CHAPTER II: Structural Basis for the Accessory Protein Recruitment by the γ-Adaptin Ear Domain

ABSTRACT

The adaptor proteins AP-1 and GGA regulate membrane traffic between the TGN and endosomes/lysosomes through ARF-regulated membrane association, recognition of sorting signals, and recruitment of clathrin and accessory proteins. The γ1-adaptin subunits of AP-1 and GGA possess homologous ear domains involved in the recruitment of accessory proteins, γ-synergin and Rabaptin-5. Here, by collaborating with Wakatsuki and colleagues, I report the crystal structure of the human γ1-adaptin ear domain, which is the first three-dimensional structure determined in the AP-1 subunits. The crystal structure of the human γ1-adaptin ear domain consists solely of an immunoglobulin-like fold, unlike the α-adaptin ear domain. Structure-based mutational analyses reveal a binding site for the accessory proteins that is composed of conserved basic residues, indicating that the recruitment mechanism in γ1-adaptin and GGA is distinct from that in α-adaptin.

INTRODUCTION

Clathrin-mediated membrane traffic is responsible for the protein transport from the plasma membrane and TGN to the endosomal/lysosomal system. Membrane recruitment of clathrin is mediated by heterotetrameric AP complexes, which interact with cargo proteins. AP-1 is one of the AP complexes composed of γ (γ1)-, β1-, μ1- and σ1-adaptins. The C-terminal protruding part of γ1-adaptin, often referred to as the ear domain, is involved in the recruitment of accessory proteins, such as γ-synergin, Rabaptin-5, and cyclin G-associated kinase, which modulate the functions of AP-1 in the membrane trafficking events (35, 26, 80, CHAPTER I). Similar to endocytic accessory proteins for α-adaptin of the AP-2 complex, γ-synergin possesses an Eps15 (a phosphorylation substrate of the epidermal growth factor receptor kinase) homology domain and interacts with Asn-Pro-Phe motifs of secretory carrier membrane protein 1, which was suggested to serve as an assembly site for the clathrin coat on the donor membrane (81). On the other hand, Rabaptin-5 is involved in the membrane fusion of transport vesicles at the target membranes.

Although, until recently, the transport of mannose 6-phoshate-modified lysosomal hydrolases from the TGN to endosomes/lysosomes have been believed to be mediated by AP-1, a recent study has suggested that AP-1 is rather responsible for retrograde endosome-to-TGN transport (32). On the other hand, a novel family of adaptor proteins possessing a γ ear-like domain, termed GGAs, have been implicated in the anterograde transport from the TGN (33-37, 82). GGAs are monomeric proteins having four functional domains, termed VHS, GAT, hinge, and GAE, from the N-terminus in order (see Fig. 4 in GENERAL INTRODUCTION). The C-terminal GAE domain seems to share a common function with the γ1 ear domain, that is, the recruitment of accessory proteins. The GAE domain interacts with γ-synergin and Rabaptin-5 (35, 37), suggesting the general importance of the γ ear-like

domain in the membrane traffic between the TGN and endosomes/lysosomes. On the other hand, the N-terminal VHS domain recognizes the ACLL sequence of cargo receptors, such as MPRs and sortilin (39-42). The signal recognition mechanism of the VHS domain has already been demonstrated by X-ray crystallography and biochemical analyses (83, 84).

Here, by collaborating with Wakatsuki and colleagues, I report the crystal structure of the human γ1-adaptin ear domain at 1.8 Å resolution, which is the first three-dimensional structure determined in the AP-1 subunits. Based on the resulting structure, I performed mutational analyses to identify the binding site for accessory proteins. The results revealed a novel mechanism of the accessory protein recruitment through the γ1 ear and GAE domains, distinct from the previously proposed model for the α-adaptin ear domain in endocytosis.

MATERIALS AND METHODS

Protein expression and purification — The DNA fragment for residues 677-822 of human γ1-adaptin was cloned into the pGEX4T-2 plasmid. The native and selenomethionine (SeMet)-substituted protein were expressed in *E. coli* BL21 and DL41 cells, respectively. An addition of preceding residues (677-702) to the γ1 ear domain (residues 703-822) improved the solubility of the protein. The GST fusion protein was purified through affinity chromatography using a glutathione-Sepharose 4B column, and cleaved by thrombin. The cleaved γ1 ear domain was further purified by Superdex 75 size-exclusion column in 100 mM NaCl and 20 mM Tris-HCl (pH 8.0).

Crystallization and data collection — Crystals of the native and SeMet-substituted proteins were obtained in a buffer solution containing 10 mg/ml protein, 20 % (wt./vol.) PEG 4000, 200 mM $MgCl_2$, and 100 mM HEPES-Na pH 7.5 with equilibration at 277 K for a week. All data sets were collected under cryogenic conditions with crystals soaked in the cryoprotectant buffer containing 20 % (wt./vol.) glycerol and cooled at 100 K in a nitrogen gas stream. Gold derivatives were prepared by soaking crystals in the cryoprotectant buffer with 10 mM $KAu(CN)_2$. Native and gold derivative data sets were collected at BL-44XU of SPring-8, and MAD data sets of the SeMet-substituted crystal were collected at BL-6A of Photon Factory. The diffraction data were integrated and scaled using the programs MOSFLM (85) and SCALA (86).

Structure determination and refinement — For SIRAS phasing with the gold derivative data, initial phases were determined using the program MLPHARE (86), and improved with a density modification procedure using the program DM (86). For MAD phasing, initial phases were determined using the program SOLVE (87), and improved using the program RESOLVE (88). The initial molecular model was built automatically using the program ARP/warp (89), which could assign the main chain of 109 out of 147 residues in the γ1 ear

construct.

Further model fitting to the electron density maps was carried out manually using the program O (90), followed by the structure refinement through the program REFMAC (86). Two γ1 ear domain monomers, termed A- and B-monomers, are present in the asymmetric unit of the crystal and are composed of residues 703-822 and 700-822, respectively. No model could be built for the N-terminal hinge region, 677-699, which are 23 residues in total 146 residues, due to disorder in the electron density. It might account for the relatively high *R*-factors; the final values for working and test sets are 22.6 and 24.8 % at 1.8 Å resolution, respectively. Stereochemical quality of the final model was assessed by the program PROCHECK (86), where 88.9 % of amino acid residues were located in the most favored regions and none in the disallowed regions.

Yeast two-hybrid assay — Reverse two-hybrid screening was performed according to the previously reported procedure (43). A cDNA fragment of the γ1 ear domain (residues 703-822) was subjected to error-prone PCR whose fidelity is reduced by the reaction mixture containing 250-500 μM $MnCl_2$ and subcloned into the pGBT9 bait vector. Y190 reporter cells harboring the pGAD10 prey vector for Rabaptin-5 were transformed with the bait vector for the mutagenized γ1 ear domain and subjected to a filter assay for β-galactosidase activity. The pGBT9 plasmids recovered from colonies that did not develop blue color were sequenced.

GST-pull down assay — N-terminally HA-tagged γ-synergin was expressed in hEK-293 cells (37). The supernatant of cell lysate was incubated for 2h at 4℃ with 20 μg of recombinant GST, GST-γ1 ear domain (37) or its mutant prebound to glutathione-Sepharose 4B beads as described in CHAPTER I. The beads were then washed three times with homogenization buffer containing 0.5% Triton X-100. Proteins associated with the beads were subjected to immunoblot analysis using anti-Rabaptin-5 (clone 20, BD Biosciences) or

anti-HA (3F10, Roche Diagnostics) antibody. The band densities were estimated using a LAS-1000 bioimaging analyzer (Fuji Photo Film, Co.).

RESULTS AND DISCUSSION

Structure description — The crystal structure of the human γ1 ear domain was determined by a combination of SIRAS and MAD phasing (Fig. II-1). The resulting structure shows that the γ1 ear domain forms an immunoglobulin-like β-sandwich fold composed of eight β-strands with two short α-helices. Each strand is composed of residues 707-712 in β-1, 715-723 in β-2, 730-739 in β-3, 746-753 in β-4, 759-762 in β-5, 778-785 in β-6, 795-802 in β-7, and 805-812 in β-8, respectively. Two β-sheets of the sandwich structure consist of the strands β-1, 2, 3, 5 and 6, and the strands β-4, 7 and 8, respectively. The helix α-1 is composed of residues 772-774 and located between the strands β-5 and 6, and α-2 is 818-820 after β-8. Topology of the entire γ1 ear domain is quite similar to those of the N-terminal subdomains in the α-and β2-adaptin ear domains of the AP-2 complex (94-96) (Fig. II-2).

A preliminary sequence comparison suggested that the γ1 ear domain is only half as long as the α and β2 ear domains and that it might be an immunoglobulin-like fold. However, it has very low sequence identity (about 10 %) and homology (about 20 %) to the N-terminal immunoglobulin-like subdomains of the α and β2 ear domains. Biochemical analyses of the α and β2 ear domains have indicated that it is not the N-terminal subdomains but rather their C-terminal platform subdomains that are critical for the interaction with accessory proteins (94-96). On the other hand, the result described above confirms that the γ1 ear domain indeed folds into an immunoglobulin-like structure and lacks the C-terminal platform subdomain, which leaves a question of how the γ1 ear domain recruits the accessory proteins.

Conserved feature of γ1-ear and GGA-GAE — The conservation of the β-sandwich structure in the GAE domain can be deduced from a structure-based sequence comparison with the γ1 ear domain (Fig. II-3). Several deletions and insertions are present between the γ1 ear and GAE domains, but all of them are attributed to the loop regions between β-strands. Most of the conserved

hydrophobic residues are buried in the core of the molecule, suggesting that they contribute to the structural integrity of the β-sandwich structure. According to the sequence alignment, the residues 703-822 of the γ1 ear domain, which form a rigid β-sandwich fold in the resulting structure, show a sequence identity of around 34 % and homology of 47 % to the corresponding regions of the GAE domains of human GGAs. Moreover, the GAE domains do not possess additional C-terminal residues downstream of the putative rigid β-sandwich structure. Thus it is highly likely that the GAE domain has a similar immunoglobulin-like domain and lacks the C-terminal platform regions.

A further analysis of the sequence alignment and the X-ray structure reveals a striking feature in the conservation of the charged amino acid residues. The basic residues at the N-terminus of the strand β-7 (Arg 793, Arg 795, and Lys 797 in the human γ1 ear domain) are highly conserved, and they form a cluster on the molecular surface. In addition, a sequence at the C-terminus of the strand β-4, $A^{753}VPK^{756}$, is almost completely conserved in all available γ-adaptin and GGA sequences, and the side chain of Lys 756 is located near the basic cluster. Altogether these conserved residues constitute a large basic surface around the C-terminus of the strand β-4 and the N-terminus of the strand β-7 (Fig. II-2). On the other hand, the N-terminal immunoglobulin-like subdomains of the α and β2 ear domains do not have conserved basic residues in the corresponding surface area, suggesting that the basic cluster is a unique feature of the γ ear-like domain.

Binding site for accessory proteins — To identify the residues interacting with accessory proteins, we carried out a reverse two-hybrid screening of γ1 ear mutants. By this screening, many mutations in the γ1 ear domain were found to abolish the ability to bind Rabaptin-5 (Fig. II-3). Most of the identified mutations were introduced into the hydrophobic or uncharged residues embedded in the core of the γ1 ear molecule, namely into the residues that would not participate in the molecular interaction. These mutations might

disrupt the β-sandwich structure, which would abolish the binding of accessory proteins altogether.

More interestingly, the screening highlighted the importance of charged residues in binding to Rabaptin-5. The identified residues are Arg 793, Lys 797 and Glu 812. Arg 793 and Lys 797 are the constituents of the conserved basic cluster, and Glu 812 is located near the cluster. We therefore introduced single mutations to all of the basic residues in the cluster (K756Q, R793Q, R795Q, and K797Q), and performed a GST pull down assay, on an assumption that the conserved basic surface is involved in the interaction with accessory proteins. We also subjected Ala 753 (A753Q) and Glu 812 (E812K) mutants to the pull down assay; the former is the first residue of the highly conserved sequence, $A^{753}VPK^{756}$, and is surrounded by the three basic residues in the cluster, Lys 756, Arg 795, and Lys 797.

As shown in Fig. II-4, the γ1 ear domain could no longer bind to γ-synergin if either of the basic residues was mutated, indicating the importance of the basic surface in binding to the accessory protein. Furthermore, the E812K mutation abolished the binding to γ-synergin. Glu 812 interacts with the proximal basic residues, Arg 793 and Arg 795, via salt bridges in the present crystal structure, suggesting that the E812K mutation perturb the spatial arrangement of the basic residues in the binding surface. The A753Q mutation moderately affected the binding. With the small side chain, Ala 753 might increase the accessibility of the binding surface to the accessory protein. Therefore, the bulky side chain of the glutamine residue might perturb the interaction at least in the case of γ-synergin. On the other hand, the previous binding analyses have shown that the middle portion of human γ-synergin (residues 518-786) directly interacts with the γ1 ear domain (79). γ-Synergin possesses five acidic sequences DDFxD/EF (x represents any amino acid), three of which are located in the γ1 ear-binding region. Provided that the acidic sequence serves as a γ1 ear-binding motif, the basic surface of the γ1 ear

domain would be suitable for the signal recognition through electrostatic interactions.

Essentially the same results were obtained for Rabaptin-5, except for the mutations A753Q and R795Q. The R795Q mutation had a marginal effect on the binding to Rabaptin-5, even though Arg 795 is located at the center of the basic cluster. My two-hybrid analysis in CHAPTER I has revealed that the γ1 ear domain interacts with the C-terminal coiled-coil region of Rabaptin-5. However, no acidic sequence like DDFxD/EF in γ-synergin is identified in this region. Although the binding mode of Rabaptin-5 seems different from that of γ-synergin, the above results strongly indicate that the basic cluster is critical for interaction in both cases. In contrast, the previous structural analyses of the α-adaptin ear domain showed that the N-terminal immunoglobulin-like subdomain only serves as the spacer between the hinge region and the C-terminal platform subdomain. Furthermore, the α ear domain possesses a shallow hydrophobic pocket between the β-sheet and the sheet-crossing α-helix on the top of the C-terminal subdomain (Fig II-2), which recognizes the DPF/W motifs of endocytic accessory proteins (94). Taken together, these findings indicate that γ1-adaptin and GGA interact with the accessory proteins in a manner different from that observed in α-adaptin of the AP-2 complex. Ultimate confirmation of the recognition mechanism of the accessory proteins awaits the crystal structure of the complex between the γ1 ear domain and γ-synergin or Rabaptin-5. Nevertheless, the combination of the γ1 ear structure and the structure-based mutational analyses strongly suggest a novel mechanism in which γ1-adaptin and GGA recruit the accessory proteins through the basic cluster of the common immunoglobulin-like ear domain.

FIGURES

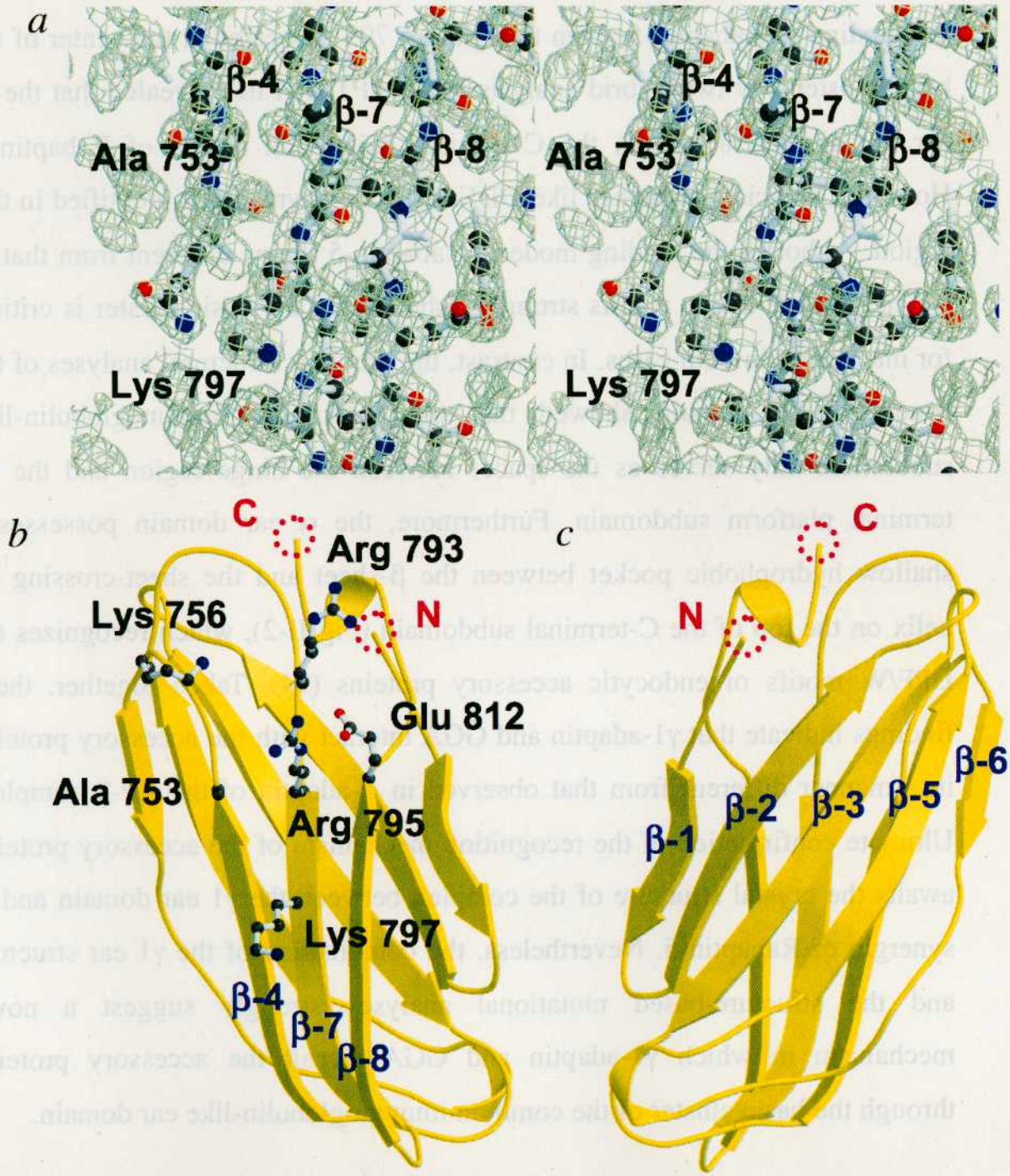

a
β-4
β-7
β-8
Ala 753
Lys 797
β-4
β-7
β-8
Ala 753
Lys 797
b
C
Arg 793
Lys 756
N
Glu 812
Ala 753
Arg 795
Lys 797
β-4
β-7
β-8
c
C
N
β-1
β-2
β-3
β-5
β-6

Fig. II-1. Structure of the human γ1-ear domain. *a* Stereo-view of an initial electron density map from the SIRAS phasing. The electron density map around Ala 753 and Lys 797 is contoured at 1.0 s. The refined model of the γ1 ear domain is superimposed on the map. Water molecules are excluded for clarity. *b, c* Ribbon diagram of the human γ1-adaptin ear domain. The immunoglobulin-like β-sandwich fold of the γ1 ear domain is composed of eight β-sheets. Two β-sheets of the sandwich folds are composed of the strands β-4, 7, and 8 in *b* and of the strands β-1, 2, 3, 5, and 6 in *c*, respectively. The key residues in the accessory protein recruitment are highlighted with the ball-and-stick models. The N- and C-termini of each structure are indicated by the dotted circles. Fig. II-1 and Fig. II-2 were prepared using MOLSCRIPT (92).

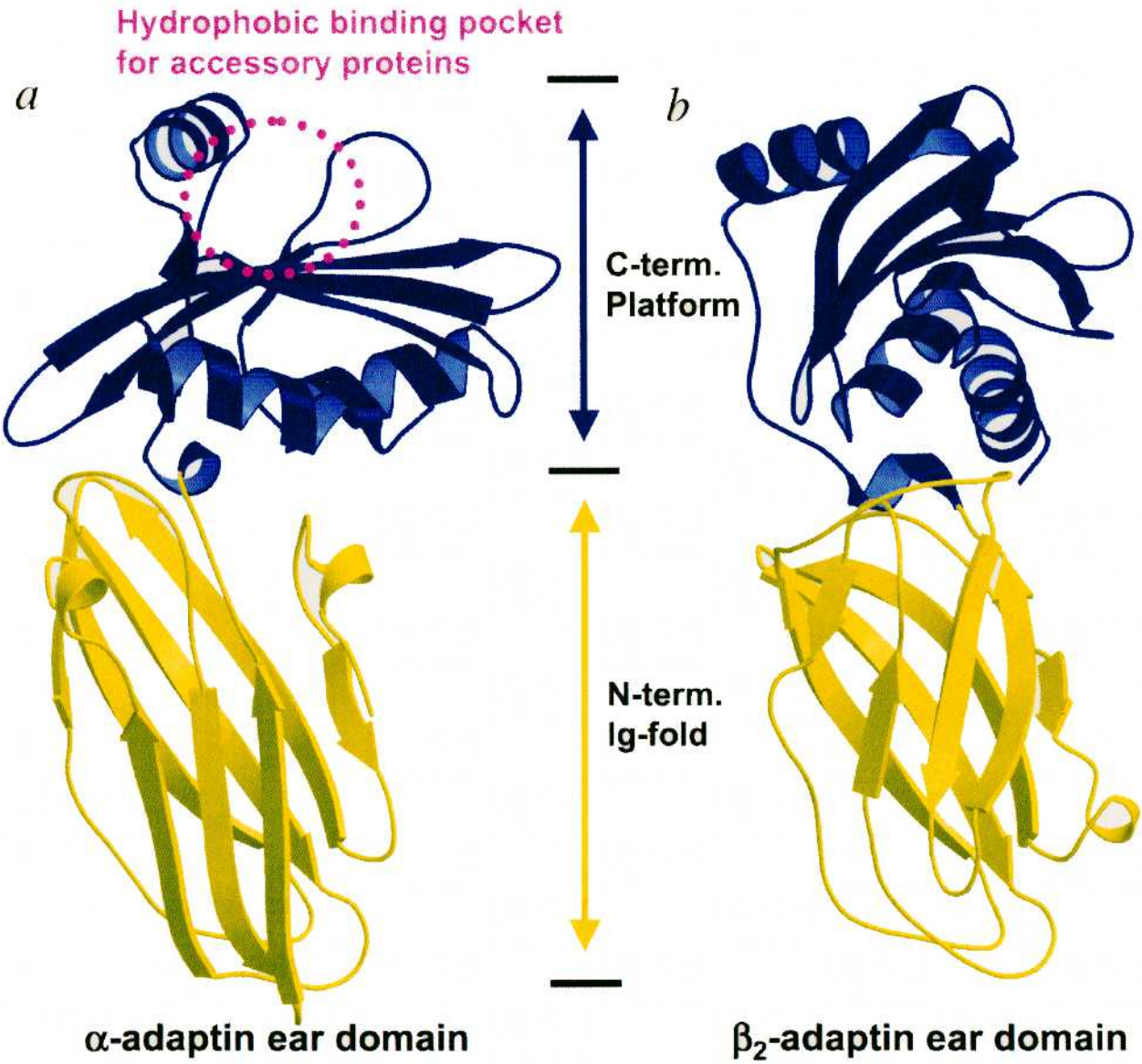
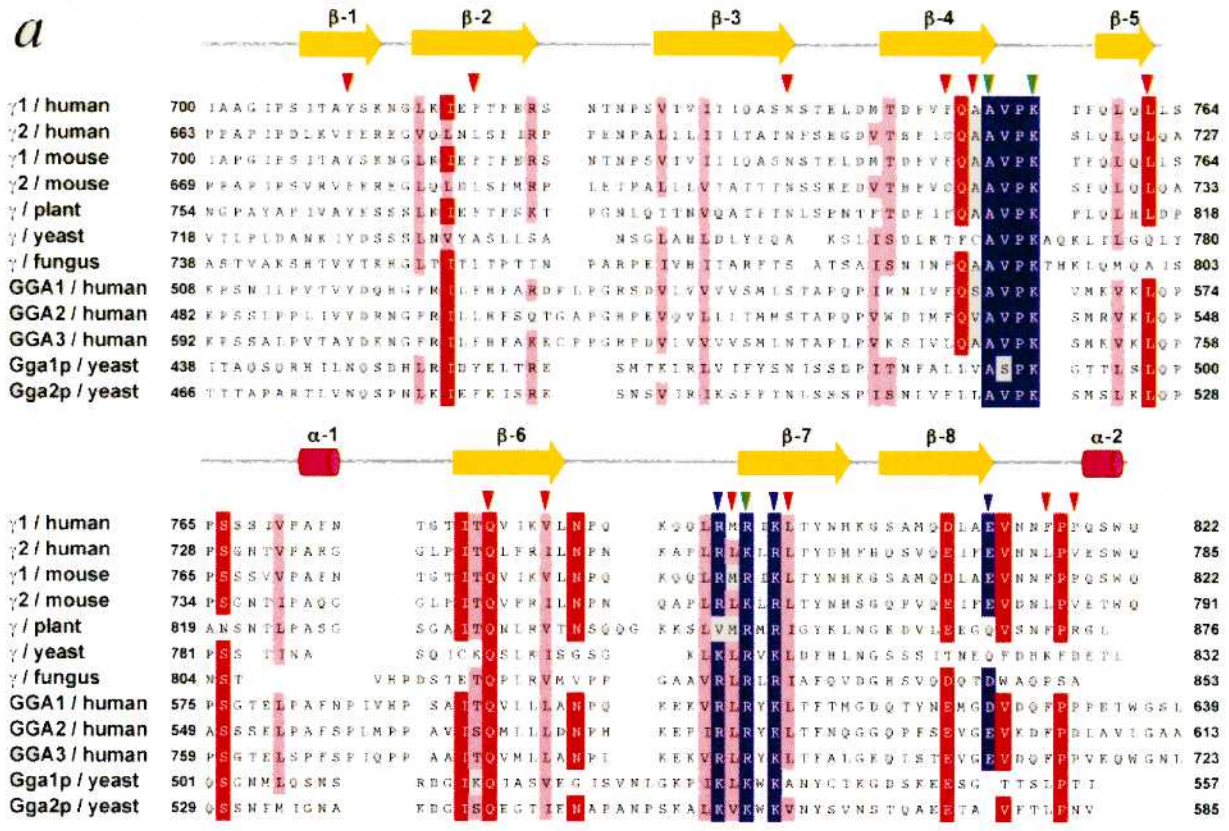

Hydrophobic binding pocket
for accessory proteins
a
b
C-term.
Platform
N-term.
Ig-fold
α-adaptin ear domain
β2-adaptin ear domain

Fig. II-2. Structure of the α-ear and β2-ear domains. *a* Ribbon diagram of the mouse α-adaptin ear domain (PDB accession code: 1B9K). *b* Ribbon diagram of the human β2-adaptin ear domain (1E42). In addition to the N-terminal immunoglobulin-like subdomain (yellow), the a and β2 ear domains possess the C-terminal platform subdomain (blue). In terms of the Cα trace, the γ1 ear domain is more similar to the α ear domain than the β2 ear domain, which is consistent with the fact that γ1-adaptin of the AP-1 complex is the counterpart of α-adaptin of AP-2. The α and γ1 ear domains can be superimposed with a root mean square deviation of 1.1 Å for the 57 Ca atoms in the 8 β-strands. Root mean square deviations in the structure superposition were calculated using the program LSQKAB (86). In the α ear domain, the C-terminal subdomain serves as the recruitment platform for accessory proteins. A shallow hydrophobic binding pocket is located at the top of the platform subdomain (the magenta dotted circle in *a*).


a
β-1
β-2
β-3
β-4
β-5
γ1 / human
γ2 / human
γ1 / mouse
γ2 / mouse
γ / plant
γ / yeast
γ / fungus
GGA1 / human
GGA2 / human
GGA3 / human
Gga1p / yeast
Gga2p / yeast
α-1
β-6
β-7
β-8
α-2

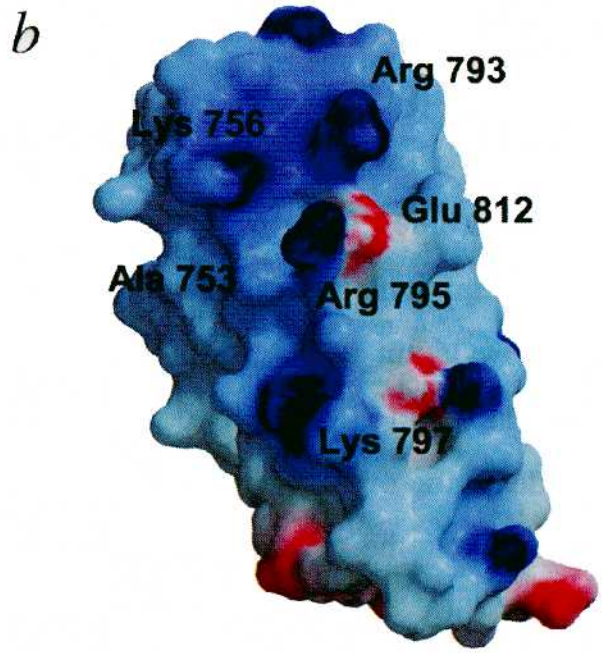

b
Arg 793
Lys 756
Glu 812
Asp 753
Arg 795
Lys 797

Fig. II-3. Conserved basic surface. *a* Sequence alignment of γ ear and GGA-GAE. The numbering for the human γ1 ear domain follows the sequence data with accession ID: AB015317 in GenBank/EBI/DDBJ database. In this sequence, the total number of amino acid residues is 822. In human and mouse, γ2-adaptin has been found in addition to γ1-adaptin. The sequences indicated as plant, yeast, and fungus correspond to those of *Arabidopsis thaliana*, *Saccharomyces cerevisiae*, and *Ustilago maydis*, respectively. The diagram above the sequences depicts the secondary structures of the human γ1 ear domain. The residues conserved among at least 9 out of 12 sequences are colored in red for identity, and in pink for similarity. In addition, the residues in the vicinity of the highly conserved basic cluster, as shown in *b*, are highlighted in blue. Triangles on the sequences indicate the residues of which the point mutations were shown to abolish the interaction with γ-synergin and rabaptin-5 in yeast two-hybrid screen; the red triangles for the residues buried in the core of the γ1 ear molecule, and blue for the exposed residues. In addition, the green triangles indicate the additional point mutants that were designed for GST pull down assays, on the basis of the structural data. *b* Electrostatic surface potential of the human γ1-adaptin ear domain. The molecular surface prepared by GRASP (93) is shown in the same orientation as in Fig. II-1*b*. A large basic surface consists of the conserved residues around the C-terminus of the strand β-4 and the N-terminus of the strand β-7. The GST pull down assays in Fig. II-4 have indicated that the basic surface serves as the binding sites for γ-synergin and Rabaptin-5.

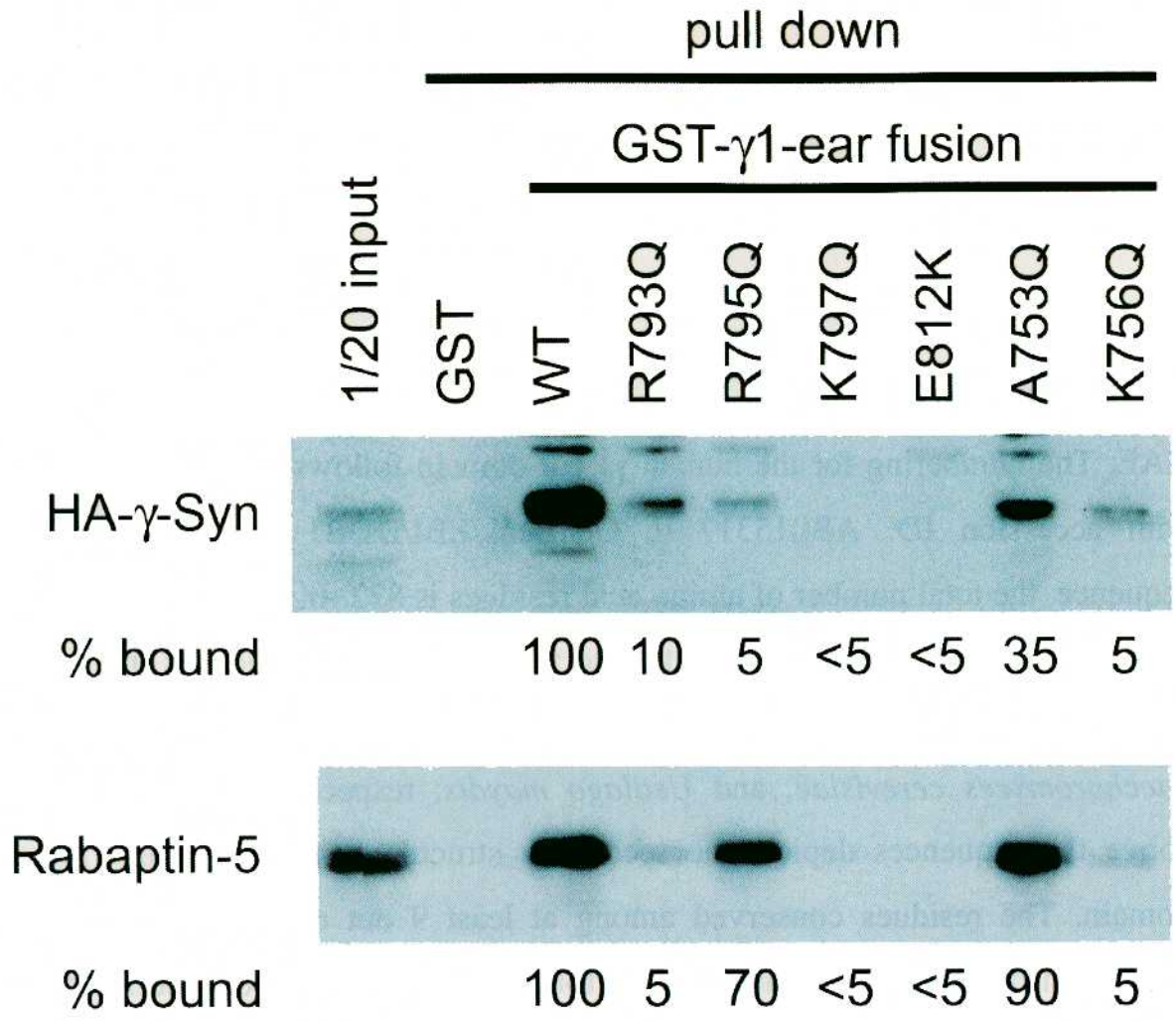


Fig. II-4. Pull down assay. Lysates from hEK-293 cells transfected with HA-tagged γ-synergin (HA-γ-Syn) were pulled down with GST or GST fused to wild type (WT) γ1-adaptin ear domain or its mutant and subjected to immunoblotting with anti-HA or anti-Rabaptin-5 antibody. GST and WT are the negative and positive control, respectively. The estimated band density of HA-γ-synergin or Rabaptin-5 pulled down with GST-γ1-ear (WT) is expressed as 100%. Left lane, one-twentieth volume of the lysate subjected to pull down was directly electrophoresed. Structural integrity of the six mutants was confirmed by measuring circular dichroic (CD) spectra of the proteins after their GST portions were cleaved with thrombin. Their CD spectra were almost identical, suggesting that they share a secondary structure similar to that of the wild type (data not shown).