Minireview

VHS domain – a longshoreman of vesicle lines

Olli Lohi[a], Anssi Poussu[b], Yuxin Mao[c], Florante Quiocho[c], Veli-Pekka Lehto[b,d,*]

[a] *Department of Pediatrics, University of Tampere, Tampere, Finland*
[b] *Department of Pathology, FIN-90014 University of Oulu, Oulu, Finland*
[c] *Howard Hughes Medical Institute and Department of Biochemistry, Baylor College of Medicine, Houston, TX, USA*
[d] *Department of Pathology, FIN-00014, University of Helsinki, Helsinki, Finland*

Received 6 November 2001; revised 12 November 2001; accepted 3 December 2001

First published online 20 December 2001

Edited by Gianni Cesareni and Mario Gimona

Abstract The VHS (Vps-27, Hrs and STAM) domain is a 140 residue long domain present in the very NH_2-terminus of at least 60 proteins. Based on their functional characteristics and on recent data on the involvement of VHS in cargo recognition in trans-Golgi, VHS domains are considered to have a general membrane targeting/cargo recognition role in vesicular trafficking. Structurally, VHS is a right-handed superhelix of eight helices with charged surface patches probably serving as sites of protein–protein recognition and docking.

Key words: GGA; Signaling; Vesicular trafficking; VHS domain

1. Introduction

The VHS domain was originally identified in a database screen based on the multiple occurrence of stretches of sequences in signal transduction proteins [1]. The name VHS derives from its occurrence in Vps-27, Hrs and STAM. As originally defined, the VHS domain is ~140 residues long and is found in at least 60 proteins. Remarkably, in all of them, VHS occupies the N-terminal end of the polypeptide suggesting that such a topology is important for its function [2].

2. VHS marks proteins involved in vesicular trafficking

Based on the domain/motif entourage, VHS-containing proteins can be divided into four groups. Their representatives are schematically shown in Fig. 1. The first group consists of proteins of the STAM/EAST/Hbp family which all share the domain composition VHS-SH3-ITAM. The second consists of proteins with a FYVE domain C-terminal to VHS. The third consists of GGA proteins with a domain composition VHS-GAT-'ear' and the fourth of proteins with a VHS domain alone or with domains other than those mentioned above.

Importantly, the members of the first and the second groups also carry one or two ubiquitin-interacting motifs (UIM), an approximately 20 residue long motif that is usually found in proteins involved in ubiquitination and ubiquitin metabolism [3]. STAM/EAST/Hbp proteins also carry NPF motifs, a short domain recognizing the Eps15 homology domain [4].

The STAM/EAST/Hbp family [5–7] consists of eight members which are well conserved from yeast to mammals. Among them, EAST (EGF receptor-associated protein with SH3 and TAM domains) associates with the epidermal growth factor (EGF) receptor and is phosphorylated in response to EGF [5]. It colocalizes with clathrin and associates also with Eps15, a protein which is required for clathrin-mediated receptor endocytosis [8]. STAM (signal transducing adapter molecule), on the other hand, is involved in cytokine-mediated intracellular signal transduction and regulation of e.g. myc expression [6,9,10]. Moreover, through binding of Hrs (hepatocyte growth factor-regulated tyrosine kinase substrate), mediated by the ITAM motif, STAMs can also be part of the endocytic and exocytic machineries [7,11–14]. A further link to receptor signaling is provided by AMSH (associated molecule with the SH3 domain of STAM) which binds the SH3 domain of STAM and is involved in signaling pathways initiated by interleukin-2, granulocyte/macrophage colony-stimulating factor and bone morphogenetic protein [15,16].

VHS-FYVE proteins include the yeast Vps27 protein, implicated in the membrane trafficking through the prevacuolar/endosomal compartment in *Saccharomyces cerevisiae* [17], and its mammalian homolog Hrs, found on the surface of early endosomes and involved in endocytic trafficking and ligand-induced degradation of e.g. EGF receptor [12,18,19]. Hrs-2, originally identified as a distinct molecule and described as a modulator of vesicle trafficking in neurotransmission [20], has more recently been shown to be identical to Hrs [7].

FYVE (see also Stenmark et al., this issue), the domain which in these proteins reside C-terminal to VHS, derives its name from its occurrence in Fab1b, YOTB, Vac1p, and EEA1 [21]. It binds specifically to the membrane lipid phosphatidylinositol 3-phosphate [22,23] and seems to be important for the subcellular targeting of Hrs [24,25] although there are also conflicting results [26].

The presence of UIM in both STAM/EAST/Hbp and Hrs is a further indication that these proteins are important for endocytosis and in further sorting for degradation, respectively, of at least some plasma membrane receptors. This is based on what is generally known about the role of ubiquitination in receptor endocytosis; tagging with ubiquitin promotes internalization of the receptor, appears to regulate the activity of

*Corresponding author. Fax: (358)-8-537 5953.
E-mail address: lehto@csc.fi (V.-P. Lehto).

PII: S0014-5793(01)03287-2

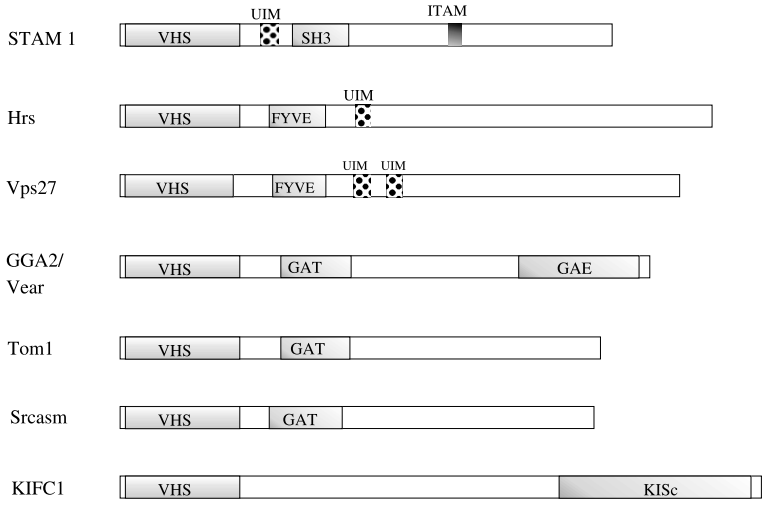

Fig. 1. Domain composition of representatives of different subclasses of VHS-containing proteins. FYVE, domain present in Fab1b, YOTB, Vac1p, and EEA1; GAE, gamma-adaptin, ear; GAT, domain present in GGA and TOM1; KISc, kinesin motor catalytic domain; SH3, Src homology 3 domain; UIM, ubiquitin-interacting motif; VHS, domain present in Vps-27, Hrs, STAM.

the endocytic apparatus, and, as shown recently, sorts the receptors to multivesicular body vesicles to be further ferried to lysosomes for degradation [27]. Moreover, as to the VHS protein in particular, Hbp was recently shown to bind, via its SH3 domain, to the deubiquitinating enzyme UBPY [28], and the yeast Vps27 interacts genetically with a deubiquitinating enzyme Doa4p [29]. Thus, the UIM-containing members of the VHS proteins could, by directly recognizing the ubiquitinated cargo, also be part of the ubiquitin-based machinery of receptor endocytosis and degradation.

GGAs are a recently identified family of proteins with three members (GGA1–3) characterized in humans and two (Gga1p and Gga2p) in *S. cerevisiae* [30–35] and with at least one related expressed sequence tag identified in the *Drosophila*, *Caenorhabditis elegans* and *Schizosaccharomyces pombe* genome sequencing projects, each. They are composed of an N-terminal VHS domain, the GAT (GGA and Tom1) homology domain which is the best conserved and bears homology to Tom1 protein, a flexible hinge region which contains clathrin-binding motifs, and the C-terminal GAE (gamma-adaptin ear) domain with homology to the ear domain of γ-adaptin. GGAs are ARF-binding proteins. Hence the name GGA for Golgi-localizing, gamma-adaptin ear homology domain, ARF-binding proteins. The founding member of the family was originally named Vear (VHS and ear domain-containing protein [36]) and is identical to GGA2.

GGAs localize predominantly to the *trans*-Golgi network (TGN) with specific GGAs associating also with other vesicular elements of the cytoplasm. Functional studies with mammalian cells supported the idea that GGAs operate at the level of TGN facilitating the vesicular trafficking from the Golgi. This was corroborated by the studies in yeast in which a specific role for GGAs in a transport from TGN to early and/or to late endosomes could be shown. Just recently, it was demonstrated by several groups that GGAs act as sorting adapters; in TGN they recognize and interact directly, via their VHS domains, with specific cargos and mediate their sorting to vesicles that are destined to endosomal/lysosomal compartments [37–40].

3. Structure of VHS

In Fig. 2, CLUSTALX alignment of the VHS domains of 15 different VHS proteins is given. There is a distinct grouping to different subclasses according to the domain composition as described above. Above the alignment, the positions of the α-helices based on the crystal structure of *Drosophila* Hrs [41] are given.

Crystal structures of the VHS domains from both *Drosophila* Hrs [41] and human Tom1 [42] have been determined. The VHS domain contains eight α-helices and a C-terminal extension (Fig. 3A). The eight α-helices fold into a curved double-layer superhelical structure with a size of $20\times35\times45$ Å. The concave face contains three α-helices, α2, α4 and α7, while the convex layer consists of four helices, α1, α3, α6 and α8. The two-turn α5 connects the two α-helix layers resulting in a slightly larger distance between the two layers on one side. Structural comparison indicates that this double-layered superhelical structure is built up mainly by three α-helical repeats (Fig. 3). The first two two-helix hairpins (α1 and α2, and α3 and α4) are very similar to the HEAT repeat [43]. In contrast, the third repeat, consisting of α5, α6 and α7, resembles the three-helix ARM repeat [44]. The hybridization of two different α-helix repeats makes the VHS domain a unique member among the superhelical structure family.

The overall sequence similarity is very low (∼25%) among members of the VHS domain family. On the basis of the three-dimensional structure, most of the conserved residues are mapped in a channel created by the double helix layers. These residues are mostly hydrophobic and make up the core of the domain. The distribution pattern of these hydrophobic residues in the channel is a major determinant of the packing geometry of the superhelical structure of the VHS domain. Several other conserved residues are localized on the surface of the domain. Of particular interest are Trp-23, Leu-27, and Asp-31 in Hrs, which form a cluster interfacing the following FYVE domain [41]. The existence of some conserved patches on the surface of the VHS domain suggests that VHS domains may be involved in protein–protein recognition and docking.

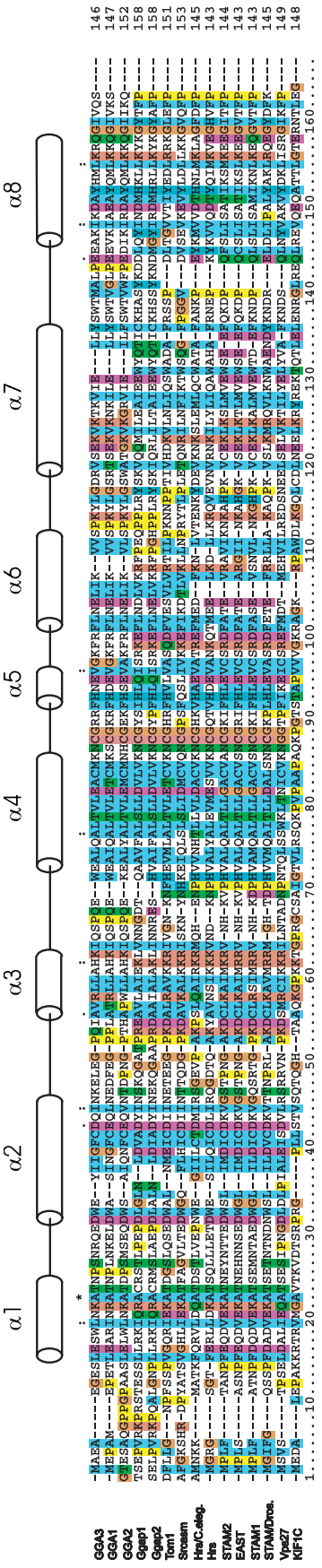

Fig. 2. Multiple sequence alignment (CLUSTALX) of VHS-containing proteins. The accession numbers are as follows: EAST, gi-3417246; GGA1, gi-14548066; GGA2, gi-14548065; GGA3, gi-14548064; Ggap1, gi-14548059; Ggap2, gi-731696; Hrs, gi-4758528; Hrs/C.eleg., gi-1326383; K1F1C, gi-4050097; Srcasm, gi-15077847; STAM/Dros., gi-5006441; STAM1, gi-3645903; STAM2, gi-3650488; Tom1, gi-3256185; Vps27, gi-1353233; The α-helices based on the structure of Hrs VHS domain (cf. Fig. 3A) are shown above the alignment.
←

Interestingly, the structure of the VHS domain is very similar to the recently reported structure of the ENTH (Epsin NH_2-terminal homology) domain, superimposing with an rms deviation of ~1.8 Å over the first seven α-helices (Fig. 3B) [45,46]. The major difference between the two structures is the orientation of the C-terminal α8. In the VHS domain, the α8 helix packs parallel with α6 and sits in the groove between α6 and α7, while in the ENTH domain, α8 runs perpendicularly across α4 and α2. Since the sequence similarity between the VHS and ENTH domains cannot be detected by sequence analysis programs, the ENTH domains constitute a distinct domain family [47]. Nevertheless, some conserved hydrophobic residues reside in the channel between the two α-helix layers. The structural similarity of the two domains and the observation that both VHS and ENTH domain-containing proteins are involved in membrane trafficking suggest that both domains may have arisen from a common ancestor. The superhelical fold of the VHS and ENTH domains is preserved by the conserved hydrophobic core, while functional differences are obtained through the variation of the surface residues.

4. Function of VHS

The most definite results concerning the function of VHS come from the studies mentioned above showing that the VHS domains of GGAs interact with some sorting receptors that traffic and transfer cargo between TGN and the endosomal compartment. They include sortilin [37], a sorting receptor for cargo such as neurotensin, and the cation-dependent and -independent mannose 6-phosphate receptors (M6PR) which act as carriers of the mannose 6-phosphate-tagged hydrolases to be ferried to lysosomes [38–40]. The minimal recognition sequence in M6PR is an acidic-cluster–dileucine motif. Attesting to the specificity of the interaction, many other transmembrane receptors that lack such complete recognition motifs do not bind GGAs, and, conversely, VHS domains from other proteins do not share the binding properties of GGAs. From these studies, combined with the results from functional studies using mutant receptors, it is now well established that, via VHS, the GGAs interact directly with the sorting signals of the cargo and facilitate its incorporation into vesicles that depart TGN.

It is important to note that the interaction of the VHS domain with the receptor of the VHS alone is not sufficient to recruit GGA from the cytosol to the TGN; overexpressed VHS of GGA3 alone shows a diffuse cytoplasmic distribution without any clear membrane association [31]. Thus, most probably, the membrane localization of GGAs is due to multiple membrane contacts which also involve other domains. In fact, the GAT domain, by virtue of its binding to Golgi-specific ARF, is the principal driver of the membrane association

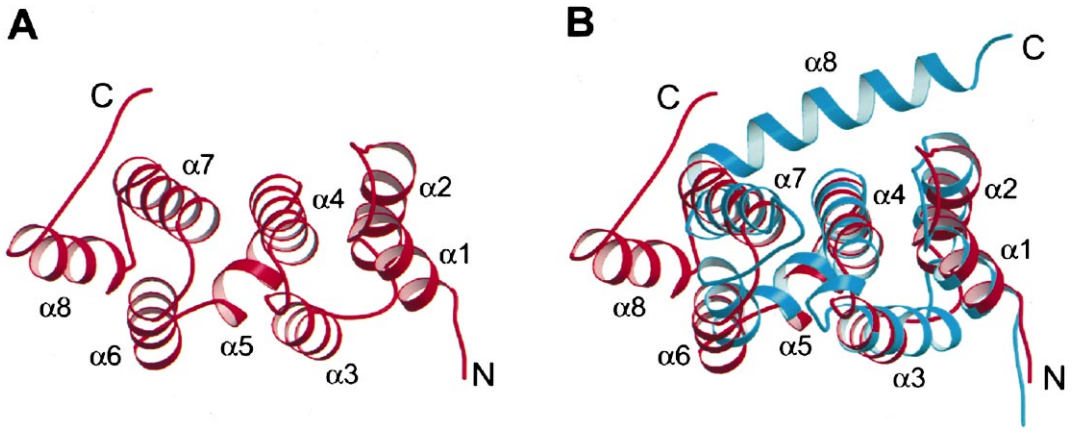

Fig. 3. A: Ribbon diagram of the structure of the VHS domain of Hrs with the helices labeled [41]. B: Superimposed ribbon diagram of the structures of the VHS (red) and ENTH (cyan) [41,45].

of GGA as evidenced by experiments using brefeldin A, by direct interaction studies, and by studies with mutant constructs [30–32,36]. Also the hinge region of GGAs may be involved due to its binding to clathrin [48,49]. Thus, as to GGAs, a picture is emerging in which multiple interactions are used to target the protein to a specific membrane compartment, followed by a VHS-mediated recognition/recruitment of the cargo. At this point it is not clear whether the GGAs function as coat proteins to nucleate their own coated vesicles or serve to mediate the entry of the cargo into forming AP-1 clathrin-coated vesicles. The GGAs could perform both of these functions.

Unlike GGAs, the functions of the VHS domains of other proteins are poorly known. In view of a high degree of intragroup similarity of the VHS domain especially within STAM/EAST/Hbp (70% identity) and also within VHS-FYVE proteins ($>35\%$ identity), which is higher than the overall similarity between the subgroups ($\sim 25\%$), it is to be expected that they interact with specific cargos as well. One crucial topic of future studies is to try to identify these putative, as yet unidentified target (receptor)s.

Regardless of the binding target of the VHS domains, it seems that, similar to GGAs, the primary targeting of STAM/EAST/Hbp is dictated by a coordinate action of several domains. This is evident from our own studies in which separately expressed N-terminal (containing VHS) and C-terminal domains of EAST displayed distinctly different subcellular association with overlapping but different vesicular structures [5]. Interestingly, the localization of the exogenously expressed VHS domain rather than of the C-terminal half was closer to that of the native EAST, suggesting the primacy of the VHS domain as a targeting domain in this class of VHS proteins. In STAM, a coiled-coil region which overlaps with the ITAM motif could provide one such target-specifying region through its binding of the early endosome-associated Hrs [11]. Another targeting mechanism for EAST and STAM could be provided through their association either directly or in a multimolecular complex, with the EGF receptor and Eps15 [5], or with several cytokine receptors [6,9], respectively.

It is also worthy of note that the VHS domain of EAST binds actin, and, upon overexpression, brings about changes in the organization of the cortical actin cytoskeleton [50]. This could be of significance in view of the by now well-established role of the actin cytoskeleton in the orderly functioning of the endocytic machinery [51].

Hrs is the best studied of the VHS-FYVE proteins. Its targeting to early endosomes, its primary residence, is through the FYVE domain along with a coiled-coil domain as demonstrated by mutant studies [25]. On the other hand, there are also results pointing to a dispensability of FYVE in this targeting [26]. The functional features of VHS in this class of proteins are poorly known except that it seems to be non-essential for endocytic targeting of Hrs [26].

5. Conclusions

The uniform presence of VHS in the very N-terminus of its residence proteins suggests that it serves a closely similar function in all of them. Based on the by now well-established cargo recognition functions of VHS in GGAs, it is reasonable to assume that the VHS domains of other proteins also are involved in such actions. Thus, it could be anticipated that the recognition of the cargo, either directly or indirectly, by VHS plays a role in the STAM/EAST/Hbp-associated endocytosis of a selected set of receptor molecules. VHS-FYVE proteins operate primarily at the level of early endosomes. Could their VHS domains sort incoming receptors for some specific pathway from these vesicles onwards? Whether VHS is only one, 'non-specific', longshoreman that recognizes the cargo based on a common code (with the specificity coming from the interactions of other domains which dictate the location of the protein), or whether there are in fact several longshoremen which only work specific cargos (and load only certain types of vesicles without any help in recognition from the neighboring domains), and whether the functions include not only loading but also unloading, remains to be seen.

Acknowledgements: Y.M. is an Associate and F.A.Q. is an Investigator in the Howard Hughes Medical Institute (Houston, TX, USA). This study was supported in part by the Finnish Academy and the Finnish Cancer Foundation. We want to thank Dr. Stuart Kornfeld for comments and Ms. Hilkka Penttinen and Mr. Hannu Wäänänen for skillful secretarial and image processing assistance.

References

[1] Schultz, J., Milpetz, F., Bork, P. and Ponting, C.P. (1998) Proc. Natl. Acad. Sci. USA 95, 5857–5864.

[2] Lohi, O. and Lehto, V.P. (1998) FEBS Lett. 440, 255–257.
[3] Hofmann, K. and Falquet, L. (2001) Trends Biochem. Sci. 26, 347–350.
[4] Wong, W.T., Schumacher, C., Salcini, A.E., Romano, A., Castagnino, P., Pelicci, P.G. and Di Fiore, P. (1995) Proc. Natl. Acad. Sci. USA 92, 9530–9534.
[5] Lohi, O., Poussu, A., Merilainen, J., Kellokumpu, S., Wasenius, V.M. and Lehto, V.P. (1998) J. Biol. Chem. 273, 21408–21415.
[6] Takeshita, T., Arita, T., Asao, H., Tanaka, N., Higuchi, M., Kuroda, H., Kaneko, K., Munakata, H., Endo, Y., Fujita, T. and Sugamura, K. (1996) Biochem. Biophys. Res. Commun. 225, 1035–1039.
[7] Takata, H., Kato, M., Denda, K. and Kitamura, N. (2000) Genes Cells 5, 57–69.
[8] Carbone, R., Fre, S., Iannolo, G., Belleudi, F., Mancini, P., Pelicci, P.G., Torrisi, M.R. and Di Fiore, P.P. (1997) Cancer Res. 57, 5498–5504.
[9] Takeshita, T., Arita, T., Higuchi, M., Asao, H., Endo, K., Kuroda, H., Tanaka, N., Murata, K., Ishii, N. and Sugamura, K. (1997) Immunity 6, 449–457.
[10] Endo, K., Takeshita, T., Kasai, H., Sasaki, Y., Tanaka, N., Asao, H., Kikuchi, K., Yamada, M., Chenb, M., O'Shea, J.J. and Sugamura, K. (2000) FEBS Lett. 477, 55–61.
[11] Asao, H., Sasaki, Y., Arita, T., Tanaka, N., Endo, K., Kasai, H., Takeshita, T., Endo, Y., Fujita, T. and Sugamura, K. (1997) J. Biol. Chem. 272, 32785–32791.
[12] Chin, L.S., Raynor, M.C., Wei, X., Chen, H.Q. and Li, L. (2001) J. Biol. Chem. 276, 7069–7078.
[13] Komada, M. and Kitamura, N. (2001) Biochem. Biophys. Res. Commun. 281, 1065–1069.
[14] Murai, S. and Kitamura, N. (2000) Biochem. Biophys. Res. Commun. 277, 752–756.
[15] Tanaka, N., Kaneko, K., Asao, H., Kasai, H., Endo, Y., Fujita, T., Takeshita, T. and Sugamura, K. (1999) J. Biol. Chem. 274, 19129–19135.
[16] Itoh, F., Asao, H., Sugamura, K., Heldin, C.H., ten Dijke, P. and Itoh, S. (2001) EMBO J. 20, 4132–4142.
[17] Piper, R.C., Cooper, A.A., Yang, H. and Stevens, T.H. (1995) J. Cell Biol. 131, 603–617.
[18] Komada, M., Masaki, R., Yamamoto, A. and Kitamura, N. (1997) J. Biol. Chem. 272, 20538–20544.
[19] Komada, M. and Kitamura, N. (1995) Mol. Cell. Biol. 15, 6213–6221.
[20] Bean, A.J., Seifert, R., Chen, Y.A., Sacks, R. and Scheller, R.H. (1997) Nature 385, 826–829.
[21] Stenmark, H., Aasland, R., Toh, B.H. and D'Arrigo, A. (1996) J. Biol. Chem. 271, 24048–24054.
[22] Gaullier, J.M., Simonsen, A., D'Arrigo, A., Bremnes, B., Stenmark, H. and Aasland, R. (1998) Nature 394, 432–433.
[23] Burd, C.G. and Emr, S.D. (1998) Mol. Cell. 2, 157–162.
[24] Raiborg, C., Bache, K.G., Mehlum, A. and Stenmark, H. (2001) Biochem. Soc. Trans. 29, 472–475.
[25] Raiborg, C., Bremnes, B., Mehlum, A., Gillooly, D.J., D'Arrigo, A., Stang, E. and Stenmark, H. (2001) J. Cell Sci. 114, 2255–2263.
[26] Hayakawa, A. and Kitamura, N. (2000) J. Biol. Chem. 275, 29636–29642.
[27] Hicke, L. (2001) Cell 106, 527–530.
[28] Kato, M., Miyazawa, K. and Kitamura, N. (2000) J. Biol. Chem. 275, 37481–37487.
[29] Amerik, A.Y., Nowak, J., Swaminathan, S. and Hochstrasser, M. (2000) Mol. Biol. Cell. 11, 3365–3380.
[30] Hirst, J., Lui, W.W., Bright, N.A., Totty, N., Seaman, M.N. and Robinson, M.S. (2000) J. Cell Biol. 149, 67–80.
[31] Dell'Angelica, E.C., Puertollano, R., Mullins, C., Aguilar, R.C., Vargas, J.D., Hartnell, L.M. and Bonifacino, J.S. (2000) J. Cell Biol. 149, 81–94.
[32] Boman, A.L., Zhang, C., Zhu, X. and Kahn, R.A. (2000) Mol. Biol. Cell 11, 1241–1255.
[33] Takatsu, H., Yoshino, K. and Nakayama, K. (2000) Biochem. Biophys. Res. Commun. 271, 719–725.
[34] Black, M.W. and Pelham, H.R. (2000) J. Cell Biol. 151, 587–600.
[35] Zhdankina, O., Strand, N.L., Redmond, J.M. and Boman, A.L. (2001) Yeast 18, 1–18.
[36] Poussu, A., Lohi, O. and Lehto, V.P. (2000) J. Biol. Chem. 275, 7176–7183.
[37] Nielsen, M.S., Madsen, P., Christensen, E.I., Nykjaer, A., Gliemann, J., Kasper, D., Pohlmann, R. and Petersen, C.M. (2001) EMBO J. 20, 2180–2190.
[38] Puertollano, R., Aguilar, R.C., Gorshkova, I., Crouch, R.J. and Bonifacino, J.S. (2001) Science 292, 1712–1716.
[39] Zhu, Y., Doray, B., Poussu, A., Lehto, V.P. and Kornfeld, S. (2001) Science 292, 1716–1718.
[40] Takatsu, H., Katoh, Y., Shiba, Y. and Nakayama, K. (2001) J. Biol. Chem. 276, 28541–28545.
[41] Mao, Y., Nickitenko, A., Duan, X., Lloyd, T.E., Wu, M.N., Bellen, H. and Quiocho, F.A. (2000) Cell 100, 447–456.
[42] Misra, S., Beach, B.M. and Hurley, J.H. (2000) Biochemistry 39, 11282–11290.
[43] Groves, M.R. and Barford, D. (1999) Curr. Opin. Struct. Biol. 9, 383–389.
[44] Huber, A.H., Nelson, W.J. and Weis, W.I. (1997) Cell 90, 871–882.
[45] Hyman, J., Chen, H., Di Fiore, P.P., De Camilli, P. and Brunger, A.T. (2000) J. Cell Biol. 149, 537–546.
[46] Mao, Y., Chen, J., Maynard, J.A., Zhang, B. and Quiocho, F.A. (2001) Cell 104, 433–440.
[47] Kay, B.K., Yamabhai, M., Wendland, B. and Emr, S.D. (1999) Protein Sci. 8, 435–438.
[48] Puertollano, R., Randazzo, P.A., Presley, J.F., Hartnell, L.M. and Bonifacino, J.S. (2001) Cell 105, 93–102.
[49] Costaguta, G., Stefan, C.J., Bensen, E.S., Emr, S.D. and Payne, G.S. (2001) Mol. Biol. Cell 12, 1885–1896.
[50] Lohi, O. and Lehto, V.P. (1998) FEBS Lett. 436, 419–423.
[51] Qualmann, B., Kessels, M.M. and Kelly, R.B. (2000) J. Cell Biol. 150, F111–F116.