A key aspect of eukaryotic intracellular trafficking is the sorting of endosomal cargoes. The ESCRT system is comprised of four major protein complexes, Crag, a regulator of protein sorting, and bacterial pore-forming proteins might mediate novel membrane interactions in trafficking. The UBA1-MVB12-associated UMA domain found in MVB12 and UBA1 defines a novel adaptor that might recruit diverse targets to ESCRT-I.


Received on March 9, 2010; revised on April 16, 2010; accepted on April 21, 2010

1 INTRODUCTION

A key aspect of eukaryotic intracellular trafficking is the sorting of cell-surface proteins into multi-vesicular endosomes or bodies (MVBs), which eventually fuse with the lysosome, where they are degraded by lipases and peptidases. This is the primary mechanism for downregulation of signaling via transmembrane receptors and removal of misfolded or defective membrane proteins (Raiborg and Stenmark, 2009). This process is also utilized by several viruses (e.g. HIV-1) to facilitate budding of their virions from the cell membrane (Morita et al., 2007). Studies in animals and fungi have shown that it depends on an intricate series of interactions, which is initiated via ubiquitination (typically one or more mono-ubiquitinations) of the cytoplasmic tails of membrane proteins by specific E3 ligases (d’Azzo et al., 2005). Ubiquitinated membrane proteins are then captured into endosomes by the ESCRT system and prevented from being recycled back to the plasma membrane via the retrograde trafficking system. The ESCRT system also folds the endosomal membranes into invaginations that are concentrated in these ubiquitinated targets and catalyzes their abscission into intraluminal vesicles inside the endosome. This largely seals the fate of these membrane proteins as targets for lysosomal degradation. The ESCRT system is comprised of four major protein complexes, ESCRT-0 to ESCRT-III, which are successively involved in the above-described steps (Raiborg and Stenmark, 2009). ESCRT-I, containing proteins with multiple Ub-binding modules, is the primary sensor for ubiquitinated membrane proteins. Both ESCRT-I and ESCRT-II have proteins with a single Ub-binding domain and are subsequent successive recipients of the ubiquitinated cargo. ESCRT-II proteins also contain lipid-binding modules and are likely to initiate invagination of the endosomal membrane. ESCRT-III, which includes the conserved AAA+ ATPase VPS4 as a component, mediates the final abscission of the invaginated membrane to form the intraluminal vesicle. In this relay, ESCRT-I is the critical bridge between the sensor of ubiquitinated targets and the membrane-binding ESCRT-II. ESCRT-I contains three subunits that are conserved between yeast and animals, namely the inactive E2-ligase protein TSG101/VPS23, VPS28 and VPS37 (Raiborg and Stenmark, 2009). Additionally, both yeast and metazoa ESCRT-I contain a fourth subunit termed MVB12 (‘multivesicular body sorting factor of 12 kD’ (Chu et al., 2006)); however, the MVB12 subunits from the two lineages do not show significant sequence similarity (Audhya et al., 2007; Chu et al., 2006; Komish et al., 2006; Morita et al., 2007). Metazoa MVB12 was shown to be critical for receptor endocytosis and also virus release (Morita et al., 2007). Given its key role in receptor downregulation, we were interested in understanding if the lack of detectable similarity with yeast MVB12 might reflect emergence of novel adaptations in animals.

 Accordingly, we analyzed the animal MVB12 proteins using sensitive sequence and structure analysis methods and identified two novel conserved domains in them. Identification of these domains allowed us to detect several putative, uncharacterized ESCRT-I subunits in animals. Characterization of these domains also provides new insights into recognition of cargo by endosomal sorting regulators.

2 METHODS

Profile searches were conducted using the PSI-BLAST program (Altschul et al., 1997) with a default profile inclusion expectation (E)-value threshold of 0.01. Profile–profile comparisons were performed using the HHpred program (Soding et al., 2005). Hidden Markov model searches were conducted using JAKHMMER from the HMMER3 package (Eddy, 2008). Multiple alignments were constructed using Kalagn (Lassmann and Sonnhammer, 2005) followed by manual adjustments based on PSI-BLAST results. Protein secondary structure was predicted using a multiple alignment as the input for the JPRED program (Cuff et al., 1998). The 3D structures were rendered using the Pymol program (http://www.pymol.org/).

3 RESULTS AND DISCUSSION

3.1 Identification of the UMA and MABP domains

To investigate the relationships of the animal MVB12, we used the closely related human paralogs MVB12A (FAM125A; gi: 24308440) and MVB12B (FAM125B; gi: 58761488) as seeds for sequence profile searches with the PSI-BLAST program and
R.F.de Souza and L.Aravind

Fig. 1. Multiple sequence alignment of the UMA (A) and MAPB (B) domains. Residues are colored according to the 85% consensus. Conserved MAPB positions are highlighted as listed in the lower box.

iterative hidden Markov model searches with the JACKHMMER program. The N-terminal region (human MVB12A, region 1-150) and the C-terminal region (MVB12A, region 210-264) recovered distinct sets of proteins. The N-terminal region of the MVB12A/B proteins hit several proteins from eukaryotes and bacteria. These included proteins typified by DENND4A/B/C from vertebrates (iteration 2, $10^{-5}$ in a PSI-BLAST search), the membrane-trafficking regulator Crag from Drosophila (iteration 3, $10^{-19}$), bacterial proteins typified by the MAC/Perforin (MACPF)-like protein plu1415 (PDB: 2QP2; iteration 4, $10^{-10}$) from *Photobacterium luminescens* and uncharacterized proteins from choanoflagellates and stramenopiles (Figs 1 and 2). In contrast, the C-terminal region produced significant hits only to metazoan proteins. These included the human ubiquitin-associated protein-1 (UBAP1; $e = 10^{-3}$, iteration 3 in PSI-BLAST), which is implicated in nasopharyngeal carcinoma risk and fronto-temporal lobar degeneration (Rollinson et al., 2009; Wu et al., 2009).

Also recovered were several other poorly characterized proteins, including at least one orthologous group of proteins conserved in vertebrates prototyped by the human protein LOC390595 (iteration 3, $10^{-5}$ in PSI-BLAST searches) and another group conserved across Metazoa typified by human tcag7.903 ($e = 10^{-4}$, iteration...
To understand the functional significance of the MABP domain and named this novel domain the MABP (Shimizu and Morikawa, 1996). However, the topology of the protein I (VMO-I) and the Bacillus thuringiensis first and third subdomain of the eukaryotic versions contains a conserved cysteine in the C-terminal domain of MVB12A/B domain, which is shared with the homologous proteins identified in the current study.

Furthermore, searches with the N-terminal domain of MVB12A/B and the equivalent domain in the DENN5A/AB/C and Crag indicated that it has an internal repeat structure of three homologous segments. Consistent with this, the structurally characterized representative, Photonhabudus plu1415, showed that this region precisely corresponds to a type-1 β-prism domain with an internal 3-fold symmetry (Rosado et al., 2007). Each of the three subdomains of the β-prism structure is a distinctive three-stranded β-sheet (Fig. 2B) that was congruent to the repeat units detected in the sequence searches (Fig. 1B). This domain shares a tridimensional structure with similar parallel β-sheets observed in the vitelline membrane outer layer protein 1 (VMO-I) and the Bacillus thuringiensis k-endotoxin (Shimizu and Morikawa, 1996). However, the topology of the strands in the β-sheet of the individual subdomains of the Photonhabudus plu1415 β-prism is entirely different (Fig. 2B). We named this novel domain the β-prism domain in at least two proteins have been found to bind monoubiquitin, a key trafficking signal (Raiborg et al., 2009; Yoshida and Tanaka, 2010). In particular, the Ub-binding ZnRs and the equivalent domain in the DENN domain is a Rab GEF that is required for Rab35-mediated recycling of endosomal proteins and trafficking of surface proteins to the apical membrane (Allaire et al., 2010). VPS13 and APOC-C domains have been implicated in protein cycling through the trans-Golgi network and formation of vesicles targeting for autophagy (Rampoldi et al., 2001). The other fusions are to DUBs and deglycanases that are also involved in the sorting of cargo in the trafficking process (Raiborg and Stenmark, 2009).

In eukaryotes, several architectures are observed including fusions to peptide-N-glycanase-type transglutaminase and PUG domains (Phytophthora infestans PITG_02329), to 8 EF-HANDs (EFh) and two Ub-binding ZnR domains (ZuH in Fig. 2A, P. infestans PITG_06730 and to a Sec7 domain (Phaeodactylum PHATDRKRAFT_49198). Two MABP domains are also found inserted into a deubiquitinating peptidase (DUB) domain in another P. infestans protein (PITG_02561; it also contains six N-terminal ZuH Ub-binding ZnRs). (v) In bacteria, the MABP domain occurs as a solo (e.g. Frankia PRAAL0413), fused to the C-terminus of a MACPF domain (e.g. plu1415) or at the N-terminus of a protein with two types of β-helix repeats (bH1/2) and a novel cysteine-containing domain (CCD) that are typical of cell-wall proteins (e.g. Clostridium CLO250_02048; Fig. 2A and Supplementary Material). In eukaryotes, several of the focal connections suggest that the MABP domain has a membrane-associated function, perhaps even specific interactions with membrane components. The structure of the MABP domain in plu1415 reveals several exposed hydrophobic residues that are

![Diagram of MABP and UMA domains](image)

**Fig. 2.** (A) Domain architectures of UMA and MABP containing proteins. (B) Structure of the MABP domain from Phaeodactylum plu1415 (PDB: 2QP2). Conserved residues P and Y of the second strand’s signature (PXGY, see Fig. 1) are represented as spheres. Only known domains are represented above, with unknown or uncharacterized regions omitted for simplicity. See text for domain name abbreviations.

3.2 Domain architectures and functional interactions of MABP and UMA domain proteins

To understand the functional significance of the MABP and UMA domains, we systematically determined domain architectures of the proteins which contain them (Fig. 2A). In addition to co-occurring with the UMA domain in MVB12 proteins found in all metazoans, the MABP domain is found independently of it but fused to several other domains: (i) In a group of related proteins typified by Crag and DENND4A/B/C found in metazoans and ciliates, it is present N-terminal to the triad of domains known as uDENN, DENN and dDENN (Levivier et al., 2001). Additionally, C-terminal to the DENN triad, these proteins have a pentatricopeptide repeat (PPR), a novel Zn-ribbon (ZdrK) and an uncharacterized α-helical domain. (ii) Two MABP domains are inserted into the choanoflagellate VPS13 ortholog, which also contains APG2-C and Dysferlin (DysF) domains. (iii) Stand-alone MABP domains are found in certain fungi. (iv) In stramenopiles, several architectures are observed including fusions to peptide-N-glycanase-type transglutaminase and PUG domains (Phytophthora infestans PITG_02329), to 8 EF-HANDs (EFh) and two Ub-binding ZnR domains (ZuH in Fig. 2A, P. infestans PITG_06730) and to a Sec7 domain (Phaeodactylum PHATDRKRAFT_49198). Two MABP domains are also found inserted into a deubiquitinating peptidase (DUB) domain in another P. infestans protein (PITG_02561; it also contains six N-terminal ZuH Ub-binding ZnRs). (v) In bacteria, the MABP domain occurs as a solo (e.g. Frankia PRAAL0413), fused to the C-terminus of a MACPF domain (e.g. plu1415) or at the N-terminus of a protein with two types of β-helix repeats (bH1/2) and a novel cysteine-containing domain (CCD) that are typical of cell-wall proteins (e.g. Clostridium CLO250_02048; Fig. 2A and Supplementary Material). In eukaryotes, several of the fused domains have been implicated in trafficking machinery: the DENN domain is a Rab GEF that is required for Rab35-mediated recycling of endosomal proteins and trafficking of surface proteins to the apical membrane (Allaire et al., 2010). VPS13 and APOC-C domains have been implicated in protein cycling through the trans-Golgi network and formation of vesicles targeted for autophagy (Rampoldi et al., 2001). The other fusions are to DUBs and deglycanases that are also involved in the sorting of cargo in the trafficking process (Raiborg and Stenmark, 2009; Yoshida and Tanaka, 2010). In particular, the Ub-binding ZuHs associated with the MABP domain in at least two proteins have been found to bind monoubiquitin, a key trafficking signal (Raiborg and Stenmark, 2009). MABP domain-containing Drosophila Crag protein localizes to endosomal vesicle and plasma membranes (Deneil et al., 2008). Likewise, bacterial proteins with MABP-MACPF domains have been suggested to target membranes (Rosado et al., 2007). Vertebrate MACPF proteins contain a fusion to the lipid-binding C2 in place of the MABP domain. These contextual connections suggest that the MABP domain has a membrane-associated function, perhaps even specific interactions with membrane components. The structure of the MABP domain in plu1415 reveals several exposed hydrophobic residues that are
consistent with such an interaction. Hence, it is plausible that the 
eukaryotic MABP domains are adaptors that help linking other 
associated domains found in the same polypeptide to vesicular 
membranes.

In MVB12, the region including the UMA domain, but not the 
MABP domain, has been shown to interact with the N-terminal 
part of VPS37 and the C-terminal part of TSG101, both ESCRT-I 
components (Morita et al., 2007). This suggests that the UMA 
domain probably specifically recruits MVB12 to the ESCRT-I 
complex to form a quaternary complex. In UBAP1 and LOC390595, 
the UMA domain is fused to three C-terminal UBA domains, which 
are known to bind ubiquitin (Raiborg and Stenmark, 2009). Hence, 
they could interact via the UBA domains with ubiquitinated tails 
of membrane proteins, while their UMA domains recruit them to the 
core ESCRT-I complex. The remaining UMA domain proteins (e.g. 
tcag7.903 group; Fig. 2A) have their own conserved N-terminal 
extensions that could potentially interact with specific protein 
partners. Based on these observations, we propose that the different 
UMA domain proteins might function as alternative MVB12-like 
subunits that recruit different targets via their specific interaction 
modules (such as MABP or UBA or the specific extensions) to the 
ESCR-T-I complex. Thus, different types of UMA domains are likely 
to be required for downregulation of different sets of receptors in 
animals.

4 GENERAL CONCLUSIONS

Identification of the MABP and UMA domains throws light 
on two vital aspects of vesicular trafficking. First, the MABP 
domain could be a common denominator in the recognition of 
specific membrane-associated features by a functionally diverse 
set of trafficking proteins in eukaryotes and bacterial proteins 
involved in pore formation and cell-wall interaction. The prediction 
that the diverse metazoan UMA domain proteins are alternative 
MVB12-like proteins implies that the recruitment of ESCRT-I to 
endosomal structures could occur via diverse mechanisms, including 
the possible direct recognition of membranes by the MABP 
domain, interaction with ubiquitinated peptides or other protein– 
protein interactions. This could have been a response to the vast 
expansion of diverse signaling receptors such as receptor tyrosine 
kinas, ion channels and 7TM receptors in the metazoan lineage. 

Intriguingly, we found that plants (e.g. Arabidopsis ATSG53330) 
have a conserved protein that has a series of C-terminal UBA 
domains closely related to those found in UBAP1. While we failed 
to find statistically significant similarity between the N-terminal 
region of these plant proteins and the UMA domain, they share a 
few tantalizing sequence patterns. It cannot be ruled out that these 
plant proteins contain a region remotely related to the UMA domain 
and perform a comparable function in relation with the ESCRT 
system.

While certain core components of this system (e.g. VPS4 
and MIT domains of ESCRT-III) have been traced to archaea 
(Hobel et al., 2008), the MABP domain is not currently found in 
any archaea. Instead it is found in diverse bacteria, suggesting that 
the eukaryotes could have acquired it early in their evolution from a 
bacterial precursor. Thus, the eukaryotic vesicular trafficking system 
appears to have been pieced together from different components 
acquired from both archaean and bacterial precursors.

Funding: Intramural funds of the National Library of Medicine, 
National Institutes of Health, USA.

Conflict of Interest: none declared.

REFERENCES

Allaire, P.D. et al. (2010) The Connexin DENN domain is a GEF for Rab35 mediating 


Audhya, A. et al. (2007) MVB-12, a fourth subunit of metazoon ESCRT-I, functions in 

Chai, T. et al. (2006) New component of ESCRT regulates endosomal sorting complex 

Bioinformatics, 14, 892–893.

d’Azzo, A. et al. (2005) E3 ubiquitin ligases as regulators of membrane protein 
trafficking and degradation. Traffic, 6, 429–441.

Drew, N. et al. (2008) Crag regulates epithelial architecture and polarized deposition 


Hobel, C.F. et al. (2008) The Sulfolobus solfataricus AAA protein Sso0909, a homologue 

Komish, H. et al. (2006) CBP is a novel tyrosine-phosphorylated protein that might 

sequence alignment algorithm. BMC Bioinformatics, 6, 298.

Levinson, E. et al. (2001) eDENN, DENN, and dDENN: indistinguishable domains in Rab 
688–695.

Morita, E. et al. (2007) Identification of human MVB12 proteins as ESCRT-I subunits 
that function in HIV budding. Cell Host Microbe, 2, 41–53.

Raiborg, C. and Stenmark, H. (2009) The ESCRT machinery in endosomal sorting of 

Rampoldi, L. et al. (2001) A conserved sorting-associated protein is mutant in chorea-

Rollinson, S. et al. (2009) Ubiquitin associated protein 1 is a risk factor for 

Rosado-C.J. et al. (2007) A common fold mediates vertebrate defense and bacterial 


Soling, J. et al. (2005) The HYTpred interactive server for protein homology detection 

Wu, M. et al. (2009) Signaling transduction network mediated by tumor 
suppressor/susceptibility genes in NPC. Curr. Genomics, 10, 216–222.