifs (*boxed*), the putative EF-hand Ca^{2+}-binding domain (*dotted lines*), the 28-amino acid insert not present in FLJ20080 (gi 31377758 ref NP_060127.2) (boldface type), and the clathrin-binding motif (*boldface* and *underlined*). Other differences with FLJ20080 include the substitution of Pro^{822} by Leu^{850} (*underlined*) and the deletion of Ser^{860} (*arrow*). The cDNA clone encoding aftiphilin was isolated from a human brain cDNA library. *B*, schematic representation of aftiphilin indicating the position of relevant motifs. *C–E*, confocal immunofluorescence microscopy of HeLa cells transfected with GFP-aftiphilin. Fixed cells were subjected to double immunostaining using mouse monoclonal anti-γ1-adaptin and rabbit anti-GFP followed by Alexa 568-conjugated anti-mouse and Alexa 488-conjugated anti-rabbit antisera. *Scale bar*, 10 μm.

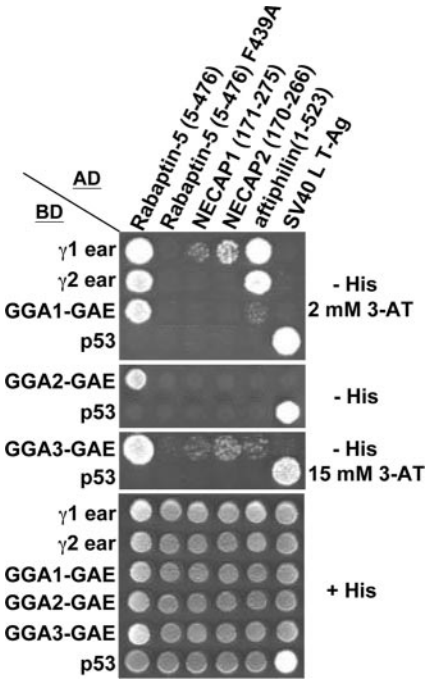

FIG. 7. **Yeast two-hybrid analysis of the interaction of GAE domains with NECAP and aftiphilin.** Constructs encoding the indicated fragments of NECAP 1, NECAP 2, and aftiphilin were subcloned into the pGAD-424 vector and assayed for interactions with the BD-GAE fusions. The AD fusions to Rabaptin-5-(5–476) and to an F439A mutant of this fragment were used as positive and negative controls, respectively.

binding experiments contain Gly at +1 (Table I). Gly at +1 was also present in 59 of 63 peptides selected in the screening of the phage display library with human GAE domains, being replaced by Ala, Val, Ser, or Thr in the remaining four. In addition, Gly at +1 was also present in the GAE-binding sequences of the newly identified partners NECAP 1, NECAP 2, and aftiphilin. Finally, substitution of Gly^{440} in Rabaptin-5 (position +1 of the motif) by other residues markedly decreased interactions with GAE domains. The G440A substitution, used as a control in the present study, warrants further discussion. Consistent with our previous results using a larger Rabaptin-5 template, the G440A substitution in the context of the Rabaptin-5-(406–476) fragment abolished the interactions with GAE domains, except for a weak binding to GGA3-GAE (Fig. 4). These results indicate that although Ala may be permitted at +1 in some contexts, it is clearly less favorable than Gly.

The critical role of the Gly at +1 is also supported by the structure of the liganded GGA-GAE domain. Although the Gly at +1 in the Rabaptin-5 peptide is only involved in main-chain interactions with the GGA3-GAE domain, this residue allows a kink in the peptide that shifts the register of the peptide β-strand. This positions the side chains of the Phe and Leu at 0 and +3, respectively, on the same side of the strand such that they can bind to their corresponding pockets on the GAE domain (26).

It should be noted that others have measured a K_D of 13–15 μM for the binding of the γ1-adaptin ear domain to a γ-synergin peptide containing Glu at +1 (peptide P1 corresponding to residues 453–464 of human γ-synergin in Ref. 23; see also Ref. 27). *Saccharomyces cerevisiae* Ent3p contains a GAE-binding sequence with Ser (residue 273) at +1, in addition to a canonical sequence with Gly (residue 380) at +1. One of the peptides selected in our phage display screen also contained Ser at +1. Thus, Glu and Ser can substitute for Gly at position +1, at least in some sequence contexts and for some GAE domains. Indeed, the reported interaction of the GAE domains of yeast Gga2p and γ-adaptin with the DEEEDDDDEF sequence from Ent3p (24) hints at a possible difference in the mechanism of recognition of yeast and mammalian GAE domains. This sequence consists of a long string of acidic residues followed by a Phe but lacks other critical elements of the canonical ΨG(P/D/E)(Ψ/L/

M). It will be of interest to determine whether this sequence binds to the same site on the GAE domain defined by the crystallographic studies.

Pro, Asp, and Glu Are Favored at Position +2—The substitution analysis in the Rabaptin-5 fragment showed that, although Pro was the most favorable residue at position +2, Asp and Glu were also active (Fig. 4). This finding is consistent with the presence of acidic residues at the corresponding positions in the active γ-synergin and enthoprotin peptides (Table I) and in the NECAP 1, NECAP 2, and aftiphilin sequences (Figs. 5 and 6). Likewise, the peptide library screening also yielded a majority (48 of 63) of peptides with either Pro or Glu at +2. All of the phage-displayed peptides selected with the GAE domains of γ1-adaptin, GGA1, and GGA2, as well as the two most frequent binders to the γ2-adaptin GAE domain, fit this pattern. The GGA3-GAE domain exhibited a more heterogeneous pattern, including a frequent binder with Ala at +2. Only one of the peptides selected with the γ2-adaptin ear domain had a Gly at +2. The low frequency of Gly at +2 is consistent with the finding that its substitution by Glu resulted in a 5-fold increase in the affinity of interactions of the p56 peptide with the GGA1 and γ1-adaptin GAE domains (27). This finding was interpreted to mean that the restricted main-chain torsion angles of acidic residues would favor binding through a reduction of the flexibility associated with a Gly at +2 (27). We concur with this interpretation in light of the strong interactions that we observe when the +2-position is occupied by the conformationally restricted Pro.

Contribution of Residues at Position +4—We have not directly examined the effect of substitutions at position +4, due to the fact that the V443A mutation in Rabaptin-5-(5–476) results in a partially active sequence (25). However, the presence of a Val at +4 in the Rabaptin-5 motif and in the majority of peptides selected in the phage display screening (25 of the 37 that do not contain Cys) indicates that Val is indeed favored at this position. This is consistent with crystallographic data showing that the Val at +4 in the Rabaptin-5 peptide displays hydrophobic interactions with Val^{654} and the aliphatic moiety of Lys^{655} of the GGA3-GAE domain (26). Ser also appears to be allowed at +4, based on its presence in the active peptides derived from γ-synergin, enthoprotin, and NECAP 1 (Table I). Other residues that occur at +4 include Thr in the NECAP 2 peptide (Fig. 6*A*), Gly in the multiple peptides from aftiphilin, and Gln in Ent3p (residue 276) and Ent5p (residue 335) as well as in the P1-γ-synergin peptide used by Mills *et al.* (23). Finally, Ala should also be considered an active +4 residue in the motif because of the interaction of the GAE domains of γ1-adaptin and GGA3 with the Rabaptin-5 V443A mutant (25). At present, we do not know whether residues other than Val at +4 contribute to the strength of peptide-GAE domain interactions or are simply permissive for interactions mediated by other residues. It is difficult to identify additional patterns at +4 other than the above listed hydrophobic or uncharged residues, and, whereas their presence supports the likelihood of a positive interaction with GAE domains, there is not enough evidence at this time to include this position in the core consensus motif.

The Problem of GAE Domain Selectivity—A problem that remains unresolved is the basis for the preferential recognition of certain peptide motifs by specific GAE domains. In contrast to the Rabaptin-5 peptide, which interacted strongly with all GAE domains tested (with the exception of GGA2-GAE), the γ-synergin, enthoprotin, and NECAP 1 peptides displayed preferential interactions with the γ1- and γ2-adaptin ear domains *in vitro* (Figs. 1 and 2). These observations are consistent with the co-localization of γ-synergin with γ1-adaptin, and with the dependence of γ-synergin membrane association on AP-1 (15). A similar preference of enthoprotin for the γ1-adaptin ear has also been reported (22, 23). However, replacement of the FGPL sequence in Rabaptin-5-(406–476) by the FGEF or FGDW sequences of γ-synergin-(666–678) or enthoprotin-(368–380), respectively, did not confer preferential recognition by the γ-adaptin ear domains in yeast two-hybrid assays (Fig. 4, *A* and *B*). These results indicate that residues outside the 0 to +3 box may determine selective recognition by the GAE domains. In this regard, whereas a relatively short (10-mer) biotinylated peptide ending at the +5-position of the γ-synergin DFGEFS motif (positions 666–675; Fig. 1) only bound to the γ2-adaptin ear, a longer (13-mer) peptide including additional carboxyl-terminal amino acids (residues 666–678) exhibited a noticeable, albeit weak, interaction with GGA1-GAE (Fig. 2). The role of residues carboxyl-terminal to the core tetrapeptide motif in fine-tuning recognition by different GAE domains is supported by the observed interactions of the Ala at +6 in the Rabaptin-5 motif with two Pro rings in the GAE domain of GGA3 (26).

Another potential determinant of avidity or selectivity of peptide recognition by GAE domains is the presence of closely spaced or overlapping motifs. Whereas multiple motifs could enhance the avidity of interactions by increasing the local concentration of ligand, the presence of overlapping sequences may result in generation of weaker ligands due to inadequate positioning of residues outside the 0 to +3 box. In addition, by presenting several variants of the motif to the GAE domains, they could broaden the specificity of recognition. Examples of these multiple motifs are the SADLFGGFADFGS sequence of enthoprotin (residues 340–352) and the DFGEFSLFGEYS sequence of γ-synergin (residues 669–680). The enthoprotin 340–352 peptide encompasses a canonical motif defined by Phe^{344} and Phe^{347} at 0 and +3 as well as a second sequence that fits the consensus at positions 0 (Phe^{347}), +2 (Asp^{349}), and +3 (Phe^{350}), although not at position +1 (Ala^{348}) (this last residue would be expected to result in a low affinity interaction). We found that this biotinylated peptide bound not only the γ-adaptin ear, as observed for another enthoprotin peptide (residues 368–380) containing a canonical DFGEFS sequence, but also to the GGA2-GAE. This indicates that the presence of overlapping motifs may modify the selectivity of recognition by GAE domains.

Comparison with the YXXØ Motif—The recognition of the GAE-binding motif can be considered a variation of the "two-pronged" mechanism exemplified by the binding of Y*XX*Ø signals to the μ2 subunit of AP-2 (33). Both interactions involve the binding of peptides in an extended conformation and the anchoring of bulky hydrophobic residues at positions 0 and +3 into hydrophobic pockets. Another similarity is the preference for Pro at position +2, which probably reflects the stabilization of conformations favorable for the binding of the anchoring residues (Fig. 4 in this study) (34, 35). In addition, the backbones of both motifs complement β-strands in their corresponding recognition modules. However, the details of the interactions are different. The hydroxyl group of the Tyr in the Y*XX*Ø motif establishes a network of hydrogen bonds with several residues located at the base of its pocket. Although other hydrophobic residues such as Phe could fit into this pocket, they would not recreate the hydrogen bonding of the Tyr. This explains why only Tyr is accepted at position 0 in the Y*XX*Ø motif. In contrast, the pocket for residue 0 of the ΨG(P/D/E)(Ψ/L/M) motif can accommodate not only Phe but also Tyr or Trp. The pocket for the +3 residue is also deeper in μ2 than in the GAE domains. In both cases, however, these pockets can accommodate a range of hydrophobic residues at +3. Finally, the activity of both the Y*XX*Ø and ΨG(P/D/E)(Ψ/L/M) motifs is

dependent on their sequence context. In the case of YXXØ motifs, another hydrophobic residue amino-terminal to the Tyr at 0 can contact μ2 at another site to establish a "three-pronged" interaction (36). It would be interesting to examine whether overlapping ΨG(P/D/E)(Ψ/L/M) motifs can engage in a similar type of interaction.

Conclusion—The validity of the GAE-binding consensus motif here defined has been demonstrated by the identification of the novel GAE-binding partners NECAP and aftiphilin. These proteins also contain clathrin-binding motifs and are thus likely to be physiologically relevant partners for AP-1 and the GGAs. Future studies will be required to elucidate the regulatory implications of the interaction of NECAP and aftiphilin with the adaptors.

Acknowledgments—We are grateful to J. Hurley for critical reading of the manuscript and to J. Philie, A. San Miguel, and X. Zhu for expert technical assistance.

REFERENCES

1. Kirchhausen, T. (2000) *Annu. Rev. Biochem.* **69,** 699–727
2. Brodsky, F. M., Chen, C. Y., Knuehl, C., Towler, M. C., and Wakeham, D. E. (2001) *Annu. Rev. Cell Dev. Biol.* **17,** 517–568
3. Kirchhausen, T. (2002) *Cell* **109,** 413–416
4. Robinson, M. S., and Bonifacino, J. S. (2001) *Curr. Opin. Cell Biol.* **13,** 444–453
5. Robinson, M. S. (1990) *J. Cell Biol.* **111,** 2319–2326
6. Lewin, D. A., Sheff, D., Ooi, C. E., Whitney, J. A., Yamamoto, E., Chicione, L. M., Webster, P., Bonifacino, J. S., and Mellman, I. (1998) *FEBS Lett.* **435,** 263–268
7. Slepnev, V. I., and De Camilli, P. (2000) *Nat. Rev. Neurosci.* **1,** 161–172
8. Benmerah, A., Begue, B., Dautry-Varsat, A., and Cerf-Bensussan, N. (1996) *J. Biol. Chem.* **271,** 12111–12116
9. Brett, T. J., Traub, L. M., and Fremont, D. H. (2002) *Structure* **10,** 797–809
10. Owen, D. J., Vallis, Y., Noble, M. E., Hunter, J. B., Dafforn, T. R., Evans, P. R., and McMahon, H. T. (1999) *Cell* **97,** 805–815
11. Traub, L. M., Downs, M. A., Westrich, J. L., and Fremont, D. H. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 8907–8912
12. Owen, D. J., Vallis, Y., Pearse, B. M., McMahon, H. T., and Evans, P. R. (2000) *EMBO J.* **19,** 4216–4227
13. Kent, H. M., McMahon, H. T., Evans, P. R., Benmerah, A., and Owen, D. J. (2002) *Structure* **10,** 1139–1148
14. Nogi, T., Shiba, Y., Kawasaki, M., Shiba, T., Matsugaki, N., Igarashi, N., Suzuki, M., Kato, R., Takatsu, H., Nakayama, K., and Wakatsuki, S. (2002) *Nat. Struct. Biol.* **9,** 527–531
15. Lui, W. W., Collins, B. M., Hirst, J., Motley, A., Millar, C., Schu, P., Owen, D. J., and Robinson, M. S. (2003) *Mol. Biol. Cell* **14,** 2385–2398
16. Hirst, J., Lui, W. W., Bright, N. A., Totty, N., Seaman, M. N., and Robinson, M. S. (2000) *J. Cell Biol.* **149,** 67–80
17. Zhu, Y., Doray, B., Poussu, A., Lehto, V. P., and Kornfeld, S. (2001) *Science* **292,** 1716–1718
18. Page, L. J., Sowerby, P. J., Lui, W. W., and Robinson, M. S. (1999) *J. Cell Biol.* **146,** 993–1004
19. Takatsu, H., Yoshino, K., and Nakayama, K. (2000) *Biochem. Biophys. Res. Commun.* **271,** 719–725
20. Wasiak, S., Legendre-Guillemin, V., Puertollano, R., Blondeau, F., Girard, M., de Heuvel, E., Boismenu, D., Bell, A. W., Bonifacino, J. S., and McPherson, P. S. (2002) *J. Cell Biol.* **158,** 855–862
21. Kalthoff, C., Groos, S., Kohl, R., Mahrhold, S., and Ungewickell, E. J. (2002) *Mol. Biol. Cell* **13,** 4060–4073
22. Hirst, J., Motley, A., Harasaki, K., Peak Chew, S. Y., and Robinson, M. S. (2003) *Mol. Biol. Cell* **14,** 625–641
23. Mills, I. G., Praefcke, G. J., Vallis, Y., Peter, B. J., Olesen, L. E., Gallop, J. L., Butler, P. J., Evans, P. R., and McMahon, H. T. (2003) *J. Cell Biol.* **160,** 213–222
24. Duncan, M. C., Costaguta, G., and Payne, G. S. (2003) *Nat. Cell Biol.* **5,** 77–81
25. Mattera, R., Arighi, C. N., Lodge, R., Zerial, M., and Bonifacino, J. S. (2003) *EMBO J.* **22,** 78–88
26. Miller, G. J., Mattera, R., Bonifacino, J. S., and Hurley, J. H. (2003) *Nat. Struct. Biol.* **10,** 599–606
27. Collins, B. M., Praefcke, G. J., Robinson, M. S., and Owen, D. J. (2003) *Nat. Struct. Biol.* **10,** 607–613
28. Duncan, M. C., and Payne, G. S. (2003) *Trends Cell Biol.* **13,** 211–215
29. Dell'Angelica, E. C., Klumperman, J., Stoorvogel, W., and Bonifacino, J. S. (1998) *Science* **280,** 431–434
30. Sidhu, S. S., Lowman, H. B., Cunningham, B. C., and Wells, J. A. (2000) *Methods Enzymol.* **328,** 333–363
31. Deshayes, K., Schaffer, M. L., Skelton, N. J., Nakamura, G. R., Kadkhodayan, S., and Sidhu, S. S. (2002) *Chem. Biol.* **9,** 495–505
32. Ritter, B., Philie, J., Girard, M., Tung, E. C., Blondeau, F., and McPherson, P. S. (2003) *EMBO Rep.* **4,** 1089–1093
33. Owen, D. J., and Evans, P. R. (1998) *Science* **282,** 1327–1332
34. Ohno, H., Aguilar, R. C., Yeh, D., Taura, D., Saito, T., and Bonifacino, J. S. (1998) *J. Biol. Chem.* **273,** 25915–25921
35. Aguilar, R. C., Boehm, M., Gorshkova, I., Crouch, R. J., Tomita, K., Saito, T., Ohno, H., and Bonifacino, J. S. (2001) *J. Biol. Chem.* **276,** 13145–13152
36. Owen, D. J., Setiadi, H., Evans, P. R., McEver, R. P., and Green, S. A. (2001) *Traffic* **2,** 105–110