Annu. Rev. Biochem. 2003. 72:395–447
doi: 10.1146/annurev.biochem.72.121801.161800
First published online as a Review in Advance on March 6, 2003

SIGNALS FOR SORTING OF TRANSMEMBRANE PROTEINS TO ENDOSOMES AND LYSOSOMES*

Juan S. Bonifacino[1] and Linton M. Traub[2]
[1]*Cell Biology and Metabolism Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892; email: juan@helix.nih.gov*
[2]*Department of Cell Biology and Physiology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261; email: traub+@pitt.edu*

Key Words endocytosis, plasma membrane, *trans*-Golgi network, receptor downregulation, multivesicular bodies, vacuole, phosphoinositides

■ **Abstract** Sorting of transmembrane proteins to endosomes and lysosomes is mediated by signals present within the cytosolic domains of the proteins. Most signals consist of short, linear sequences of amino acid residues. Some signals are referred to as tyrosine-based sorting signals and conform to the NPXY or YXXØ consensus motifs. Other signals known as dileucine-based signals fit [DE]XXXL[LI] or DXXLL consensus motifs. All of these signals are recognized by components of protein coats peripherally associated with the cytosolic face of membranes. YXXØ and [DE]XXXL[LI] signals are recognized with characteristic fine specificity by the adaptor protein (AP) complexes AP-1, AP-2, AP-3, and AP-4, whereas DXXLL signals are recognized by another family of adaptors known as GGAs. Several proteins, including clathrin, AP-2, and Dab2, have been proposed to function as recognition proteins for NPXY signals. YXXØ and DXXLL signals bind in an extended conformation to the μ2 subunit of AP-2 and the VHS domain of the GGAs, respectively. Phosphorylation events regulate signal recognition. In addition to peptide motifs, ubiquitination of cytosolic lysine residues also serves as a signal for sorting at various stages of the endosomal-lysosomal system. Conjugated ubiquitin is recognized by UIM, UBA, or UBC domains present within many components of the internalization and lysosomal targeting machinery. This complex array of signals and recognition proteins ensures the dynamic but accurate distribution of transmembrane proteins to different compartments of the endosomal-lysosomal system.

CONTENTS

INTRODUCTION

Lysosomes are the final destination for many macromolecules taken up by endocytosis from the extracellular space and from the cell surface (Figure 1). Nutrient carrier proteins such as cholesterol-laden low density lipoprotein (LDL) particles, for example, bind to specific receptors on the surface of cells and are rapidly internalized into endosomes. The acidic pH of endosomes induces dissociation of the LDL-receptor complexes, after which the LDL is delivered to lysosomes for degradation while the receptor returns to the cell surface for additional rounds of endocytosis. Binding of hormones, growth factors, and other signaling molecules to their cognate receptors can also trigger downregulation of the receptors by inducing their internalization and delivery to lysosomes. Finally, some endosomal and lysosomal proteins traffic via the plasma membrane en route to their main sites of residence at steady state.

Lysosomes can also be accessed via the biosynthetic pathway (Figure 1). In a process that mirrors receptor-mediated endocytosis, newly synthesized acidic hydrolases modified with mannose 6-phosphate groups bind to mannose 6-phosphate receptors (MPRs) in the *trans*-Golgi network (TGN), after which they are transported to endosomes following an intracellular route. The hydrolase-MPR

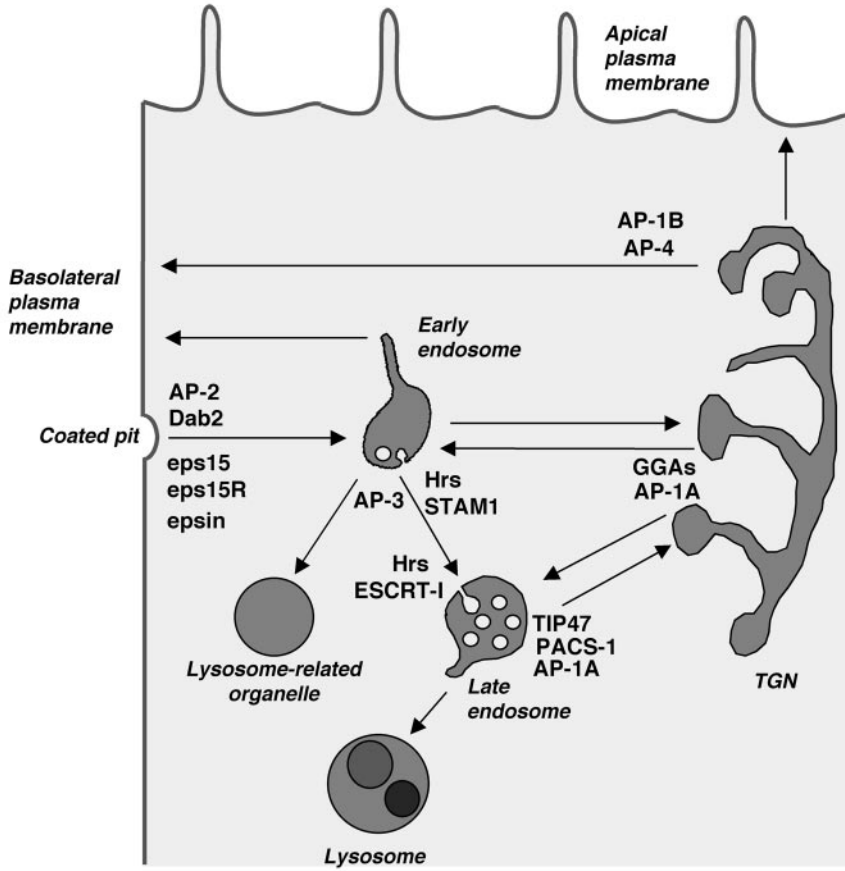


Figure 1 Sorting pathways leading to endosomes, lysosomes, and related organelles. The presumed location and site of action of different coat proteins are indicated by their placement next to an organelle or arrow origin. These should be considered tentative, as in most cases they have not been definitively established. Lysosome-related organelles include melanosomes, platelet-dense bodies, antigen-processing compartments, lytic granules, and other organelles that share some biogenetic pathways with endosomes and lysosomes.

complexes dissociate in the acidic pH of the endosomes, and then the hydrolases are transported to lysosomes with the fluid phase while the MPRs return to the TGN for further rounds of hydrolase sorting. Abnormal or excess proteins targeted for degradation, as well as many lysosomal resident proteins, also travel from the TGN to endosomes and lysosomes intracellularly.

Both the endocytic and biosynthetic routes are critical for the biogenesis and function of endosomes and lysosomes. The proper operation of these transport

routes requires that several important sorting decisions be made along the way. At the plasma membrane, proteins can either remain at the cell surface or be rapidly internalized into endosomes. At the TGN, the choice is between going to the plasma membrane or being diverted to endosomes. In endosomes, proteins can either recycle to the plasma membrane or go to lysosomes. These decisions are governed by a complex system of sorting signals in the itinerant proteins and a molecular machinery that recognizes those signals and delivers the proteins to their intended destinations.

This article focuses on sorting signals present in transmembrane proteins that are targeted to endosomes, lysosomes, and related organelles, and on the proteins that recognize these signals. These issues have been the subject of previous reviews (1–3), but recent advances in the understanding of the nature of the signals, the identification of new signals, and the elucidation of their mechanisms of recognition allow us to provide a more complete and updated account of this topic.

GENERAL PROPERTIES OF ENDOSOMAL-LYSOSOMAL SORTING SIGNALS

Most endosomal-lysosomal sorting signals characterized to date are contained within the cytosolic domains of transmembrane proteins. A list of known signals and their recognition proteins or domains is presented in Table 1. In general, the signals consist of short, linear arrays of amino acid residues. These arrays are not exactly conserved sequences but degenerate motifs of four to seven residues of which two or three are often critical for function. The critical residues are generally bulky and hydrophobic, although charged residues are also important determinants of specificity for some signals. Two major classes of endosomal-lysosomal sorting signals are referred to as "tyrosine-based" and "dileucine-based" owing to the identity of their most critical residues. Sorting mediated by these signals is saturable in vivo, indicating that it relies on recognition of the signals by a limited number of "receptor-like" molecules. In cases in which these recognition molecules have been identified and characterized, interactions have been found to occur with low affinity relative to other protein-protein interactions, such as those between polypeptide hormones and their receptors. Sorting signals can be considered a subset of a larger group of peptide-motif recognition-domain interactions (including those listed in Table 2) that are a characteristic feature of the endocytic and lysosomal sorting machineries. The defining property that sets sorting signals apart from other motifs is their occurrence within the cytosolic domains of transmembrane proteins.

Not all signals are short peptide motifs, however. In some cases, the sorting determinants appear to be folded structures in which the critical residues are not necessarily colinear (4). A striking example of this type of conformational determinant is the protein ubiquitin, which can function as an endosomal-

TABLE 1 Endosomal/lysosomal sorting signals

Signal type	Proposed recognition protein or domain	Functions
NPXY	Clathrin terminal domain, μ2 subunit of AP-2, PTB domain of Dab2	Internalization
YXXØ	μ subunits of AP complexes	Internalization, lysosomal targeting, basolateral targeting
[DE]XXXL[LI]	μ and/or β subunits of AP complexes	Internalization, lysosomal targeting, basolateral targeting
DXXLL	VHS domain of the GGAs	TGN-to-endosomes sorting
Acidic cluster	PACS-1	Endosomes-to-TGN sorting
FW- or P-rich	TIP47	Endosomes-to-TGN sorting
NPFX(1,2)D	SHD1 domain of *Saccharomyces cerevisiae* S1a1p	Internalization
Ubiquitin	UIM, UBA, and UBC domains	Internalization, lysosomal/vacuolar targeting

Motifs in this and other tables are denoted using the PROSITE syntax (*http://www.expasy.ch/prosite/*). Amino acid residues are designated according to the single letter code as follows: A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan, and Y, tyrosine. X stands for any amino acid and Ø stands for an amino acid residue with a bulky hydrophobic side chain. Abbreviations: PTB, phosphotyrosine-binding; Dab2, disabled-2; AP, adaptor protein; VHS domain present in Vps27p, Hrs, Stam; GGAs, Golgi-localized, γ-ear-containing, ARF-binding proteins; PACS-1, phosphofurin acidic cluster sorting protein 1; TIP47, tail-interacting protein of 47 kDa; SHD1, Sla1p homology domain 1; UBA, ubiquitin associated; UBC, ubiquitin conjugating; UIM, ubiquitin interaction motif.

lysosomal sorting signal upon covalent conjugation onto the cytosolic domain of some transmembrane proteins (5).

SIGNAL-MEDIATED SORTING OCCURS WITHIN COATED AREAS OF MEMBRANES

Long before the identification of sorting signals and their recognition proteins, it had become widely accepted that sorting in the endosomal-lysosomal system occurs through coated areas of membranes. Receptor-mediated endocytosis of lipoproteins, for instance, was found to involve concentration of the lipoprotein-receptor complexes within plasma membrane "coated pits" that become deeply invaginated and eventually give rise to "coated vesicles" (6, 7). A similar process was shown to occur at the TGN, where acidic hydrolases bound to the MPRs gather within coated membrane domains before budding in a different type of coated vesicle (8, 9). Recent evidence has revealed the existence of endosomal coated domains where protein sorting also takes place (10–12). In light of these observations, it was logical to hypothesize that sorting involved interactions

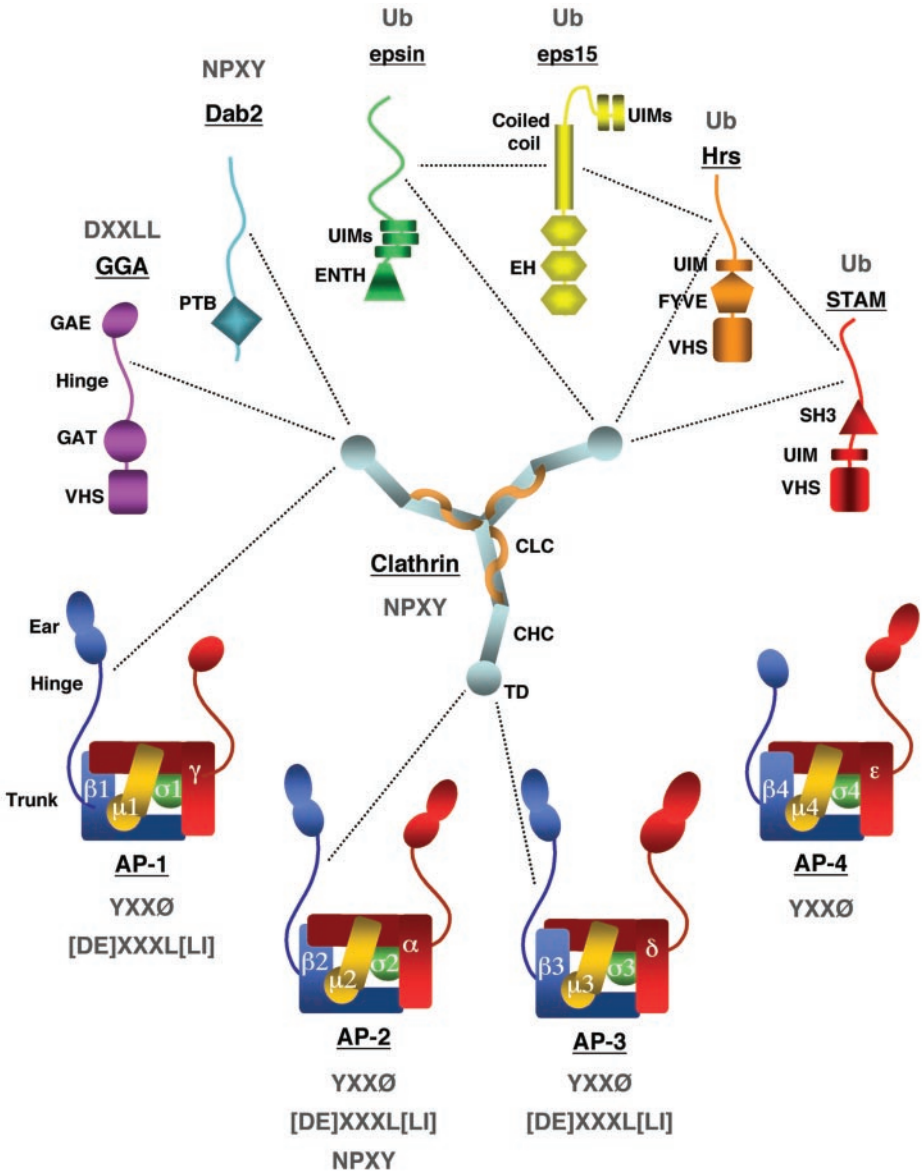
TABLE 2 Endocytic machinery motifs and their recognition domains

Motif	Proposed recognition protein or domain
DP[FW]	Ear domain of the α subunit of AP-2
FXDXF	Ear domain of the α subunit of AP-2
NPF	EH domains
LØXØ[DE]	Clathrin terminal domain
LLDLL	Clathrin terminal domain
PWDLW	Clathrin terminal domain

Motifs are denoted as indicated in the legend to Table 1. EH, eps15 homology.

between signals present within the cytosolic domains of the receptors and components of the protein coats (13). The formulation of this hypothesis provided the impetus for the isolation and biochemical characterization of coat proteins (Figure 2). We now know of the existence of several coats that function in the endosomal-lysosomal system. Clathrin coats are composed of the structural protein clathrin (14), the heterotetrameric adaptor protein (AP) complexes, AP-1, AP-2, or AP-3 (15), and various accessory factors (16). Plasma membrane clathrin coats contain AP-2 and accessory factors such as AP180, epsin 1, eps15, eps15R, intersectin 1, disabled-2 (Dab2), and others. TGN and endosomal clathrin coats contain AP-1 and/or the monomeric adaptors, GGA1, GGA2, and GGA3. AP-3 is also found on clathrin coats associated with endosomes, although this complex may be able to function in the absence of clathrin. Another type of clathrin coat associated with endosomes contains the protein Hrs but no AP complexes. Finally, a fourth AP complex, AP-4, appears to be part of a nonclathrin coat associated with the TGN. Because of their localization to sites of signal-mediated protein sorting (Figure 1), these coat proteins are considered prime candidates to function in signal recognition.

Figure 2 Clathrin and adaptors. The basic unit of clathrin is a "triskelion" composed of three heavy chains (CHC) and three light chains (CLC). At the amino terminus of each CHC there is globular domain known as the terminal domain (TD), which serves as a binding site for many adaptor proteins, as indicated in the figure. The structure of the adaptor-protein (AP) complexes AP-1/4 has been modeled after that of AP-2 (69) (Figure 4). This structure consists of a brick-like core comprising the trunk domains of the two large subunits plus the μ and σ subunits, with two hinge-like sequences that connect the core to two ear domains. AP-1/3 interact with clathrin whereas AP-4 does not. Dab2, the Golgi-localized, γ-ear-containing, ADP-ribosylation factor-binding proteins (GGAs), epsin, eps15, Hrs, and STAM1 are modular clathrin-associated proteins that may also function as adaptors. Their domain structure is indicated, but proteins are not drawn to scale. All of the above proteins are part of coats associated with cytosolic faces of membranes. The sorting signals proposed to interact with each of these proteins are indicated in gray letters.


Ub
epsin
Ub
eps15
NPXY
Dab2
Coiled coil
UIMs
Ub
Hrs
DXXLL
GGA
UIMs
ENTH
EH
UIM
FYVE
VHS
Ub
STAM
GAE
PTB
Hinge
GAT
VHS
SH3
UIM
VHS
CLC
Clathrin
NPXY
CHC
TD
Ear
Hinge
Trunk
β1
μ1
σ1
γ
AP-1
YXXØ
[DE]XXXL[LI]
β2
μ2
σ2
α
AP-2
YXXØ
[DE]XXXL[LI]
NPXY
β3
μ3
σ3
δ
AP-3
YXXØ
[DE]XXXL[LI]
β4
μ4
σ4
ε
AP-4
YXXØ

TYROSINE-BASED SORTING SIGNALS

Discovery of Tyrosine-Based Sorting Signals

The cloning of several genes encoding endocytic receptors in the early 1980s failed to reveal the presence of conserved sequences that could correspond to the hypothetical endocytic signals. This perplexing finding began to be explained with the discovery by the group of Brown & Goldstein that substitution of a cysteine codon for a tyrosine codon in the cytosolic domain of the LDL receptor, detected in a patient with familial hypercholesterolemia, abrogated the rapid internalization of the receptor (17). Later studies demonstrated that the critical tyrosine residue was part of the sequence motif NPXY (18), found not only in the LDL receptor but also in other cell surface proteins such as the LDL receptor-related protein 1 (LRP1), megalin, the β subunits of integrins, and the β-amyloid precursor protein (Table 3). Many other receptors known at the time, such as the transferrin receptor, the asialoglycoprotein receptor, and the cation-dependent and cation-independent mannose 6-phosphate receptors (CD- and CI-MPRs, respectively), however, lacked NPXY motifs but were rapidly internalized, suggesting the existence of other types of endocytic signals. An observation by Lazarovits & Roth suggested that tyrosine residues could nonetheless be key elements of these other signals (19). These investigators found that artificial replacement of a tyrosine residue for a cysteine residue in the cytosolic domain of influenza hemagglutinin (HA) enabled the protein to undergo rapid internalization via clathrin-coated pits. Several groups subsequently showed that substitution of tyrosine residues in the cytosolic domains of various endocytic receptors devoid of NPXY motifs impaired internalization (20–24). Systematic mutational analyses performed by Kornfeld and colleagues soon led to the definition of another tyrosine-based motif, YXXØ, as the major determinant of endocytosis of the MPRs (25, 26) as well as many other transmembrane proteins (Table 4). This motif is now known to mediate not only rapid internalization from the plasma membrane, but also targeting of certain proteins to lysosomes (27, 28) or to the basolateral plasma membrane domain of polarized epithelial cells (29, 30), sorting events that are thought to occur within clathrin-coated areas of the TGN or endosomes. The concept thus emerged of two types of tyrosine-based sorting signals represented by the NPXY and YXXØ consensus motifs.

NPXY-Type Signals

CHARACTERISTICS OF NPXY SIGNALS To date, NPXY signals have been shown to mediate only rapid internalization of a subset of type I integral membrane proteins and not other intracellular sorting events. These signals tend to occur in families such as members of the LDL receptor, integrin β, and β-amyloid precursor protein families (Table 3). They also seem to be conserved among different metazoans, including *Caenorhabditis elegans*, *Drosophila*, and mammals (Table 3), suggesting evolutionary conservation of the sorting mechanism

TABLE 3 NPXY-type signals

Protein	Species	Sequence
LDL receptor	Human	Tm-10-INFD**NP**V**Y**QKTT-29
LRP1 (1)	Human	Tm-21-VEIG**NP**T**Y**KMYE-64
LRP1 (2)	Human	Tm-55-TNFT**NP**V**Y**ATLY-33
LRP1	*Drosophila*	Tm-43-GNFA**NP**V**Y**ESMY-38
LRP1 (1)	*C. elegans*	Tm-54-TTFT**NP**V**Y**ELED-91
LRP1 (2)	*C. elegans*	Tm-140-LRVD**NP**L**Y**DPDS-4
Megalin (1)	Human	Tm-70-IIFE**NP**M**Y**SARD-125
Megalin (2)	Human	Tm-144-TNFE**NP**I**Y**AQME-53
Integrin β-1 (1)	Human	Tm-18-DTGE**NP**I**Y**KSAV-11
Integrin β-1 (2)	Human	Tm-30-TTVV**NP**K**Y**EGK
Integrin β (1)	*Drosophila*	Tm-26-WDTE**NP**I**Y**KQAT-11
Integrin β (2)	*Drosophila*	Tm-35-STFK**NP**M**Y**AGK
APLP1	Human	Tm-33-HGYE**NP**T**Y**RFLE-3
APP	Human	Tm-32-NGYE**NP**T**Y**KFFE-4
APP-like	*Drosophila*	Tm-38-NGYE**NP**T**Y**KYFE-3
Insulin receptor	Human	Tm-36-YASS**NP**E**Y**LSAS-379
EGR receptor (1)	Human	Tm-434-GSVQ**NP**V**Y**HNQP-96
EGR receptor (2)	Human	Tm-462-TAVG**NP**E**Y**LNTV-68
EGR receptor (3)	Human	Tm-496-ISLD**NP**D**Y**QQDF-34

Numbers in parentheses indicate motifs that are present in more than one copy within the same protein. The signals in this and other tables should be considered examples. Not all of these sequences have been shown to be active in sorting; some are included because of their conservation in members of the same protein family. Key residues are indicated in bold type. Numbers of amino acids before (i.e., amino-terminal) and after (i.e., carboxy-terminal) the signals are indicated. Abbreviations: Tm, transmembrane; LDL, low density lipoprotein; LRP1, LDL receptor related protein 1; APP, β-amyloid precursor protein; APLP1, APP-like protein 1.

in which they participate. Substitution of alanine for the N, P, or Y residues in the LDL receptor signal largely abolishes rapid endocytosis, whereas substitution of phenylalanine for tyrosine does not (18). Although NPXY is the minimal motif shared by all proteins with this type of signal (Table 3), alone it may not be sufficient for internalization. In the LDL receptor, a phenylalanine residue at two positions amino-terminal to the critical asparagine is also required for optimal internalization, indicating that the complete signal for this receptor is the sequence FDNPVY (18). Moreover, transplantation of only NPVY into the transferrin receptor fails to direct rapid endocytosis, but insertion of FDNPVY does (31). Similarly, internalization of the β-amyloid precursor protein is mediated by the longer GYENPTY sequence, in which the first of the two tyrosines is the most critical for function (32). Other proteins also have either a

TABLE 4 YXXØ-type signals

Protein	Species	Sequence
LAMP-1	Human	Tm-RKRSHAG**Y**QT**I**
LAMP-2a	Human	Tm-KHHHAG**Y**EQ**F**
LAMP-2a	Chicken	Tm-KKHHNTG**Y**EQ**F**
LAMP-2b	Chicken	Tm-RRKSRTG**Y**QS**V**
LAMP-2c	Chicken	Tm-RRKSYAG**Y**QT**L**
LAMP	*Drosophila*	Tm-RRRSTSRG**Y**MS**F**
LAMP	Earthworm	Tm-RKRSRRG**Y**ES**V**
CD63	Human	Tm-KSIRSG**Y**EV**M**
GMP-17	Human	Tm-HCGGPRPG**Y**ET**L**
GMP-17	Mouse	Tm-HCRTRRAE**Y**ET**L**
CD68	Human	Tm-RRRPSA**Y**QA**L**
CD1b	Human	Tm-RRRS**Y**QN**I**P
CD1c	Human	Tm-KKHCS**Y**QD**I**L
CD1d	Mouse	Tm-RRRSA**Y**QD**I**R
CD1	Rat	Tm-RKRRRS**Y**QD**I**M
Endolyn	Rat	Tm-KFCKSKERN**Y**HT**L**
Endolyn	*Drosophila*	Tm-KFYKARNERN**Y**HT**L**
TSC403	Human	Tm-KIRLRCQSSG**Y**QR**I**
TSC403	Mouse	Tm-KIRQRHQSSA**Y**QR**I**
Cystinosin	Human	Tm-HFCLYRKRPG**Y**DQ**L**N
Putative solute carrier	Human	Tm-12-SLSRGSG**Y**KE**I**
TRP-2	Human	Tm-RRLRKG**Y**TP**L**MET-11
HLA-DM β	Human	Tm-RRAGHSS**Y**TP**L**PGS-9
LmpA	Dictyostelium	Tm-KKLRQQKQQG**Y**QA**I**INNE
Putative lysosomal protein	Dictyostelium	Tm-RSKSNQNQS**Y**NL**I**QL
LIMP-II	Dictyostelium	Tm-RKTFYNNNQYNG**Y**NI**I**N
Transferrin receptor	Human	16-PLS**Y**TR**F**SLA-35-Tm
Asialoglycoprotein receptor H1	Human	MTKE**Y**QD**L**QHL-29-Tm
CI-MPR	Human	Tm-22-SYK**Y**SK**V**NKE-132
CD-MPR	Human	Tm-40-PAA**Y**RG**V**GDD-16
CTLA-4	Human	Tm-10-TGV**Y**VK**M**PPT-16
Furin	Human	Tm-17-LIS**Y**KG**L**PPE-29
TGN38	Rat	Tm-23-ASD**Y**QR**L**NLKL
gp41	HIV-1	Tm-13-RQG**Y**SP**L**SFQT-144
Acid phosphatase	Human	Tm-RMQAQPPG**Y**RH**V**ADGEDHA

See legends to Tables 1–3 for explanation of signal format.

phenylalanine or a tyrosine residue at this same position relative to the NPXY motif, although their functional importance has not been assessed (Table 3). These observations suggest that the most potent among these signals conform to the hexapeptide motif [FY]XNPXY.

NPXY signals are most often found within medium-length cytosolic domains ranging from ~40 to ~200 amino acid residues, although some are present within the larger cytosolic domains of signaling receptors such as the insulin receptor and the epidermal growth factor (EGF) receptor (Table 3). The distance of these signals from the transmembrane domains is variable, but none is closer than 10 residues (Table 3). They are also never found exactly at the carboxy-terminus of the proteins. Some proteins, most notably LRP-1 and megalin, contain two copies of this motif, suggesting the possibility of bivalent binding to their recognition proteins.

RECOGNITION PROTEINS FOR NPXY SIGNALS Ever since the discovery of NPXY signals, recognition proteins for these signals have been sought. As is often the case with this type of weak interaction, affinity purification approaches met with little success. The LDL receptor and other proteins having NPXY-type signals are internalized via clathrin-coated pits (6, 7), so clathrin-coat components are likely candidates for recognition proteins. Indeed, nuclear magnetic resonance (NMR) spectroscopy studies have shown that peptides containing the FDNPVY sequence from the LDL receptor bind directly to the globular, amino-terminal domain (TD) of the clathrin heavy chain (33) (Figure 2). These interactions exhibit the expected requirement for the critical tyrosine residue and occur with an equilibrium dissociation constant (K_D) of ~0.1 mM (33). The FDNPVY peptide also binds to clathrin cages, in which context the NPVY residues adopt a type 1 β-turn structure (33). A problem with the idea that clathrin functions as the primary recognition protein for NPXY signals is that clathrin is associated not only with the plasma membrane but also with the TGN and endosomes. Since NPXY signals only mediate internalization from the plasma membrane, it is unclear what would keep NPXY signals from engaging clathrin at other intracellular locations. Another caveat pointed out by Boll et al. (34) is that, because of its position within the cytosolic domain, the FDNPVY signal of the LDL receptor could extend at most 30–40Å from the transmembrane domain, whereas the terminal domain of clathrin lies at about 100Å from the membrane (35).

AP-2 and other putative adaptors (Figure 2) are situated between the clathrin lattice and the membrane and may therefore be in a better position to interact with endocytic signals. Early experiments by Pearse detected a weak interaction of AP-2 with a fusion protein containing the cytosolic domain of the LDL receptor, but the dependence of this interaction on the FDNPVY signal was not tested (36). More recently, Boll et al. (34) have reported the use of surface plasmon resonance (SPR) spectroscopy and photoaffinity labeling to demonstrate binding of FDNPVY peptides to purified AP-2. This binding is dependent on the phenylalanine and tyrosine residues of the signal. The very low affinity of these

interactions and the insolubility of the peptides at high concentrations have precluded an estimation of the K_D of the binding reaction. Interestingly, competition and mutational analyses suggest that NPXY and YXXØ signals bind to distinct sites on the μ2 subunit of AP-2 (34) (see below). This observation is in line with the finding that inducible overexpression of the FXNPXY-containing LDL receptor or the YXXØ-containing transferrin receptor does not slow the endocytic uptake of the other, even though each is able to compete with itself for incorporation into endocytic structures (37). These two receptors also saturate the endocytic machinery at different receptor densities, suggesting distinct recognition components (37).

Recent work has led to the alternative hypothesis that NPXY signals are not directly recognized by AP-2 but rather by proteins containing a domain known as phosphotyrosine-binding (PTB) or phosphotyrosine-interacting domain (PID). Although one biologic use of the PTB domain is to drive recruitment of signaling adaptors such as IRS-1 or Shc onto NPXpY motifs (pY stands for phosphotyrosine) on activated receptor tyrosine kinases, this domain demonstrates a marked plasticity in ligand recognition. Some PTB domain-containing proteins actually display a substantially higher selectivity for NPXY than for NPXpY (38, 39), and others bind sequences unrelated to this consensus motif. There is mounting evidence that PTB domain-containing proteins do function in LDL receptor internalization. Dab2 (Figure 2) is one good candidate, as this protein colocalizes extensively with clathrin coat components at the cell surface at steady state (39, 40). The PTB domain of Dab2 is ~65% identical to that of Dab1, a protein that participates in neural migration and cortical lamination in the brain (41). Dab1 operates in a signaling pathway directly downstream of the very low density lipoprotein (VLDL) receptor and apoE receptor 2 by engaging FXNPXY sequences in these receptors via the PTB domain (41). Similarly, the Dab2 PTB domain binds directly FXNPXY sequences from various LDL receptor family members in the nonphosphorylated form (39, 40, 42), and overexpression of either the Dab1 (43) or Dab2 (40) PTB domain causes the LDL receptor to back up at the cell surface, as its endocytosis is selectively impeded. Conditional disruption of *Dab2* in mice results in viable animals that are apparently healthy but display a proteinuria (44) that is similar to, but less severe than, that seen in megalin$^{-/-}$ animals (45). As the LDL receptor family member megalin plays a pivotal role in protein reabsorption in the renal proximal tubule (46), these in vivo data tie Dab2 to LDL receptor family endocytosis.

Targeting of Dab2 to the clathrin bud site is due to several endocytic interaction motifs, including DPF/FXDXF, LVDLN, and NPF sequences that engage AP-2, clathrin, and EH domain-containing proteins, respectively (Table 2) (39, 40). The PTB domain of Dab2 also binds phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P_2) via a surface separate from the FXNPXY binding site (40). These functional attributes are consistent with Dab2 being an intermediate sorting adaptor, but the mild phenotype of Dab2 nullizygous mice argues against it being singularly responsible for recognition of all NPXY-type inter-

nalization signals. In fact, there are other PTB domain proteins that could be functionally redundant with Dab2. ARH, for example, is a PTB domain protein that is defective in patients with an autosomal recessive form of hypercholesterolemia (47, 48). In cultured lymphoblasts from these patients, the LDL receptor stagnates at the plasma membrane, but the transferrin receptor continues to be internalized normally (49). This suggests that ARH is specifically involved in FXNPXY-driven endocytosis. The ARH PTB domain is most closely related to the PTB domains of Dab1/2, Numb, and GULP/Ced-6, all PTB domains that bind nonphosphorylated FXNPXY sequences preferentially. In *Drosophila*, Numb is involved in cell fate determination by segregating asymmetrically during cell division. Numb antagonizes Notch receptor signaling in the daughter cell into which it segregates by facilitating receptor endocytosis within clathrin-coated vesicles (50). The mammalian Numb orthologue, like Dab2, is located within clathrin-coated structures at steady state (51) and contains endocytic interaction motifs that facilitate interactions with AP-2 and eps15 (51).

YXXØ-Type Signals

CHARACTERISTICS OF YXXØ SIGNALS Although discovered after NPXY signals, YXXØ signals are now known to be much more widely involved in protein sorting. They are found in endocytic receptors such as the transferrin receptor and the asialoglycoprotein receptor, intracellular sorting receptors such as the CI- and CD-MPRs, lysosomal membrane proteins such as LAMP-1 and LAMP-2, and TGN proteins such as TGN38 and furin, as well as in proteins localized to specialized endosomal-lysosomal organelles such as antigen-processing compartments (e.g., HLA-DM) and cytotoxic granules (e.g., GMP-17) (Table 4). Each of these protein types has a distinct steady-state distribution, but they all share the property of trafficking via the plasma membrane to some extent. The YXXØ motifs are essential for the rapid internalization of these proteins from the plasma membrane. However, their function is not limited to endocytosis, since the same motifs have been implicated in the targeting of transmembrane proteins to lysosomes and lysosome-related organelles (27, 28, 52, 53), as well as the sorting of a subset of proteins to the basolateral plasma membrane of polarized epithelial cells (29, 30). The multiple roles of YXXØ signals suggest that they must be recognized not only at the plasma membrane but also at other sorting stations such as the TGN and endosomes. YXXØ signals are found in organisms as distantly related as protists and mammals (Table 4), suggesting that they participate in evolutionarily conserved mechanisms of sorting. As a matter of probability, sequences conforming to the YXXØ motif are common in the large cytosolic domains of signaling receptors and in cytosolic proteins. However, most of these sequences are not likely to be active as sorting signals because they are folded within the structure of the proteins and therefore are not accessible for interactions with components of the sorting machinery.

The YXXØ tetrapeptide is the minimal motif capable of conferring sorting information onto transmembrane proteins. However, the X residues and other

residues flanking the motif also contribute to the strength and fine specificity of the signals. The Y residue is essential for function and in most cases cannot even be substituted by other aromatic amino acid residues, suggesting that the phenolic hydroxyl group of the tyrosine is a critical recognition element. In this regard, YXXØ signals differ from NPXY signals, which tolerate phenylalanine in place of tyrosine [reviewed in (1, 17)]. The Ø position can accommodate several residues with bulky hydrophobic side chains, although the exact identity of this residue can specify the properties of the signal (53, 54). The X residues are highly variable but tend to be hydrophilic. A salient feature of YXXØ signals involved in lysosomal targeting is that most have a glycine residue preceding the critical tyrosine residue (Table 4). Mutation of this glycine to alanine decreases lysosomal targeting but not endocytosis (28), indicating that this residue is an important determinant of sorting to lysosomes. Lysosomal-sorting YXXØ motifs tend to have acidic residues at the X positions, and these may also contribute to the efficiency of lysosomal targeting (54).

Just as important as the actual amino acid sequence of YXXØ signals is their position within the cytosolic domains. These signals can be found in all types of transmembrane proteins, including Type I (e.g., LAMP-1 and LAMP-2), Type II (e.g., the transferrin and asialoglycoprotein receptors) and multispanning integral membrane proteins (e.g., CD63 and cystinosin). Purely endocytic YXXØ signals are most often situated at 10–40 residues from the transmembrane domains, but not at the carboxy-termini of the proteins (Table 4). This means that the carboxy-terminus of the Ø residue need not be free for function in endocytosis. In contrast, lysosomal-targeting YXXØ signals are conspicuous for their presence at 6–9 residues from the transmembrane domain and at the carboxy-terminus of the proteins (Table 4). In some proteins targeted to late endosomal or lysosomal compartments such as CD1b or cystinosin, the Ø residue is followed by only one more residue (Table 4). The importance of the distance from the transmembrane domain has been emphasized by a study showing that changing the spacing of the GYQTI signal from LAMP-1 impairs targeting to lysosomes (55). The mutant proteins continue to be internalized at the same rates but recycle to the plasma membrane, a behavior typical of endocytic receptors (55). These observations indicate that the placement of YXXØ signals at 6–9 residues from the transmembrane domain allows their recognition as lysosomal targeting signals at the TGN and/or endosomes.

RECOGNITION OF YXXØ SIGNALS BY THE μ SUBUNITS OF AP COMPLEXES As expected from morphological studies, YXXØ signals have been shown to interact with AP-1 and AP-2 (56–58). Recognition of YXXØ signals was originally ascribed to the large subunits of AP-2 complexes. Subsequent studies using the yeast two-hybrid system and in vitro binding assays revealed that YXXØ-signal recognition is instead a property of the μ subunits of the four AP complexes (56, 57, 59, 60). Binding to the μ subunits is now known to occur with apparent affinities in the 0.05–100 μM range (56, 61, 62) and to be strictly dependent on

the Y and Ø residues (56, 57, 59). Each μ subunit recognizes a distinct but overlapping set of YXXØ signals, as defined by the identity of residues other than the critical tyrosine (63, 64). The μ2 subunit of AP-2 exhibits the highest avidity and broadest specificity. These properties may allow sufficient leeway for μ2 to interact productively with YXXØ signals even when the placement of the signals or the identity of the X residues are suboptimal. This interpretation is supported by the fact that most YXXØ signals characterized to date function as endocytic signals. In addition, it has been shown that mutations in the X residues that substantially decrease interactions with μ2 have little effect on internalization (59). Moreover, mutations in or around the LAMP-1 signal that impair lysosomal targeting have minimal effects on internalization (28, 55). These observations suggest that the binding properties of μ2 endow the endocytic machinery with the ability to function efficiently with a wide variety of YXXØ signals.

Relative to μ2, μ1 (A and B isoforms), μ3 (A and B isoforms), and μ4 bind more weakly and display a narrower set of preferences for residues surrounding the critical tyrosine (63–65). μ1A preferences are relatively nondescript, whereas μ3A and μ3B exhibit clear preferences for acidic residues before and after the critical tyrosine (63). The most characteristic feature of μ4-binding signals is the presence of aromatic residues at various positions near the critical tyrosine residue (64). These properties may make interactions with μ1, μ3, and μ4 subunits more sensitive to variations in signal placement and amino acid sequence. Thus, these subunits, and by extension the AP complexes of which they are part, may be responsible for intracellular sorting events mediated by a subset of YXXØ signals. Exactly which sorting events are mediated by μ1, μ3, and μ4 subunits, however, cannot be inferred solely from their signal preferences. The μ3A and μ3B preferences hint at a role for AP-3 in protein sorting to lysosomes and lysosome-related organelles, since the signals from several proteins targeted to these organelles (i.e., LAMP-2a, CD63, and GMP-17) have acidic residues at the Y+1 position. This presumption is supported by genetic evidence linking AP-3 to the biogenesis of lysosome-related organelles in various organisms [reviewed in (66)]. A confounding fact is that, for the most part, none of the μ subunits analyzed shows a strong preference for a glycine preceding the critical tyrosine, a characteristic of lysosomal targeting signals. This suggests either that GYXXØ signals are recognized more specifically by another type of protein, or that the glycine is present for reasons other than recognition specificity. In this regard, the absence of a side chain in the glycine residue could allow more flexible or unhindered recognition of the YXXØ residues in close proximity to the membrane.

STRUCTURAL BASES FOR THE RECOGNITION OF YXXØ SIGNALS Yeast two-hybrid, proteolytic digestion, and crystallographic analyses have shown that the μ subunits are organized into an amino-terminal domain comprising one third of the protein and a carboxy-terminal domain comprising the remaining two thirds

(67–69). The amino-terminal domain mediates assembly with the β subunits of the AP complexes, whereas the carboxy-terminal domain harbors the binding site for YXXØ signals. This latter domain exhibits homology to the carboxy-terminal domain of *Drosophila* Stoned-B and the mammalian stonin 1 and stonin 2 proteins (70–72); in these proteins, however, this domain does not bind YXXØ signals (71). The crystal structure of the amino-terminal domain of μ2 in the context of the AP-2 core consists of a five-stranded β-sheet flanked on either side by α-helices (69). This domain makes extensive hydrophobic contacts with the trunk domain of β2 (69). The crystal structure of the carboxy-terminal domain of μ2 complexed with various YXXØ signal peptides has also been solved. This domain exhibits a banana-shaped, 16-strand β-sheet structure organized into two β-sandwich subdomains (*A* and *B*) (68) (Figure 3*A*). The YXXØ signal peptide binds in an extended conformation to β-sheet strands 1 and 16 of subdomain *A* and forms an extra β-strand paired with β-strand 16 of μ2. The critical Y and Ø residues fit into two hydrophobic pockets on the surface of the protein (Figure 3*B*). The Y residue interacts with residues lining the interior of the pocket via both its aromatic ring and phenolic hydroxyl group. This latter group participates in a network of hydrogen bonds with various μ2 residues, including ^{176}Asp located at the bottom of the pocket. The absence of this phenolic hydroxyl group in phenylalanine explains why this residue cannot substitute for tyrosine in YXXØ signals. The pocket for the Ø residue is lined with flexible aliphatic side chains that allow accommodation of structurally diverse bulky hydrophobic residues (68). The X residues can also be involved in backbone interactions with μ2, thus contributing to the fine specificity of binding (68). Residues amino-terminal to the critical tyrosine can provide additional points of attachment to μ2. For example, a leucine residue at position Y-3 from a YXXØ signal in P-selectin binds to a separate hydrophobic pocket on the μ2 surface (73) (Figure 3*C*). On the basis of these observations, interactions of YXXØ signals with μ2 have been likened to either two-pinned (i.e., most signals, Figure 3*B*) or three-pinned (i.e., P-selectin, Figure 3*C*) plugs fitting into complementary sockets (68, 73). The three-pin configuration may allow distribution of interaction forces over a larger surface area, such that internalization activity is less dependent on any one particular residue. Because of their homology and the conservation of residues involved in YXXØ recognition (3), other μ subunits are expected to have a similar structure and mechanism of binding.

The physiological role of μ2 in YXXØ signal recognition and the identity of the residues involved in recognition have been corroborated using a dominant-negative interference approach (74). Mutant μ2 constructs with substitutions of ^{176}Asp and ^{421}Trp are unable to bind YXXØ signals. Upon transfection into cells, these constructs are incorporated into AP-2 and impair the rapid internalization of the transferrin receptor (74).

The recent resolution of the crystal structure of the AP-2 core (69) has uncovered an unexpected complexity in the mechanism of YXXØ-signal recognition. This structure shows that the carboxy-terminal domain of μ2 rests on a

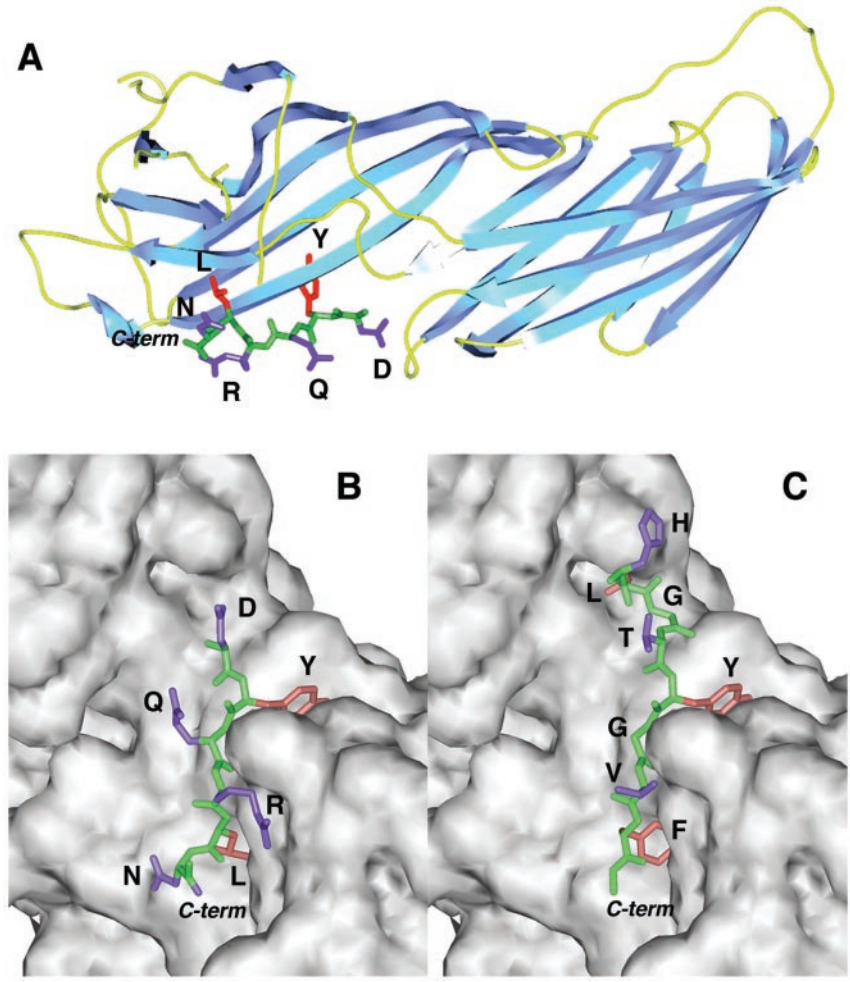


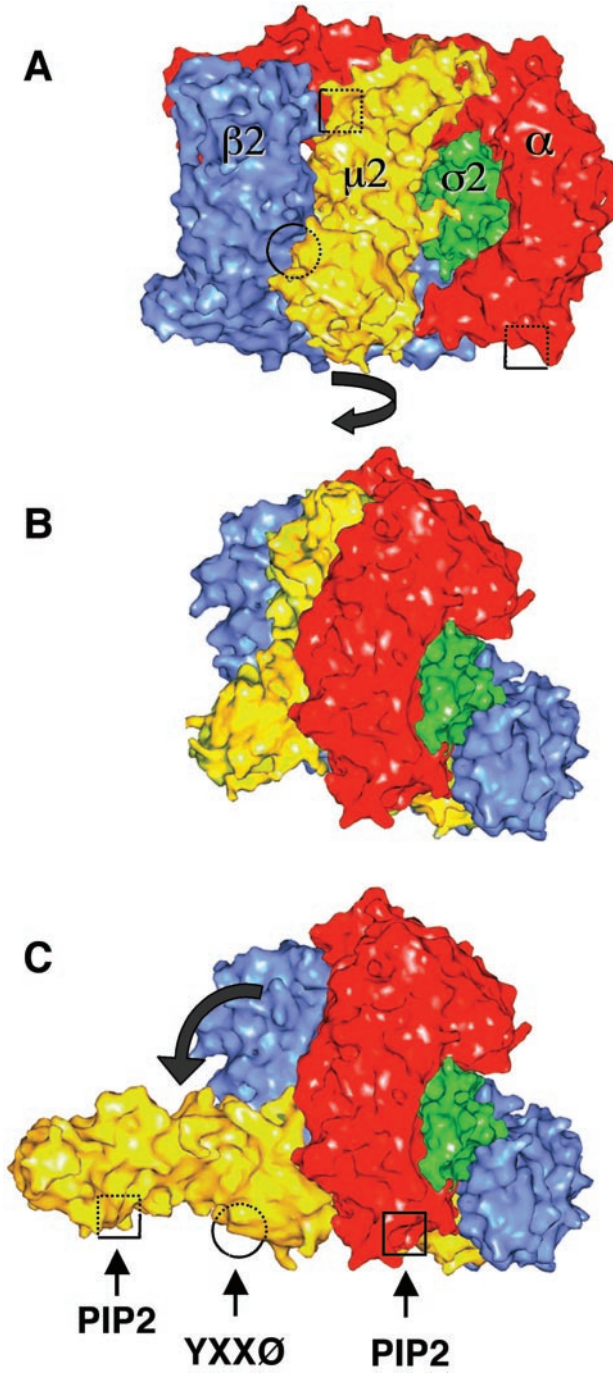
Figure 3 Structure of YXXØ-μ2 complexes. (*A*) Ribbon diagram of the carboxy-terminal domain of μ2 in a complex with the DYQRLN peptide from TGN38. Notice the orientation of the Y and L residues toward μ2. (*B*) Binding of the DYQRLN peptide from TGN38 to the surface of μ2. The peptide binds in an extended conformation with the Y and L residues fitting into two hydrophobic pockets (i.e., the "two-pinned" plug model). (*C*) Binding of the HLGTYGVF peptide from P-selectin to the surface of μ2. This peptide binds also in an extended conformation with the L, Y, and F residues fitting into three hydrophobic pockets (i.e., the "three-pinned" plug model). The peptide backbone is indicated in green, the key side chains contacting μ2 in red, and other side chains in purple. Data are from (68, 73).

furrow formed by the trunk domains of the α and β2 subunits (Figure 4, *A* and *B*). In this structure, the YXXØ-binding site on μ2 is partially occluded by an interaction with the β2-trunk domain. The solved structure thus corresponds to an inactive form of AP-2, which must be activated by a conformational change. This change may involve displacement of the carboxy-terminal domain of μ2 from the AP-2 core and its attachment to the membrane via a phosphoinositide-binding site on subdomain *B* (69, 75) (Figure 4*C*). Such a change would place the YXXØ-binding site in a favorable position to interact with sorting signals, especially with those from lysosomal membrane proteins, which are located close to the membrane.

REGULATION OF YXXØ-μ2 INTERACTIONS The large conformational change required to activate AP-2 for YXXØ-signal binding is likely a major target for regulation. Indeed, a threonine residue (^{156}Thr) in the linker sequence that connects the amino-terminal and carboxy-terminal domains of μ2 is phosphorylated (76, 77) specifically by a recently discovered serine/threonine kinase, AAK1 (78, 79). This kinase binds to the ear domain of the α subunit of AP-2 and largely colocalizes with this complex in clathrin-coated pits and vesicles (78). Phosphorylation of ^{156}Thr enhances the affinity of AP-2 for YXXØ signals by about one order of magnitude (79) and is required for normal receptor-mediated endocytosis of transferrin (77, 78). Thus, cycles of phosphorylation/dephosphorylation may allow for multiple rounds of cargo recruitment into forming clathrin-coated pits. How this phosphorylation could trigger release of the μ2 carboxy-terminal domain from the core is not apparent from the structure. In addition to phosphorylation, other factors such as electrostatic attraction between the positively charged μ2 surface and the negatively charged membrane surface could contribute to the release of μ2 required for YXXØ-signal binding.

Localized changes in the levels of phosphoinositides [PtdIns(4,5)P_2 and PtdIns(3,4,5)P_3] by specific lipid kinases (80) and phosphatases (81) might also modulate YXXØ-μ2 interactions. Studies have in fact shown that the interaction of AP-2 with phosphoinositides enhances the recognition of YXXØ signals (62). This modulation likely involves phosphoinositide-binding sites present on both μ2 (69, 75) and the amino-terminal portion of the α-trunk domain (69, 82).

Figure 4 Structure of the AP-2 core. (*A*) Surface representation of the AP-2 core with its four subunits indicated in different colors. Two sites for polyphosphatidylinositol (PIP2/PIP3) binding, one on μ2 and the other on α, are indicated by the squares. The partially occluded YXXØ-binding site on μ2 is indicated by the circle. (*B*) Rotated view of the structure shown in *A*. (*C*) Hypothetical structure of AP-2 activated for YXXØ binding. The carboxy-terminal domain of μ2 is shown displaced from its original position in the core. This conformational change would position the polyphosphatidylinositol-binding and YXXØ-binding sites on μ2 close to the membrane, where they can interact with their corresponding ligands. Data are from (69).


A
β2
μ2
σ2
α
B
C
PIP2
YXXØ
PIP2

Additional regulation might be provided by interaction of AP-2 with putative docking factors such as synaptotagmin (83). A general model incorporating all of these regulatory factors could posit that AP-2 exists in equilibrium between two conformers defined by the position of the μ2 carboxy-terminal domain: an unphosphorylated, inactive form and a phosphorylated, active form capable of binding YXXØ signals, phosphoinositides, and synaptotagmin via the μ2 carboxy-terminal domain. Phosphorylation and/or binding of any of these ligands would shift the equilibrium toward the open state of AP-2, resulting in a highly cooperative activation of the complex.

Phosphorylation of the YXXØ signals or adjacent residues can also modulate their recognition by μ2. The T cell costimulatory receptor CTLA-4, for example, has a YVKM sequence that binds to μ2 and mediates rapid internalization of the protein in resting T cells (84–86). Upon activation of the T cells, the Y residue of the signal becomes phosphorylated. This phosphorylation blocks interaction with μ2 and inhibits internalization, while at the same time allowing the recruitment of signaling molecules to the phosphotyrosine residue (84–86). Phosphorylation of a single tyrosine residue thus serves as a regulatory switch that determines whether the protein is removed from the plasma membrane in resting T cells or remains at the cell surface to transduce signals in activated T cells. A similar regulatory process has been proposed to operate for the neural cell adhesion protein L1 (87). The negative effect of tyrosine phosphorylation on interactions with μ2 can be easily explained by the impossibility of accommodating a large and negatively charged phosphate group into the Y-binding pocket. In contrast to this negative regulation, phosphorylation by casein kinase II of a serine residue immediately preceding the critical tyrosine residue in the YXXØ signal of aquaporin 4 (AQP4) enhances interactions with μ3A and lysosomal targeting of the protein (88). The structural bases for this effect, however, remain to be elucidated.

DILEUCINE-BASED SORTING SIGNALS

Discovery of Dileucine-Based Sorting Signals

At a time when the field of intracellular protein sorting was focused on the study of tyrosine-based sorting signals in endocytosis and lysosomal targeting came the surprising discovery of dileucine-based sorting signals. In the course of a deletion analysis of the cytosolic domain of the CD3-γ chain of the T-cell antigen receptor, Letourneur & Klausner noticed that a particular segment of this domain could confer on a reporter protein the ability to be rapidly internalized and delivered to lysosomes, even though it lacked any tyrosine residues (89). Further deletion analyses revealed that this segment contained a DKQTLL sequence that was responsible for those activities. An alanine scan mutagenesis showed that both leucine residues were required, whereas the other residues were dispensable

for function (89). Shortly after the publication of these observations, Johnson & Kornfeld (90, 91) reported that deletion of the LLHV and HLLPM sequences from the carboxy-termini of the CI- and CD-MPRs, respectively, impaired sorting of the MPRs from the TGN to the endosomal system. Both leucines as well as a cluster of acidic amino acid residues preceding the leucines were found to be important for sorting (90–92). Subsequent studies uncovered the presence of dileucine-based sorting signals in many other transmembrane proteins (Tables 5 and 6) and demonstrated that these signals have as broad a range of functions as that of tyrosine-based sorting signals. All of these signals had come to be thought of as variants of the same family. Recent investigations into the nature of their recognition proteins, however, have revealed that the prototypical CD3-γ and MPR signals correspond to two distinct classes represented by the motifs, [DE]XXXL[LI] and DXXLL, respectively.

[DE]XXXL[LI]-Type Signals

CHARACTERISTICS OF [DE]XXXL[LI] SIGNALS [DE]XXXL[LI] signals play critical roles in the sorting of many type I, type II, and multispanning transmembrane proteins (Table 5). The DKQTLL sequence of the CD3-γ chain is now known to be part of a regulatable SDKQTLL signal that participates in serine phosphorylation-dependent downregulation of the T-cell antigen receptor from the cell surface, a process that involves rapid internalization and lysosomal degradation of the receptor (93, 94). The CD4 coreceptor protein undergoes a similar downregulation upon phosphorylation of a serine residue within its SQIKRLL signal (95). Other transmembrane proteins that contain constitutively active forms of this signal are mainly localized to late endosomes and lysosomes (e.g., NPC1, LIMP-II), as well as to specialized endosomal-lysosomal compartments such as endocytic antigen-processing compartments (e.g., Ii), synaptic dense-core granules (VMAT1, VMAT2), stimulus-responsive storage vesicles (e.g., GLUT4, AQP4), and premelanosomes and melanosomes (e.g., tyrosinase, TRP-1, Pmel17, QNR-71). The Nef gene product of human immunodeficiency virus (HIV-1) has also been shown to contain a [DE]XXXL[LI] signal that participates in downregulation of CD4 (96, 97). This type of signal appears to be conserved throughout the animal and protist kingdoms since it is found not only in mammals but also in birds, fish, *C. elegans*, and yeast (Table 5). For *Saccharomyces cerevisiae* alkaline phosphatase, a [DE]XXXL[LI]-type signal mediates sorting of alkaline phosphatase to the vacuole, the yeast counterpart of metazoan lysosomes (98). Like YXXØ signals, the [DE]XXXL[LI] signals in mammalian proteins mediate rapid internalization and targeting to endosomal-lysosomal compartments, suggesting that they can be recognized both at the plasma membrane and intracellular locations. Sequences conforming to this motif have also been implicated in basolateral targeting in polarized epithelial cells (99, 100).

Substitution of either of the critical leucines by alanine abrogates all activities of [DE]XXXL[LI] signals (89, 101). The [DE]XXXL[LI] motif highlights the

TABLE 5 [DE]XXX[LI]-type signals

Protein	Species	Signal
CD3-γ	Human	Tm-8-S**D**KQT**LL**PN-26
LIMP-II	Rat	Tm-11-D**E**RAP**LI**RT
Nmb	Human	Tm-37-Q**E**KDP**LL**KN-7
QNR-71	Quail	Tm-37-T**E**RNP**LL**KS-5
Pmel17	Human	Tm-33-G**E**NSP**LL**SG-3
Tyrosinase	Human	Tm-8-E**E**KQP**LL**ME-12
Tyrosinase	Medaka fish	Tm-16-G**E**RQP**LL**QS-13
Tyrosinase	Chicken	Tm-8-P**E**IQP**LL**TE-13
TRP-1	Goldfish	Tm-7-E**G**RQP**LL**GD-15
TRP-1	Human	Tm-7-E**A**NQP**LL**TD-20
TRP-1	Chicken	Tm-7-E**L**HQP**LL**TD-20
TRP-2	Zebrafish	Tm-5-R**E**FEP**LL**NA-11
VMAT2	Human	Tm-6-E**E**KMA**IL**MD-29
VMAT1	Human	Tm-6-E**E**KLA**IL**SQ-32
VAchT	Mouse	Tm-10-S**E**RDV**LL**DE-42
VAMP4	Human	19-S**E**RRN**LL**ED-88-Tm
Neonatal FcR	Rat	Tm-16-D**D**SGD**LL**PG-19
CD4	Human	Tm-12-S**Q**IKR**LL**SE-17
CD4	Cat	Tm-12-S**H**IKR**LL**SE-17
GLUT4	Mouse	Tm-17-R**R**TPS**LL**EQ-17
GLUT4	Human	Tm-17-H**R**TPS**LL**EQ-17
IRAP	Rat	46-EP**R**GSR**LL**VR-53-Tm
Ii	Human	MD**D**QRD**LI**SNNEQLP**ML**GR-11-Tm
Ii	Mouse	MD**D**QRD**LI**SNHEQLP**IL**GN-10-Tm
Ii	Chicken	MAE**E**QRD**LI**SSDGSSG**VL**PI-12-Tm
Ii-1	Zebrafish	MEPDH**Q**NES**LI**QRVPSAET**IL**GR-12-Tm
Ii-2	Zebrafish	MSSEG**N**ETP**LI**SDQSSVNMGPQP-8-Tm
Lamp	Trypanosome	Tm-RPRRRT**E**EDE**LL**PE**E**AEG**LI**DPQN
Menkes protein	Human	Tm-74-P**D**KHS**LL**VGDFREDDDTAL
NPC1	Human	Tm-13-T**E**RER**LL**NF
AQP4	Human	Tm-32-V**E**TDD**LI**L-29
RME-2	*C. elegans*	Tm-104-F**E**NDS**LL**
Vam3p	*S. cerevisiae*	153-N**E**QSP**LL**HN-121-Tm
ALP	*S. cerevisiae*	7-S**E**QTR**LV**P-18-Tm
Gap1p	*S. cerevisiae*	Tm-23-E**V**DLD**LL**K-24

See legends to Tables 1–3 for explanation of signal format.

residues found in the most active among these signals. An acidic residue at position −4 from the first leucine appears to be important for targeting to late endosomes or lysosomes, though not for internalization (101, 102). The first of the two leucines is generally invariant, probably because substitution by other amino acids, including isoleucine, greatly decreases the potency of the signal (89). The second leucine, in contrast, can be replaced by isoleucine without loss of activity (89). In some cases, another acidic residue or a phosphoacceptor serine further amino-terminal to the [DE]XXXL[LI] motif adds to the strength of the signals. Some [DE]XXXL[LI] signals have arginine residues in place of the acidic residues, as is the case for signals from GLUT4 and IRAP, two proteins localized to insulin-regulated storage compartments (Table 5). This difference may bear physiological significance, since replacement of two glutamates for the two arginines in the RRTPSLL signal of GLUT4 impairs its internalization and sorting to storage compartments (102). Conversely, substitution of two arginines for the aspartate-glutamate pair in the DERAPLI signal of LIMP-II impairs internalization and lysosomal targeting (102). Hence, various [DE]XXXL[LI] signals may be recognized differently at different intracellular locales.

As is the case for YXXØ signals, the position of [DE]XXXL[LI] signals relative to the transmembrane domains and to the carboxy or amino termini also appears to influence the function of the signals. In proteins targeted to late endosomes or lysosomes (e.g., NPC1, LIMP-II), synaptic dense-core granules (e.g., VMAT1 and VMAT2), and premelanosomes or melanosomes (e.g., tyrosinase, TRP-1), the signals are very close to the transmembrane domain (i.e., 6–11 residues away). Late endosomal or lysosomal proteins also tend to display their signals near their carboxy (e.g., NPC1, LIMP-II) or amino termini (e.g., Ii). These properties are similar to those of lysosomal YXXØ signals, suggesting that both are subject to the same positional requirements for lysosomal targeting. Also in this case, proximity to the transmembrane domain and to the carboxy or amino terminus may enable optimal binding to specific recognition proteins. Indeed, a minimum of 6–7 amino acids from the transmembrane domains has been experimentally demonstrated to be optimal for downregulation of CD3-γ chimeras (103). The presence of a few residues carboxy-terminal to the second leucine or isoleucine, though common in naturally occurring signals, does not appear to be essential for function, since a chimeric protein bearing the DKQTLL sequence from CD3-γ at the carboxy terminus can undergo rapid internalization and lysosomal targeting (89).

RECOGNITION OF [DE]XXXL[LI] SIGNALS BY AP COMPLEXES Given the functional similarities of YXXØ- and [DE]XXXL[LI] signals, it is not surprising that [DE]XXXL[LI] signals have also been found to bind AP complexes in various in vitro assays (96, 104–109). This binding is dependent on the LL or LI pairs and, in some cases, on the acidic residues at positions −4 and −5 from the first leucine, thus paralleling the sequence requirements for function of the signals. Each [DE]XXXL[LI] signal exhibits distinct preferences for different AP com-

plexes. For example, the DDQRDLI and NEQLPML signals of Ii bind to AP-1 and AP-2, but not detectably to AP-3 (108). In contrast, the DERAPLI signal of LIMP-II and the EEKQPLL signal of tyrosinase bind to AP-3, but not to AP-1 or AP-2 (106). This binding specificity agrees with the observation that LIMP-II (110) and tyrosinase (111), but not class II MHC-associated Ii (112), are missorted in AP-3-deficient cells. Similarly to YXXØ signals, the fine specificity of interactions of [DE]XXXL[LI] signals may be dictated by the X residues or other contextual factors (113).

In vivo overexpression of transmembrane proteins bearing YXXØ- or [DE]XXXL[LI] signals has been shown to saturate the corresponding sorting machineries, causing missorting of proteins that have the same type of signal (114). However, YXXØ signals do not compete with [DE]XXXL[LI] signals and vice versa (114). This indicates that [DE]XXXL[LI] signals do not bind to the same site as YXXØ signals on μ2. Indeed, in vitro binding and yeast two-hybrid analyses have failed to demonstrate interactions of various [DE]XXXL[LI] signals with the carboxy-terminal domain of μ2 (56, 57, 59). Similar analyses, however, have documented interactions of [DE]XXXL[LI] signals from Ii and HIV-1 Nef with full-length μ1, μ2, and μ3A (115–117), and a phage display screen has identified the 119–123 segment of μ2 as a binding site for the Ii signals (118). Photoaffinity labeling analyses, in contrast, have demonstrated an interaction of various [DE]XXXL[LI] signals with the trunk domains of the β1 and β2 subunits of AP-1 and AP-2, respectively (119). Locating the [DE]XXXL[LI]-binding site on the AP complexes will ultimately require structural and mutational analyses of the kind that have been performed for YXXØ-μ2 interactions. Curiously, the purified AP-2 core could not be cocrystallized with any of several [DE]XXXL[LI] peptides (69), suggesting that the binding site may inaccessible. In any event, given the apparent functional diversity of [DE]XXXL[LI] signals, it cannot be ruled out that proteins other then AP complexes are involved in the recognition of certain subsets of signals. The recent discovery of a different kind of recognition protein for DXXLL signals suggests that this is a definite possibility.

DXXLL-Type Signals

CHARACTERISTICS OF DXXLL SIGNALS It has only recently become evident that DXXLL signals constitute a distinct type of dileucine-based sorting signals. These signals are present in several transmembrane receptors and other proteins that cycle between the TGN and endosomes, such as the CI- and CD-MPRs, sortilin, the LDL-receptor-related proteins LRP3 and LRP10, and β-secretase (Table 6). They also seem to be conserved in all metazoans. For the CI- and CD-MPRs, DXXLL signals appear to mediate incorporation into clathrin-coated vesicles that bud from the TGN for transport to the endosomal system (91, 120–122). The requirement for the D and LL residues in these signals is quite

TABLE 6 DXXLL-type dileucine-based signals

Protein	Species	Sequence
CI-MPR	Human	Tm-151-SFHDDSDEDLLHI
CI-MPR	Bovine	Tm-150-TFHDDSDEDLLHV
CI-MPR	Rabbit	Tm-151-SFHDDSDEDLLNI
CI-MPR	Chicken	Tm-148-SFHDDSDEDLLNV
CD-MPR	Human	Tm-54-EESEERDDHLLPM
CD-MPR	Chicken	Tm-54-DESEERDDHLLPM
Sortilin	Human	Tm-41-GYHDDSDEDLLE
SorLA	Human	Tm-41-ITGFSDDVPMVIA
Head-activator BP	Hydra	Tm-41-INRFSDDEPLVVA
LRP3	Human	Tm-237-MLEASDDEALLVC
ST7	Human	Tm-330-KNETSDDEALLLC
LRP10	Mouse	Tm-235-WVVEAEDEPLLA
LRP10	Human	Tm-237-WVAEAEDEPLLT
Beta-secretase	Human	Tm-9-HDDFADDISLLK
Mucolipin-1	Mouse	Tm-43-GRDSPEDHSLLVN
Nonclassical MHC-I	Deer mouse	Tm-6-VRCHPEDDRLLG
FLJ30532	Human	Tm-83-HRVSQDDLDLLTS
GGA1	Human	350-ASVSLLDDELMSL-275
GGA1	Human	415-ASSGLDDLDLLGK-211
GGA2	Human	408-VQNPSADRNLLDL-192
GGA3	Human	384-NALSWLDEELCL-326
GGA	*Drosophila*	447-TVDSIDDVPLLSD-116

See legends to Tables 1–3 for explanation of signal format. Serine and threonine residues are underlined.

strict since mutation of any of these residues to alanine inactivates the signals and results in increased expression of the transmembrane proteins at the cell surface (91, 92). The D position does not even tolerate isoelectric or isosteric substitutions without drastic loss of activity (92). In contrast, the X residues or other residues amino-terminal to the critical D are less important for function (92). Given these distinct requirements, it is noteworthy that the D residue is generally found in the context of a cluster of acidic residues (Table 6). Because of this, these signals are also referred to as acidic cluster-dileucine signals. Another feature of these signals is the presence of one or more serine residues upstream of the acidic cluster (Table 6). Also of note is the fact that most DXXLL signals are separated by one or two variable residues from the carboxy-termini of the proteins. The distance of the signals from the transmembrane domain, on the other hand, is longer and more variable. In the case of the CD-MPR, this distance

may be shortened in the cell by palmitoylation of two cysteine residues in the cytosolic domain (123). The palmitoyl chains likely get inserted into the lipid bilayer, pulling the rest of the cytosolic domain closer to the membrane (123).

RECOGNITION OF DXXLL SIGNALS BY THE VHS DOMAIN OF THE GGAs Unlike [DE]XXXL[LI] signals, DXXLL signals do not detectably bind to AP complexes (120, 124, 125). Instead, DXXLL signals from sortilin, the CI- and CD-MPRs, LRP3, SorLa, and β-secretase bind to the mammalian GGAs (120, 125–130), a recently described family of ADP-ribosylation factor (ARF)-dependent clathrin adaptors localized to the TGN and endosomes (131–134) (Figure 2). Molecular dissection of the GGAs has revealed that the DXXLL-binding activity resides within their amino-terminal VHS domains (120, 125–127). The VHS domains of the three human GGAs (i.e., GGA1, GGA2 and GGA3) bind DXXLL signals from various proteins with affinities ranging from 5 to 100 μM (135, 136). Despite differences in affinity, no sequence preferences for particular subsets of DXXLL signals are apparent among the three human GGAs. The VHS domains of other proteins, such as Hrs (Figure 2), STAM1, TOM1, and TOM1L1 do not bind DXXLL signals (120). Furthermore, the VHS domains of the GGAs do not bind [DE]XXXL[LI] or YXXØ signals (120). These observations emphasize the high degree of specificity of GGA-DXXLL signal interactions. In line with the sequence requirements for function in vivo, the D and LL residues, but not the X residues, of the signals are essential for interactions (120, 125, 127, 136). The physiological significance of these interactions is underscored by the findings that the CD-MPR and the GGAs exit the TGN on the same vesicular intermediates (120) and that a dominant-negative GGA construct causes retention of the CI- and CD-MPRs at the TGN (120). The mammalian GGAs are thus likely involved in the sorting of transmembrane proteins having DXXLL signals from the TGN to endosomes. In the case of the MPRs, this sorting is critical for the efficient delivery of acidic hydrolases to lysosomes. In line with this notion, the *S. cerevisiae* GGAs, Gga1p and Gga2p, have been functionally implicated in sorting hydrolases to the vacuole (132, 133, 137–139), although to date the signals that bind to the yeast GGAs have not been identified.

STRUCTURAL BASES FOR DXXLL-VHS DOMAIN INTERACTIONS Recent studies have solved the crystal structures of the VHS domains of GGA3 and GGA1 in complexes with DXXLL signal peptides from the CI- and CD-MPRs (135, 140) (Figure 5). The VHS domains of both proteins were found to be right-handed superhelices of eight helices (Figure 5*A*). Both the CI-MPR and CD-MPR peptides bind in an extended conformation to a groove formed by helices 6 and 8 (135, 140) (Figure 5*A*). The critical D and LL residues bind to an electropositive and two shallow hydrophobic pockets, respectively, whereas the X residues and flanking residues either point away from the VHS domain or are disordered (Figure 5*B*). The terminal carboxyl group appears to interact weakly with residues on the VHS domain (135, 140), which might explain why most DXXLL

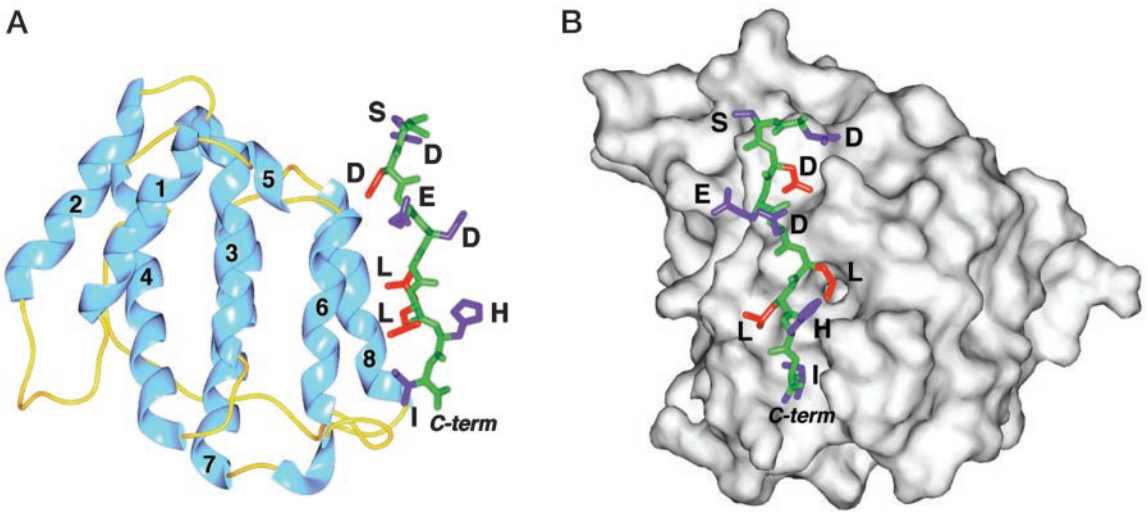


Figure 5 Structure of the complex between the GGA3 VHS domain and the SDEDLLHI signal from the CI-MPR. (*A*) Ribbon diagram showing the peptide binding in an extended conformation to helices 6 and 8 of the VHS domain. Note the position of the key D and LL residues facing the VHS domain. (*B*) Surface representation showing the key D and LL residues binding to pockets on the VHS domain. The peptide backbone is indicated in green, the key side chains contacting the VHS domain in red, and other side chains in purple. Graphics courtesy of Saurav Misra and James Hurley, NIDDK.

signals are positioned one or two residues away from the carboxy-terminus (Table 6). The identity of the DXXLL residues that contact the VHS domain is entirely consistent with the sequence requirements for the function of these signals. The structure of these complexes reveals that the failure of [DE]XXX-L[LI] signals to bind to the VHS domains is due, at least in part, to the placement of their acidic residue at position −4, instead of −3, from the first leucine. In addition, the structures of the VHS domains show that the pocket for the aspartate residue is too small to accommodate a bulkier glutamate residue, in agreement with the functional requirements of the signals (92). Thus, these structural studies provide compelling evidence for the distinct nature of [DE]XXXL[LI] and DXXLL signals. The VHS residues involved in interactions with the D and LL residues are completely conserved in the three human GGAs and mostly conserved in the GGAs in *Drosophila* and *C. elegans*, suggesting that all of these proteins may be able to bind similar signals. This conservation does not extend to the two GGAs in *S. cerevisiae*, indicating that their participation in vacuolar protein sorting probably involves recognition of a different type of signal.

REGULATION OF THE RECOGNITION OF DXXLL SIGNALS BY THE GGAs As mentioned above, serine residues are often found one to three positions amino-terminal to DXXLL signals. These serines are in a sequence context that fits the [ST]XX[DE] consensus motif for phosphorylation by casein kinase II (CKII), in which the X residues are generally acidic. For the CI- and CD-MPR, these serines have been shown to be phosphorylated both in vivo and in vitro by CKII or a CKII-like kinase (141–144). Although the physiological relevance of these phosphorylation events is still unclear (91, 92, 145), mutation of the serine residue upstream of the CI-MPR signal to alanine decreases the sorting efficiency of the receptor (92) and the interaction of its DXXLL signal with VHS domains (127, 136). Moreover, the in vitro binding affinity of a DXXLL signal containing peptide from the CI-MPR to GGA VHS domain increases ~threefold by phosphorylation of the serine residue (136). Crystallographic analyses have revealed that phosphorylation of this serine allows electrostatic interactions between one of the negatively charged phosphate oxygens and two positively charged residues on the VHS domain (136). The acidic cluster may therefore be conserved to provide a CKII recognition site for this type of regulation, rather than to interact with basic regions on the GGA VHS domain. An identically positioned serine residue next to a DXXLL signal is likely to afford the same type of regulation for sortilin (Table 6). The serine residues in other proteins are further amino-terminal from the DXXLL signals (Table 6); if and how they regulate signal recognition is not yet known.

Another form of regulation of DXXLL-GGA interactions has recently been uncovered. GGA1 and GGA3 have DXXLL motifs within their hinge domains (Table 6), which are able to bind to their own VHS domains (146). This binding is thought to be auto-inhibitory and to account for the poor binding activity of the full-length GGA1 and GGA3 relative to their isolated VHS domains (146).

Inhibition depends on CKII-mediated phosphorylation of a serine three residues upstream of the critical aspartate (146). These observations suggest that GGA1 and GGA3 may become activated by a dephosphorylation event that displaces its own DXXLL ligand and frees up the VHS domain for interaction with signals in the cytosolic domains of receptors. These interactions may in turn be enhanced by CKII-mediated phosphorylation of serine residues upstream of the receptors' signals. A recent study has presented evidence for cooperation between the GGAs and AP-1 in the sorting of MPRs (121). The GGAs were found to bind, via their hinge domains, to the ear domain of the γ subunit of AP-1. An AP-1-associated CKII then phosphorylates the GGAs, resulting in autoinhibition and possible transfer of the MPRs from the GGAs to AP-1 (121). This observation explains the presence of MPRs in TGN-derived, clathrin-coated vesicles that contain both GGAs and AP-1 (121).

OTHER SIGNALS

Acidic Clusters

Another family of sorting motifs consists of clusters of acidic residues containing sites for phosphorylation by CKII. This type of motif is often found in transmembrane proteins that are localized to the TGN at steady state, including the prohormone-processing enzymes furin, PC6B, PC7, CPD, and PAM, and the glycoprotein E of herpes virus 3 (Table 7). Several of these proteins cycle between the TGN and endosomes, and it is currently thought that the acidic clusters play a role in retrieval from endosomes to the TGN. This retrieval depends on phosphorylation of the CKII sites (147). Recent studies have identified a monomeric protein named PACS-1 (phosphofurin acidic cluster sorting protein 1), which binds to acidic clusters in a CKII phosphorylation-dependent manner (148). Interestingly, PACS-1 appears to function as a connector that links the phosphorylated acidic clusters to the clathrin-dependent sorting machinery. This connection is likely mediated by interaction of PACS-1 with AP-1 and AP-3 (149). Antisense RNA and dominant-negative interference studies support the notion that PACS-1 is required for transport of furin from endosomes to the TGN (148, 149).

Lysosomal Avoidance Signals

Discharge into the endosomal lumen of lysosomal hydrolases transported by the CD-MPR requires, in part, protonation of ^{133}Glu of the receptor (150). As the pK_a of glutamic acid is 4.1, this necessitates that the unloading occur within an endosomal compartment with a pH $<$6.0 (151). Indeed, MPRs can be visualized within the intralumenal vesicles of prelysosomal structures (152). Yet the half-life of the CD-MPR is in excess of 40 h (153); thus each receptor performs multiple delivery cycles. For the CD-MPR, endosomal retrieval

TABLE 7 Acidic cluster signals

Protein	Species	Sequence
Furin	Mouse	Tm-31-QEECPSDSEEDEG-14
PC6B (1)[a]	Mouse	Tm-39-RDRDYDEDDEDDI-36
PC6B (2)	Mouse	Tm-69-LDETEDDELEYDDES-4
PC7	Human	Tm-38-KDPDEVETES-47
CPD	Human	Tm-36-HEFQDETDTEEET-6
PAM	Human	Tm-59-QEKEDDGSESEEEY-12
VMAT2	Human	Tm-35-GEDEESESD
VMAT1	Human	Tm-35-GEDSDEEPDHEE
VAMP4	Human	25-LEDDSDEEEDF-81-Tm
Glycoprotein B	HCMV	Tm-125-KDSDEEENV
Glycoprotein E	Herpes virus 3	Tm-28-FEDSESTDTEEEF-21
Nef	HIV-1 (AAL65476)	55-LEAQEEEEV-139
Kex1p (1)	*S. cerevisiae*	Tm-29-ADDLESGLGAEDDLEQDEQLEG-40
Kex1p (2)	*S. cerevisiae*	Tm-79-TEIDESFEMTDF
Kex2p	*S. cerevisiae*	Tm-36-TEPEEVEDFDFDLSDEDH-61
Vps10p	*S. cerevisiae*	Tm-112-FEIEEDDVPTLEEEH-37

See legends to Tables 1–3 for explanation of signal format. Serine and threonine residues are underlined.

[a]The number in parentheses is the motif number.

depends upon both cytosolic palmitoylation (123) and an ^{18}FW di-aromatic sorting signal; an FW→AA mutation causes a ~tenfold increase in CD-MPR recovered within dense lysosomes (154). A functionally analogous ^{18}YF di-aromatic signal is present in the cytosolic domain of the mannose receptor (155). The CD-MPR sequence is recognized with micromolar affinity by the carboxy-terminal segment of TIP47 (tail-interacting protein of 47 kDa) (156), a protein that also binds the cytosolic domain of CI-MPR despite the absence of an FW sequence (4, 157). The binding site on the CI-MPR is, in part, localized to the sequence ^{49}PPAPRP of the cytosolic domain, adjacent to the ^{26}YSKV-internalization signal, and appears to associate with TIP47 via a hydrophobic-based interaction (4). Still, the MPR interactions are selective as TIP47 does not bind appreciably in vitro to the cytosolic tails of furin, TGN38, or the LDL receptor, or to other trafficking signals within the CI-MPR (4, 156, 157). The intracellular distribution of TIP47 is compatible with a role in retrieval of MPRs from late endosomes, and decreasing TIP47 abundance with antisense oligonucleotides reduces substantially the half-life of the CI-MPR (156).

In the presence of the nonhydrolyzable GTP analogue, GTPγS, a fourfold increase in TIP47 association with endosome membranes occurs (156). The GTP requirement appears to reflect the involvement of Rab9, as TIP47 cooperatively

binds Rab9•GTP and MPRs via separate sites (158). The improved affinity of the ternary complex for MPRs (K_d = 0.3 μM) is proposed to direct TIP47 to the appropriate intracellular site (158). Curiously, TIP47 is a member of the perilipin family of proteins associated with the cytosolic surface of intracellular lipid droplets (159), but the significance of this homology to the sorting of MPRs is not currently known.

NPFX(1,2)D-Type Signals

Another internalization sequence in *S. cerevisiae*, NPFSD, was uncovered within the cytosolic domain of the furin-like endoprotease, Kex2p, when this region was switched for the carboxy-terminal region of Ste2p (160). Kex2p does not usually traffic to the cell surface, but the NPFSD sequence, when appended to a noninternalizing reporter, independently directs efficient uptake (160–162). A related sequence, NPFSTD, near the carboxy-terminus of Ste3p is able to drive **a**-factor-dependent internalization of a truncated, nonubiquitinatable form of this receptor (160). Ste2p contains a lower-potency variant, GPFAD, which seems to function redundantly with ubiquitination (see below) in directing optimal receptor internalization (162). Although this type of internalization motif has an embedded NPF triplet, and the proline and phenylalanine are both essential side chains, the sequence does not engage EH domains (162). Instead, activity of the NPFX(1,2)D sequence is dependent upon Sla1p—specifically, the first *SLA1* homology domain (SHD1), to which it binds directly (162). Part of a known endocytic complex together with Pan1p and End3p (163), Sla1p couples cargo recognition and internalization with actin cytoskeletal dynamics, which plays a major role in endocytosis in yeast (164).

UBIQUITIN AS A SORTING SIGNAL

Discovery of Ubiquitin-Based Sorting

Ubiquitin is a 76-amino acid polypeptide virtually invariant in sequence throughout eukaryotes. The carboxy-terminal residue, ^{76}Gly, can be conjugated by an isopeptide linkage to lysine side chains on target proteins. Conjugation is the result of a sequential series of reactions involving the generation and transfer of a thioester-bound ubiquitin intermediate by the E1, E2, and E3 conjugation machinery proteins (165). Several endogenous lysines (e.g., ^{29}Lys, ^{48}Lys, ^{63}Lys) within the protein-linked ubiquitin molecule can, subsequently, be self-conjugated to additional ubiquitin molecules, generating polyubiquitin chains. Post-translational ubiquitination is at the core of regulated intracellular protein turnover, and ^{48}Lys-linked polyubiquitin chains of tetraubiquitin or greater serve as a signal for protein degradation via the 26S proteasome (166). Appreciation that ubiquitin also functions as an authentic trafficking signal came from the discovery that, in *S. cerevisiae*, internalization of the G protein-coupled α-factor

(Ste2p) and **a**-factor (Ste3p) pheromone receptors is accompanied by ubiquitination of these proteins (167, 168). Genetic background mutations that prevent endocytosis cause ubiquitinated species to accumulate, whereas compromising the cellular ubiquitination machinery retards surface uptake and prolongs receptor half-life (167, 168). It was quickly established that a multitude of yeast plasma membrane receptors and permeases/transporters (e.g., Ste6p, Gal2p, Gap1p, Fur4p, Pdr5p, Zrt1p,) use ubiquitination as an endocytic signal (169–171).

Apparently the major endocytic signal in budding yeast, in higher organisms ubiquitin addition also regulates the internalization of certain transmembrane proteins. One of the first indications came from studies on the growth hormone (GH) receptor in the CHO-ts20 cell line that exhibits a temperature-sensitive E1 ubiquitin-activating enzyme. At the nonpermissive temperature, no uptake or lysosomal degradation of transfected GH receptor occurs (172). Mutation of ^{327}Phe in the cytosolic domain of the receptor blocks endocytosis and also inhibits receptor ubiquitination (173). Specific inhibitors of the proteasome, which cause depletion of free ubiquitin within the cell, abrogate ligand-stimulated destruction of the GH receptor, but transferrin receptor endocytosis is unaffected (174, 175). Also, blocking endocytosis by cyclodextran-mediated cholesterol depletion, or with a dominant-negative dynamin mutant, causes ubiquitinated GH receptor to accumulate on the plasma membrane (176). These experiments firmly establish that internalization of the GH receptor requires the ubiquitin-conjugation machinery and that the receptor is ubiquitinated directly.

There is now good evidence for ubiquitination regulating the internalization of the EGF (177–179), MET (180), and CSF-1 (181, 182) receptors, the epithelial sodium channel (ENaC) (171), the aggregated IgG-bound Fc receptor FcRγIIA (183), and the transmembrane Notch ligand Delta (184–186). The α1 subunit of the glycine receptor is modified with one to three ubiquitin molecules at the plasma membrane as a prelude to uptake (187), and ubiquitination of E-cadherin precedes internalization (188). There is also evidence that endocytosis of the pre-T cell receptor in thymocytes is dependent upon ubiquitination (189). In *C. elegans*, targeted neural overexpression of ubiquitin reduces the surface density of glutamate receptors (190) in a manner that is dependent on clathrin-mediated endocytosis, as mutant *unc-11* (AP180) counteracts the effect of the excess ubiquitin. Direct ubiquitination of the glutamate receptor can be demonstrated biochemically (190). Final corroboration of the role of ubiquitination in endocytic mobilization comes from the demonstration that the Karposi's sarcoma-associated herpes virus protein K3 can route surface MHC class I for lysosomal degradation by possessing inherent E3 ubiquitin ligase activity that ubiquitinates a single MHC class I lysine (191, 192).

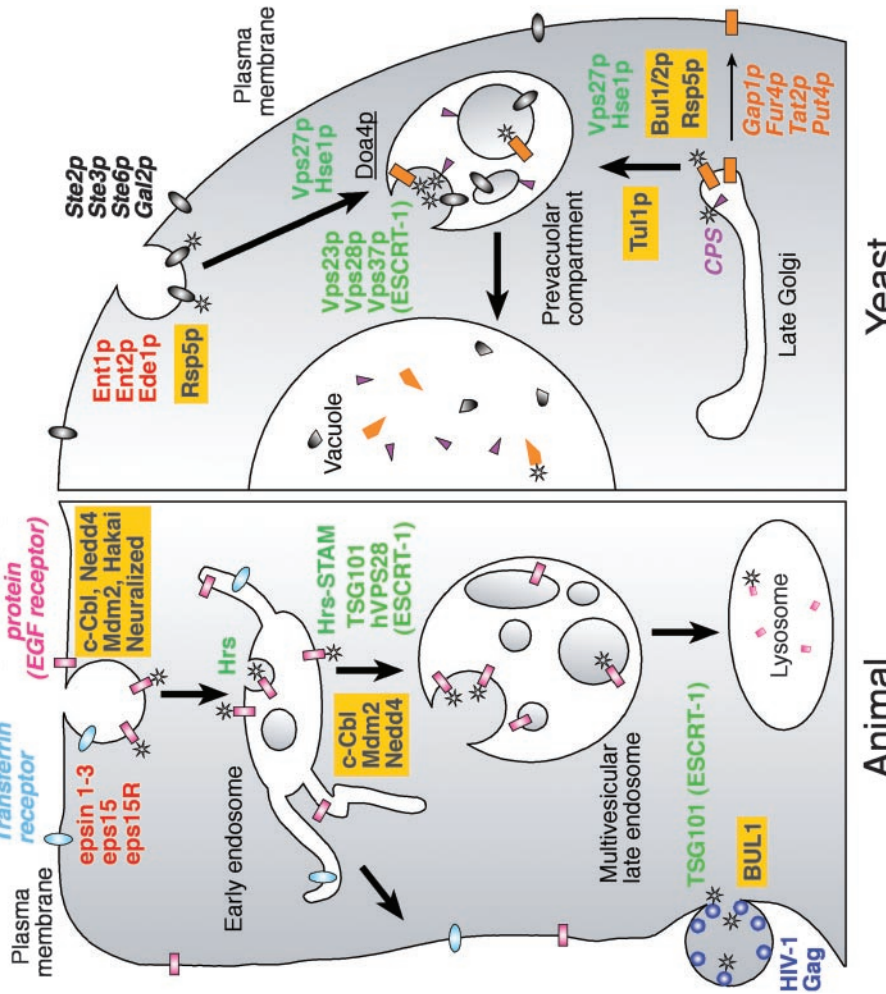
Multiple Sorting Steps Involving Ubiquitin

Ubiquitination has now been found to effect protein sorting at other intracellular stations as well. Ubiquitin conjugation is a cyclical process; because ubiquitin is

a long-lived protein, the bulk of activated ubiquitin generated by the E1 enzyme comes from ubiquitin recovered by deubiquitinating enzymes (193). Inactivation of one deubiquitinating enzyme in yeast, Doa4p, results in pleiomorphic effects, including defects in surface uptake (194–197). Bypass suppressors of a *doa4* mutation turn out to be six class E *vps* (vacuolar protein sorting) mutants (194). Class E *vps* mutants are defective in multivesicular body formation, and as these suppressors essentially negate the role of Doa4p, the data indicate that the late endosome/prevacuolar compartment is a major sink for ubiquitin (198). However, even though overexpression of ubiquitin overcomes the *doa4* phenotype (195, 199), this does not indicate that Doa4p simply retrieves and recycles ubiquitin from tagged proteins only originating from the cell surface (198). Biosynthetic delivery of newly synthesized vacuolar proteins carboxypeptidase S (CPS) and the polyphosphatase Phm5p into the vacuolar lumen is also ubiquitin dependent (196, 200). These proteins are ubiquitinated directly; a single major lysine acceptor has been mapped in both Phm5p (196) and CPS (200), but other potential acceptors are also utilized in Phm5p (196). Biosynthetic ubiquitination of CPS occurs post Golgi but prior to delivery to late endosomes (200). If ubiquitination of CPS is blocked, the unprocessed form accumulates on the limiting membrane of the vacuole instead of inside (200). The same is true for Phm5p (196). In-frame addition of either a single unextendable ubiquitin (196, 201) or a ubiquitination-directing sequence (200) routes into the vacuole lumen a variety of proteins that normally traffic through the late endosome/prevacuolar compartment without entering intralumenal structures. This does not involve passage through the plasma membrane, showing that ubiquitin-directed sorting of proteins not originating from cell surface certainly occurs (Figure 6).

In mammalian cells, signaling receptors, like the EGF receptor, move into intralumenal membranes of multivesicular endosomes following ligand stimulation and internalization (202). TGFα, an alternate ligand for the EGF receptor, does not induce this translocation, nor substantial receptor degradation, because the ligand dissociates from the receptor within early endosomes (203, 204). TGFα also does not cause protracted ubiquitination of the EGF receptor (204). Proteasome inhibitors prevent both the sequestration of EGF-activated receptors into the interior of multivesicular bodies (204) and degradation of internalized receptors (204, 205). Also, lysosomal turnover of the HIV coreceptor CXCR4 requires a degradation motif (206). Mutation of the three lysine residues within this motif abolishes degradation of the G protein–coupled receptor while monoubiquitination of the receptor correlates with lysosomal delivery (206). Interestingly, proteasome inhibitors selectively disrupt delivery of membrane proteins into the lysosome; soluble protein traffic is normal (175, 204). Together, these studies reveal that transmembrane protein relocation into the interior elements of the multivesicular late endosome is a regulated step in lysosomal delivery that is under the control of ubiquitin.

Sorting decisions are, of course, made before reaching the late endosomal compartment. Minutes after endocytosis, EGF and GH receptors are actively


Plasma membrane
Transferrin receptor
epsin 1-3
eps15
eps15R
Transmembrane protein (EGF receptor)
c-Cbl, Nedd4
Mdm2, Hakai
Neuralized
Hrs
Early endosome
Hrs-STAM
TSG101
hVPS28
(ESCRT-1)
c-Cbl
Mdm2
Nedd4
Multivesicular late endosome
Lysosome
TSG101 (ESCRT-1)
BUL1
HIV-1 Gag
Animal
Plasma membrane
Ste2p
Ste3p
Ste6p
Gal2p
Ent1p
Ent2p
Ede1p
Rsp5p
Vps27p
Hse1p
Doa4p
Vps23p
Vps28p
Vps37p
(ESCRT-1)
Vacuole
Prevacuolar compartment
Tul1p
Vps27p
Hse1p
Bul1/2p
Rsp5p
CPS
Late Golgi
Gap1p
Fur4p
Tat2p
Put4p
Yeast

segregated away from components that recycle back to the cell surface (11) (Figure 6). Transferrin receptors with an in-frame ubiquitin at the amino terminus are not recycled to the surface as efficiently as the wild-type receptor (12). Cells treated with the proteasome inhibitor lactacystin (177, 204) or expressing a ubiquitination-defective receptor (179) display elevated EGF receptor recycling at the expense of receptor degradation. As proteasome inhibitors deplete cellular ubiquitin, this again suggests that ubiquitin participates in partitioning molecules away from recycling cargo. Lysosomal destruction of IL-2 (207, 208) and glycine (187) receptors is also prevented by proteasome inhibitors. Because proteins return to the plasma membrane instead of being delivered to the limiting membrane of the lysosome, this work reveals that ubiquitin participates in additional sorting decisions at the early endosome as well (Figure 6). In fact, the discovery that ubiquitin is also involved in sorting within the endosomal compartment makes definitive interpretation of the many studies making use of proteasome inhibitors difficult as all steps are likely to be perturbed under these conditions.

Finally, the trafficking of certain amino acid permeases in yeast is governed by nutritional status. Under appropriate environmental conditions, newly synthesized general amino acid permease, Gap1p, proline permease, Put4p, uracil permease, Fur4p, and tryptophan permease, Tat2p, are diverted to the vacuole for destruction (198, 209–212). This change in permease routing is ubiquitin dependent, a decision that appears to occur around Golgi exit (211), before delivery to the prevacuolar compartment (212) (Figure 6). Thus, a ubiquitin sorting tag can clearly function at multiple branchpoints en route to the lysosome/vacuole.

Characteristics of the Ubiquitin Signal

In yeast, addition of only one or two ubiquitin moieties to Ste3p is sufficient for constitutive endocytosis (213), whereas uptake of the galactose transporter, Gal2p, can occur within cells expressing an unextendable ubiquitin, devoid of any lysines (199). This demonstrates that the extent of ubiquitin derivation required to generate a functional ubiquitination signal is low. In fact, simply

←

Figure 6 Ubiquitin-directed endosomal trafficking. Schematic depiction of the likely site of action of ubiquitin-conjugating machinery (boxed in yellow), endocytic recognition/sorting components (bold type), and deubiquitinating enzymes (underlined) displayed in relation to the relevant subcellular organelles in either in animal (*left*) or yeast (*right*) cells. With the exception of the transferrin receptor, the trafficking itinerary of the various cargo proteins (italic type) shown is regulated by enzymatic ubiquitin addition (✲). In yeast, the activity of Rsp5p and Bul1/2p at the Golgi complex diverts the Gap1p permease from delivery to the cell surface by driving polyubiquitination and trafficking to multivesicular bodies. See text for complete details.

appending a single ubiquitin molecule to the carboxyl end of Ste2p is sufficient to drive endocytosis of this receptor (197). Similarly, fusion of ubiquitin to Pma1p, the resident plasma membrane H^+-ATPase, promotes internalization (214). In mammalian cells, addition of a single ubiquitin molecule essentially in place of the amino terminus of invariant chain (Ii) also changes the steady-state distribution from the plasma membrane to internal endocytic elements (215). In these systems, uptake is apparently not dependent upon extension of the fused ubiquitin as no difference in trafficking behavior of K48R mutants is seen (197, 215). A conjugation-incompetent, fully arginine-substituted form of ubiquitin expressed in a *doa4* background, which has defective Ste2p uptake, rescues α-factor internalization (197). Together, these studies indicate that a single molecule is the minimal ubiquitin tag required for endocytosis, and it is suggested that this distinguishes the endocytic ubiquitin signal from the proteasome degradation signal (216).

However, in vivo, the extent of ubiquitination and the type of linkage vary in both yeast and mammals. Even Ste2p is multiubiquitinated in vivo (197). Both Gap1p and Fur1p permeases are polyubiquitinated on two lysine residues within the cytosolic amino terminus, ^{9}Lys and ^{16}Lys in Gap1p (212), and ^{38}Lys and ^{41}Lys in Fur4p (217). In these proteins, simultaneous mutation of both acceptor lysines traps the permease at the surface (212, 217). A single utilizable lysine is sufficient to drive endocytosis, but the rate is slowed (212), as it is when only monoubiquitination occurs (218), suggesting that multiubiquitin is a more effective signal. Further, using Lys→Arg substituted ubiquitins, it has been established that chain extension in both permeases is via a ubiquitin ^{63}Lys linkage (212, 218). Similarly, a single ubiquitin fused to the transferrin receptor retards recycling of only ~25% of internalized transferrin receptors; most still recycle efficiently (12), suggesting that multiubiquitin might be the optimal signal form.

In many instances, it has not been established whether the multiple ubiquitin-positive species detected biochemically represent limited assembly of multiubiquitin chains or monoubiquitination of several distinct lysine residues. In vertebrates, a range from limited to extensive ubiquitination can be seen. Deubiquitinating enzymes, the largest family of ubiquitin-modifying enzymes, rapidly disassemble/salvage conjugated ubiquitin (193), possibly causing underestimation of the extent of ubiquitination in vivo. The efficiency of anti-ubiquitin antibodies is also poor, and consequently, in many studies ubiquitination is detected using overexpressed, epitope-tagged ubiquitin. Thus, the precise molecular details of the ubiquitin signal remain to be determined. Because fusion of only residues 36–44 of ubiquitin to a cytosolically truncated form of the α chain of the IL-2 receptor facilitates its internalization, it was suggested that ^{43}Leu-^{44}Ile are part of a [DE]XXXL[LI] signal within ubiquitin (215). However, in yeast, a fusion of this dileucine-bearing segment of ubiquitin to a truncated Ste2p does not promote internalization whereas whole ubiquitin does (214). Although this might be due to failure of this putative [DE]XXXL[LI] signal to interface with

yeast AP-2, accessibility of the dileucine within folded ubiquitin might be limited.

Substitution of aliphatic hydrophobic side chains buried within the internal core of ubiquitin disrupts folding and abrogates the sorting capacity of a Ste2p-ubiquitin chimera (214). This observation suggests that in *S. cerevisiae*, the ubiquitin tag is detected as a folded interaction surface rather than as a linear sequence in an extended conformation. Systematic mutation of a ubiquitin appended to Ste2p as the only internalization information uncovered two roughly contiguous surface patches containing important side chains. Most important are ^{4}Phe and ^{44}Ile, with ^{8}Leu and ^{70}Val, in proximity to ^{44}Ile, also participating in the generation of the internalization signal (214). Interestingly, ^{8}Leu, ^{44}Ile, and ^{70}Val are also important elements of the polyubiquitin determinant necessary for proteasomal degradation (219), which suggests that cellular identification of these ubiquitin signals might employ related components (see below).

Finally, although ubiquitination of the GH receptor occurs at the plasma membrane, and ubiquitinated GH receptor is found in clathrin lattices (176), receptor-bound ubiquitin is not necessary for internalization (220). Simultaneous Lys→Arg substitution of all 16 potential ubiquitin acceptors in the cytosolic domain of the GH receptor is still compatible with internalization (220). Intriguingly, this internalization is still dependent upon both a competent cellular ubiquitination machinery and a ubiquitin-dependent endocytosis (UbE) motif within the GH receptor cytosolic domain (220). This suggests that, in this system, either ubiquitination of another component serves as an internalization signal for the receptor, or that ubiquitination of an inhibitory molecule promotes exposure of a ubiquitin-independent sorting signal, or that the docked ubiquitination machinery tags the complex for uptake (220). An example of the last possibility is the recent discovery that CIN85 bridges a receptor-bound E3 ligase and endophilin, a known endocytic protein, to promote internalization (221, 222).

Generation of the Ubiquitin Sorting Signal

Selection of proteins for ubiquitination is mediated by E3 ubiquitin ligases, which contain substrate-recognition modules. Two general classes of E3s can be distinguished on the basis of sequence homology and mode of operation with respect to the ubiquitin-carrying E2 component (165, 223). HECT (homologous to *E*6-AP carboxy terminus) domain E3s are often large proteins with an ~350-residue HECT domain, commonly positioned at the carboxy terminus. The HECT domain transfers activated ubiquitin from a designated E2 onto an invariant HECT cysteine prior to conjugation onto the substrate; thus, HECT E3s conjugate ubiquitin directly. RING (Really Interesting New Gene) finger E3 ligases all contain an ~70-amino acid sequence with conserved cysteine and histidine residues positioned to coordinate two zinc atoms (165, 223). The folded finger provides an E2 interaction surface and this class of E3 provides target specificity while working in catalytic conjunction with an E2, as there is no thiol ester-linked ubiquitin directly associated with these E3s.

In many instances, ubiquitination of proteins at the plasma membrane is preceded by phosphorylation that generates a docking site for a particular E3 ubiquitin ligase. In *S. cerevisiae*, the major and best-characterized E3 ligase operative in endocytic trafficking is Rsp5p (171) (Figure 6). In various *rsp5* mutants, plasma membrane proteins fail to be ubiquitinated and, consequently, internalized (171). Rsp5p is a HECT-type E3 that recognizes its targets via three amino-terminal, tandem WW domains. The ~40-amino acid WW fold, bearing two conserved tryptophan residues, is a protein–protein interaction domain that recognizes several proline-rich ligand motifs, including the PPXY and PPLP motifs, polyproline stretches with included glycine and methionine residues, as well as phosphoserine- and phosphothreonine-containing sequences (224). The three Rsp5p WW domains appear to function redundantly, although differential contributions to fluid-phase uptake and protein endocytosis can be mapped (225, 226). For Ste2p ubiquitination, Rsp5p recognizes a hyperphosphorylated ^{331}SINNDAKSS sequence and ubiquitinates the included ^{337}Lys as well as other lysines within the carboxy-terminal segment (197). A related sequence, DAKTI, which controls Ste6p endocytosis (227), is likely also an Rsp5p target. Unexpectedly, internalization of a Ste2 p-ubiquitin chimera is defective in an *rsp5* background (161), indicating that a monoubiquitin tag in the absence of functional endocytic conjugation machinery is insufficient to promote pheromone receptor internalization. These results are in general accord with the GH receptor data (220) and suggest that additional ubiquitination of core endocytic machinery might be a prerequisite for efficient endocytosis (see below).

G protein-coupled receptor internalization in vertebrates is also triggered by phosphorylation, but in contrast to the pheromone receptors, the intermediate adaptor β-arrestin provides critical internalization information and interfaces with the clathrin-coat machinery (228). Still, β-arrestin and the β2-adrenergic receptor itself are ubiquitinated in a stimulus-dependent fashion (229). The E3 ligase in this instance is Mdm2, a RING E3 better known for promoting the proteasomal degradation of the cell-cycle regulator p53 (230). Mdm2-null cells show that Mdm2-catalyzed ubiquitination of arrestin is required for β2-adrenergic receptor internalization, whereas Mdm2-driven polyubiquitination of the receptor is required for lysosomal degradation because a lysine-free receptor form is turned over very slowly (229).

Nedd4 is an Rsp5p orthologue in vertebrates that is linked to internalization of the $\alpha_2\beta\gamma$-ENaC heterotetramer, as overexpression of catalytically defective Nedd4 results in increased channel number and activity at the cell surface (171, 231). Human Nedd4 has four WW domains that bind, primarily via WW domain 2 and/or 3 (232, 233), to PPPXY (PY) motifs found in the cytosolic segment of each ENaC subunit (171). The physiologically relevant PY motifs are contained within the β and γ subunits (234, 235), and inherited mutations of the PY motifs in either the β- or γ-ENaC subunits cause Liddle's syndrome (236, 237), a form of severe hypertension. Disease appears to be due to elevated channel activity because ubiquitin-dependent endocytosis is disrupted and direct, lysine-depen-

dent ubiquitination of α- and γ-ENaC subunits by Nedd4 has been demonstrated (231). Additional regulation of the E3-substrate interaction also occurs; phosphorylation of a threonine residue flanking the PY motif within the carboxy-terminal segment of either β- or γ-ENaC subunits increases the affinity for Nedd4 WW domains by threefold (238). *Drosophila* Nedd4 is involved in internalization of Commissureless (Comm), a regulatory transmembrane protein operative in axonal migration (239). DNedd4 binds, via PY—WW-domain interactions, to Comm to induce ubiquitination and facilitate Comm-dependent downregulation of the Roundabout (Robo) receptor from the cell surface (239). A Nedd4-like E3 also appears likely to regulate internalization of the ClC-5 chloride channel by interacting with a PPYTPP internalization sequence (240), and Itch is yet another Nedd4-like HECT-domain E3 that recognizes Notch via WW-dependent recognition to prompt ubiquitin addition (241).

c-Cbl, a RING finger E3 ligase, is recruited onto the EGF receptor via an SH2-like domain adjacent to the E2-binding RING finger domain (177, 178) (Figure 6). Phosphorylation of ^{1045}Tyr in the cytosolic domain of the receptor generates a major docking site for c-Cbl. A Y1045F substitution diminishes internalization of the receptor while potentiating recycling and mitogenic signaling (179). Dominant-negative Cbl diminishes EGF receptor ubiquitination and reduces lysosomal degradation (204), whereas H_2O_2 inhibition of EGF receptor uptake correlates with inhibition of receptor ubiquitination (242). This system also illustrates nicely how ongoing ubiquitination can guide a receptor along its trajectory toward the interior of the lysosome. Ubiquitination begins at the cell surface; the Y1045A mutant (179), or wild-type EGF receptor in dominant-negative dynamin-expressing cells (243), arrests at the cell surface. Yet maximal ubiquitination of the receptor occurs as it penetrates the endosomal compartment (204). Internalized TGFα-activated EGF receptors recycle back to the cell surface because the ligand dissociates from the receptor within early endosomes and ubiquitination is terminated (203, 204). As kinase-deficient (202) or Y1045F-substituted (179) EGF receptors do not translocate into the multivesicular body interior but instead recycle to the cell surface, sustained activation of the receptor, to maintain the ubiquitinated state, is required for directed passage to the lysosome. In fact, c-Cbl and the EGF receptor traffic toward the lysosome bound together (204, 244, 245); after 60 min both the EGF receptor and c-Cbl are found within the intralumenal vesicles of multivesicular bodies (245) (Figure 6). Thus, progressive ubiquitination generates sorting signals to ensure downregulation of the receptor.

c-Cbl is also involved in negative regulation of other receptor tyrosine kinases including the platelet-derived growth factor (PDGF) receptor (246). CSF-1 receptor internalization is slow in c-Cbl mutant cells (181), and c-Cbl also translocates onto tyrosine-phosphorylated Notch1, mediating lysosomal degradation of the uncleaved transmembrane form of the receptor (247). The immunological phenotypes of c-Cbl$^{-/-}$ mice suggest that this E3 is also involved in negative regulation of antigen receptors (189). Neuralized, a RING E3 ligase in

flies ubiquitinates the Notch ligand Delta in vitro and using the RING domain, facilitates endocytosis and lysosome degradation of Delta (184–186). Another RING E3, Hakai, drives internalization of E-cadherin (188). Tyrosine phosphorylation facilitates Hakai recruitment, and overexpression of Hakai in epithelial cells perturbs cell-cell adhesion by inducing E-cadherin endocytosis (188).

An interesting variation on the generation of a ubiquitin sorting signal is seen in the vacuole-directed diversion of the Gap1p permease. Here, Rsp5p is the responsible E3 (211), in accord with localization of the protein to both cell surface and perinuclear structures (248). Bul1p/Bul2p are additionally required at the Golgi complex to deflect Gap1p from surface delivery, as a *bul1*Δ*bul2*Δ strain delivers excess permease to the surface (211) (Figure 6). Bul1p is ~50% identical to Bul2p, and they act redundantly by binding, via a PY motif, to the WW-domain region of Rsp5p and driving extensive polyubiquitination of Gap1p (211) as a signal for intracellular vacuolar delivery (211). A subsequent study (212) revealed that Rsp5p, Bul1/2p, Doa4p, and ^{9}Lys and ^{16}Lys of Gap1p are also all required for internalization. As ^{63}Lys-ubiquitin linkages are generated (212), perhaps Bul1/2p are required to modulate Rsp5p activity specifically to generate the poly-^{63}Lys-linked chains.

Tul1p is an atypical ligase that ubiquitinates biosynthetic cargo for intracellular delivery to the vacuole. Unusual in that the RING domain is located at the carboxyl end of a membrane protein with seven predicted transmembrane domains, Tul1p localizes to the Golgi complex at steady state (249) (Figure 6). Tul1p appears to be the E3 that directs both CPS and Phm5p into the interior of the vacuole as, in a *tul1*Δ background, GFP-tagged forms of these two vacuolar proteins accumulate on the limiting vacuolar membrane instead of internally (249). Missorting correlates with a fivefold decrease in CPS ubiquitination, and Tul1p-ubiquitinated proteins proceed to the vacuole without passing through the cell surface, showing that sorting occurs at the Golgi (249). The mode of substrate recognition is also unusual; transmembrane domains harboring polar side chains mark proteins for Tul1p modification (249), explaining how a single L→D substitution within the transmembrane domain of Pep12p causes ubiquitin-dependent translocation of this endosomal SNARE to the vacuole interior (250). This mode of substrate recognition has led to the suggestion that, in addition to routing normal vacuolar components, Tul1p acts as a quality control device to send abnormal proteins that reach the Golgi complex for destruction (249). The dichotomy between Tul1p and Rsp5p-Bul1/2p in directing different transmembrane cargo along a similar path to the vacuole probably reflects the nutritionally inducible switching required for Gap1p, Fur4p, and other permeases.

Recognition of Ubiquitinated Transmembrane Proteins

Since added ubiquitin serves as a sorting determinant at the cell surface, at endosomes, and at the TGN (Figure 6), ubiquitinated transmembrane cargo must interface with several different sorting components. At present, it appears that

decoding the ubiquitin sorting tag might be performed, primarily, by a limited group of ubiquitin-binding modules embedded within compartment-specific components. The 26S proteasome degrades polyubiquitinated targets, and S5a (Rpn10p in *S. cerevisiae*) is a ubiquitin-binding subunit of the 19S regulatory component of the proteasome. Using a bivalent ubiquitin-binding sequence from S5a (251) as a search model, the endocytic proteins Ent1p/Ent2p/epsin 1–3, eps15/eps15R, Vps27/Hrs, Hse1p/STAM1, and STAM2 (also termed Hrs binding protein, Hbp) were each revealed to contain one to three copies of a degenerate consensus sequence termed the ubiquitin-interacting motif (UIM) (252). The UIMs in the epsins, eps15/eps15R, Hrs/Vps27p, and Hse1p all bind directly to ubiquitin, to monoubiquitin relatively weakly, and to polyubiquitin chains with better apparent affinity (12, 205, 253–256). Point mutation of conserved side chains in the Vps27p and Hrs UIMs, or UIM deletion, blocks ubiquitin association (12, 205, 253, 256), and interaction of yeast epsin, Ent1p, with ubiquitin is sensitive to a ubiquitin mutation (I44A) that disrupts Ste2p internalization (256). UIM binding to ubiquitin chains with alternate linkages has not yet been addressed.

The overall domain architecture of these UIM-bearing endocytic proteins (Figure 2) suggests that they could act as cargo-dedicated intermediate adaptors for sorting ubiquitinated cargo at discrete intracellular sites. The epsins superficially resemble Dab2; a PtdIns(4,5)P_2-binding ENTH domain is followed by a largely unstructured carboxy-terminal segment (257, 258) containing interaction motifs for engaging AP-2, clathrin, and eps15, respectively (Table 2). The UIMs are positioned directly following the ENTH domain and thus proximal to the membrane. eps15, on the other hand, has a central coiled-coil domain that can hetero-oligomerize with eps15R (or intersectin), flanked by amino-terminal tandem EH domains and multiple AP-2 binding triplets at the carboxy-terminal end. Two UIMs are located at the extreme carboxyl terminus of eps15/eps15R. As epsin and eps15 are binding partners, and epsin contributes to clathrin lattice assembly, they could cooperatively capture ubiquitinated cargo within clathrin buds at the cell surface. This idea is supported by UIM deletions in the yeast epsin, Ent1p (Figure 6). Although an *ent1*Δ*ent2*Δ strain expressing a UIM-deleted Ent1p from a plasmid exhibits minimally impeded Ste2p endocytosis (256), this is apparently due to functional redundancy with Ede1p, the orthologue of mammalian eps15, in which the UIM is replaced with a ubiquitin-associated (UBA) domain. The UBA domain is a ~40-amino acid region found in some E2, E3, and several other proteins and binds multiubiquitin chains (259). In *ent1*Δ*ent2*Δ*ede1*Δ yeast, production of the UIM-deleted Ent1p does not support α-factor internalization, whereas wild-type Ent1p does. As Ste2p is normally ubiquitinated in the UIM-deleted strain, the data reveal the importance of the UIM and functionally analogous UBA domains in endocytic sorting.

Analogously to Ent1p and Ede1p, Vps27p hetero-oligomerizes with Hse1p (253), forming a complex similar to the Hrs-STAM complex in mammals. Both CPS and Ste2p delivery to the vacuolar lumen is blocked in either *vps27*Δ or

vps27 UIM-mutated strains (256), which confirms that this sorting complex gathers both biosynthetic and endocytic cargo for transport to the vacuole (Figure 6). That a complex of two UIM-harboring proteins works redundantly is shown by the correct delivery of ubiquitinated Ste3p to the vacuole interior in *vps27-ΔUIM* or in *hse1-ΔUIM* strains but not in a *vps27-ΔUIMhse1-ΔUIM* double mutant, where the pheromone receptor ends up missorted on the vacuolar outer membrane (253). Indeed, vacuoles isolated from the double mutant do not contain ubiquitinated polypeptides as do normal vacuoles (253). A vacuolar protein that does not require ubiquitin for delivery, Sna3p (196), is still appropriately delivered to the lumen, which also contains internal membranes, in a *vps27/hse1* UIM-deleted strain (253). This again highlights the sorting function of the UIM.

In animal cells, localization of an Hrs-STAM complex to clathrin-scaffold regions upon early endosomes places the UIMs in an appropriate position to retain ubiquitinated cargo at this site (Figure 6). The importance of the Hrs lipid-binding FYVE domain (Figure 2) is revealed by the disappearance of the endosomal clathrin coats in the presence of the PtdIns3-kinase inhibitor wortmannin (11, 12), a drug that prevents sequestration of EGF receptors within multivesicular bodies (260). Overexpression of Hrs causes redistribution of clathrin onto early endosomes (261). Coexpression of Hrs and a ubiquitin-tagged transferrin receptor markedly increases retention of this receptor within endosomal structures, a phenomenon overcome by inactivating the Hrs UIM by mutagenesis (12). Overexpressed Hrs alone causes relocation of endogenous ubiquitinated proteins to endosomes and slows degradation of EGF (205, 261, 262). Thus, increasing the density of the UIM-containing proteins on endosomal structures concentrates ubiquitinated cargo at the site. These recent studies all show a strong correlation between normal function of UIM-containing traffic proteins and appropriate sorting of ubiquitin.

A severely truncated form of Hrs expressed in mutant *Drosophila* causes strikingly enlarged endosomes and persistent activation of several receptor tyrosine kinases that are not appropriately downregulated (254). Targeted disruption of Hrs in mice is embryonically lethal, but primary embryonic cultures again show defects in early, but not late, endosome morphology (263). The cultured $Hrs^{-/-}$ cells contain numerous large transferrin receptor-positive vacuoles, mirroring wortmannin-treated cells (263). Although this is all in accord with the clathrin/Hrs sorting scaffold preferentially operating at the early endosomes, where Hrs has consistently been localized, Hrs-mutant flies have a pronounced defect in multivesicular body formation (254), and Vps27p, the homologue in *S. cerevisiae*, is a class E vps protein involved in multivesicular body formation (256). This suggests either that, in the absence of sorted ubiquitinated cargo, no multivesicular body formation occurs, or that Hrs is also involved in the involution process directly. Support for the latter idea comes unexpectedly from

recent work on HIV budding from the surface of infected cells where, with respect to the cytosol, virus budding is topologically analogous to endosome involution. The HIV-1 Gag viral polyprotein is alone sufficient to drive virion budding. The p6 late domain of Gag carries a PTAP peptide sequence that is required for budding, and this sequence binds directly to a cellular component, tumor-susceptibility gene 101 (TSG101) (264). Remarkably, TSG101 is the orthologue of Vps23p, which, in a 350-kDa functional complex with Vps28p and Vps37p termed ESCRT-I (endosomal sorting complex required for transport), is involved in sorting ubiquitinated cargo into multivesicular endosomes (200, 265). RNAi-mediated knockdown of TSG101 protein level arrests virus separation and the release of infectious particles (264). The same cellular machinery that retroviruses usurp to promote efficient viral particle release normally facilitates lysosomal delivery. After 60 min in A-431 cells, which express very high levels of the EGF receptor, internalized EGF colocalizes with early endosome autoantigen 1 (EEA1), ubiquitinated proteins, TSG101, and hVPS28 at large scattered intracellular structures (205). *tsg101* mutant cells recycle to the surface, instead of degrading, ligand-bound EGF receptors and also missort the CI-MPR and lysosomal hydrolases (265). Antibodies against hVPS28 or depletion of TSG101 with RNAi arrests EGF receptor traffic in a similar fashion to overexpression of Hrs (205). TSG101 RNAi also prevents viral E3-directed degradation of MHC class I (191).

Budding of several retroviruses, including HIV, is inhibited by proteasome inhibitors, and spherical viral particles, still tethered to the plasma membrane, accumulate (266, 267). The effect can be bypassed by overexpression of ubiquitin, and Gag itself is ubiquitinated (266). In fact, a number of viral Gag proteins display PY motifs, and a Nedd4-like HECT E3 with two WW domains, termed BUL1, appears to be involved in retrovirus budding (268). For virus release, the PTAP sequence does not function independently in the absence of ubiquitination (269) because the interaction of Gag with TSG101 is bipartite, utilizing both the PTAP sequence and conjugated ubiquitin (264). TSG101/Vps23p contains a ubiquitin recognition module termed the ubiquitin conjugating (UBC)-variant, or UEV domain. E2 conjugating enzymes, which receive activated ubiquitin from E1, exhibit a conserved ubiquitin-conjugation domain, termed UBC, containing an invariant catalytic cysteine for thioester linkage to ubiquitin. Both Vps23p and TSG101 bear sequence homology to the UBC domain but lack the active site cysteine. Like the UIM and UBA, the UBC domain imparts to both proteins the capacity to bind ubiquitin (200, 205, 270). The Gag PTAP sequence binds directly to the UEV domain to replace an α helix normally found in the canonical UBC fold, but which is absent in TSG101 (270), and although TSG101 binds the Gag PTAP sequence directly, the affinity increases roughly tenfold if ubiquitin is conjugated to the Gag p6 (264). These studies indicate that the viral Gag polyprotein uses both PTAP and ubiquitin to redirect the endosomal ESCRT-I complex to the plasma membrane (Figure 6). If pirating a regular cellular function, what contributes the PTAP sequence to TSG101/ESCRT-I normally at the endosomal sorting site?

Intriguingly, Hrs has a PSAP sequence, and in fact, direct interaction between Hrs and ESCRT-I has been reported (12). Thus, Hrs could be part of the machinery that recruits ESCRT-I to the surface of the nascent multivesicular body. Hrs would play a sequential, but coupled role, beginning at the early endosome by segregating ubiquitinated cargo away from the recycling population and then driving ESCRT-I translocation to initiate intralumenal budding (Figure 6). This is in full accord with the limited number of intralumenal membrane elements seen in early endosomes containing Hrs-clathrin bilayered coats (11). This model would account for the observed phenotype of $Hrs^{-/-}$ flies and mice, and it integrates ubiquitin-dependent lysosomal sorting with the morphological transition that accompanies early to late endosome maturation.

An unexpected observation is that the UIM-sequence also directs monoubiquitination of proteins containing this sequence, or of extraneous proteins fused to a UIM region, as the acceptor lysine is not contained within the UIM (255, 271–273). The E3 ligase involved appears to be Nedd4 (255, 273), but the significance of this modification has not been precisely defined. It has been suggested that the UIM could participate directly in E3 binding by associating with the ligase-bound ubiquitin, or bind to the newly conjugated ubiquitin, in either case possibly favoring only monoubiquitination (255). Although a substantial amount of eps15/eps15R is monoubiquitinated upon addition of EGF, the fraction of ubiquitinated epsin and Hrs is, however, low (255, 271). Further, the effect of monoubiquitination on the capacity of UIM-containing proteins to bind ubiquitin has not yet been reported. The observed monoubiquitin modification could be linked to the requirement for a functional ubiquitination machinery to facilitate endocytosis of GH receptor (220) or Ste2p-ubiquitin (161). Ubiquitination of the endocytic machinery could amplify, or create, via UIM-ubiquitin interactions, a specifically required protein-interaction web at the appropriate sorting site. One argument for this is that the amount of ubiquitinated protein amassed on endosomes in Hrs-overexpressing cells stimulated with EGF cannot be accounted for by ubiquitinated EGF receptor alone (205). In this model, mirroring HIV-1 Gag, monoubiquitinated Hrs and internal PSAP sequence, and not ubiquitinated cargo itself, might be the physiologic binding partner of the ESCRT-I complex. The Hrs UIM could synchronously bind a ubiquitin tag attached to cargo. That ESCRT-I might not necessarily bind ubiquitinated cargo directly is indicated by the ability of overexpressed ESCRT-II complex, which functions downstream of ESCRT-I in multivesicular body formation, to rescue sorting in an ESCRT-I-mutant (*vps23*Δ) strain (274). As ESCRT-II is not known to interact with ubiquitin, perhaps the Vps27p and Hse1p UIMs mediate the critical cargo sorting (253).

Some of the ubiquitin signal on early endosomes could also be derived from eps15, as this protein binds to Hrs2 (275) which, in neurons, is enriched focally on multivesicular endosomes (275), highly reminiscent of the bilayered clathrin scaffold sorting sites (11). A physiological role for direct ubiquitination of the endocytic machinery is also supported by functional studies of the *Drosophila*

epsin homologue, Liquid facets (Lqf) (276). Fat facets, a deubiquitinating enzyme, interacts genetically with Lqf and is specifically required for Lqf function, possibly by recycling the ubiquitinated epsin for further rounds of endocytic sorting (276, 277). In this regard, it has also been proposed that ubiquitination of UIM-bearing endocytic proteins causes functional inactivation by the conjugated ubiquitin engaging the UIM (273).

Finally, Doa4p appears to deubiquitinate cargo proteins just before final packaging into the involuting buds (200) (Figure 6), accounting for the strong sorting phenotype of *doa4* mutants; in fact, Doa4p accumulates in the expanded class E structure in class E mutants (194). However, deubiquitination is not essential for progression into intralumenal vesicles. Ubiquitinated forms of yeast and mammalian proteins are detectable in multivesicular bodies and vacuoles (204, 205, 253).

CONCLUDING REMARKS

The information reviewed in this article illustrates how far we have come in the understanding of the molecular mechanisms of signal-mediated protein trafficking to endosomes and lysosomes. Many issues remain unresolved, but with the powerful tools now available we can expect these to be elucidated in the not too distant future. Among the outstanding questions are the identification of the complete repertoire of sorting signals and their binding partners, the explanation of the structural bases for the recognition of all signals, the demonstration of the exact sorting events mediated by particular signals and recognition proteins, the detailed molecular description of complex pathways of signal addition and detection (as is the case for ubiquitin), the regulation of signal recognition, and the contribution of luminal/extracellular and transmembrane domains to sorting. The current knowledge of the mechanisms of signal-mediated protein sorting has already contributed to the elucidation of the pathogenesis of genetic diseases such as familial hypercholesterolemia, Hermansky-Pudlak syndrome type 2, and Liddle's syndrome. A more complete understanding of protein sorting is likely to illuminate the participation of the endosomal-lysosomal system in more common physiological and pathological processes as well as to provide novel avenues for therapeutic intervention.

ACKNOWLEDGMENTS

We thank Mickey Marks, Markus Boehm, and James Hurley for helpful discussions and critical reading of the manuscript, and Matthew Wood and Saurav Misra for help with the crystallography figures.

The *Annual Review of Biochemistry* is online at http://biochem.annualreviews.org

LITERATURE CITED

1. Trowbridge IS, Collawn JF, Hopkins CR. 1993. *Annu. Rev. Cell Biol.* 9:29–61
2. Sandoval IV, Bakke O. 1994. *Trends Cell Biol.* 4:292–97
3. Bonifacino JS, Dell'Angelica EC. 1999. *J. Cell Biol.* 145:923–26
4. Orsel JG, Sincock PM, Krise JP, Pfeffer SR. 2000. *Proc. Natl. Acad. Sci. USA* 97:9047–51
5. Hicke L. 2001. *Cell* 106:527–30
6. Roth TF, Cutting JA, Atlas SB. 1976. *J. Supramol. Struct.* 4:527–48
7. Anderson RG, Brown MS, Goldstein JL. 1977. *Cell* 10:351–64
8. Schulze-Lohoff E, Hasilik A, von Figura K. 1985. *J. Cell Biol.* 101:824–29
9. Klumperman J, Hille A, Veenendaal T, Oorschot V, Stoorvogel W, et al. 1993. *J. Cell Biol.* 121:997–1010
10. Futter CE, Gibson A, Allchin EH, Maxwell S, Ruddock LJ, et al. 1998. *J. Cell Biol.* 141:611–23
11. Sachse M, Urbe S, Oorschot V, Strous GJ, Klumperman J. 2002. *Mol. Biol. Cell* 13:1313–28
12. Raiborg C, Bache KG, Gillooly DJ, Madshus IH, Stang E, Stenmark H. 2002. *Nat. Cell Biol.* 4:394–98
13. Pearse BM, Robinson MS. 1990. *Annu. Rev. Cell Biol.* 6:151–71
14. Kirchhausen T. 2000. *Annu. Rev. Biochem.* 69:699–727
15. Kirchhausen T. 1999. *Annu. Rev. Cell Dev. Biol.* 15:705–32
16. Slepnev VI, De Camilli P. 2000. *Nat. Rev. Neurosci.* 1:161–72
17. Davis CG, Lehrman MA, Russell DW, Anderson RG, Brown MS, Goldstein JL. 1986. *Cell* 45:15–24
18. Chen J-J, Goldstein JL, Brown MS. 1990. *J. Biol. Chem.* 265:3116–23
19. Lazarovits J, Roth M. 1988. *Cell* 53: 743–52
20. Lobel P, Fujimoto K, Ye RD, Griffiths G, Kornfeld S. 1989. *Cell* 57:787–96
21. McGraw TE, Maxfield FR. 1990. *Cell Regul.* 1:369–77
22. Collawn JF, Stangel M, Kuhn LA, Esekogwu V, Jing SQ, et al. 1990. *Cell* 63:1061–72
23. Alvarez E, Girones N, Davis RJ. 1990. *Biochem. J.* 267:31–35
24. Breitfeld PP, Casanova JE, McKinnan WC, Mostov KE. 1990. *J. Biol. Chem.* 265:13750–57
25. Canfield WM, Johnson KF, Ye RD, Gregory W, Kornfeld S. 1991. *J. Biol. Chem.* 266:5682–88
26. Jadot M, Canfield WM, Gregory W, Kornfeld S. 1992. *J. Biol. Chem.* 267: 11069–77
27. Williams MA, Fukuda M. 1990. *J. Cell Biol.* 111:955–66
28. Harter C, Mellman I. 1992. *J. Cell Biol.* 117:311–25
29. Hunziker W, Harter C, Matter K, Mellman I. 1991. *Cell* 66:907–20
30. Rajasekaran AK, Humphrey JS, Wagner M, Miesenböck G, Le Bivic A, et al. 1994. *Mol. Biol. Cell* 5:1093–103
31. Collawn JF, Kuhn LA, Liu L-FS, Tainer JA, Trowbridge IS. 1991. *EMBO J.* 10:3247–53
32. Perez RG, Soriano S, Hayes JD, Ostaszewski B, Xia W, et al. 1999. *J. Biol. Chem.* 274:18851–56
33. Kibbey RG, Rizo J, Gierasch LM, Anderson RG. 1998. *J. Cell Biol.* 142: 59–67
34. Boll W, Rapoport I, Brunner C, Modis Y, Prehn S, Kirchhausen T. 2002. *Traffic* 3:590–600
35. Musacchio A, Smith CJ, Roseman AM, Harrison SC, Kirchhausen T, Pearse BM. 1999. *Mol. Cell* 3:761–70
36. Pearse BM. 1988. *EMBO J.* 7:3331–36
37. Warren RA, Green FA, Stenberg PE, Enns CA. 1998. *J. Biol. Chem.* 273: 17056–63
38. Howell BW, Lanier LM, Frank R,

Gertler FB, Cooper JA. 1999. *Mol. Cell. Biol.* 19:5179–88
39. Morris SM, Cooper JA. 2001. *Traffic* 2:111–23
40. Mishra SK, Keyel PA, Hawryluk MJ, Agostinelli NR, Watkins SC, Traub LM. 2002. *EMBO J.* 21:4915–26
41. Herz J, Bock HH. 2002. *Annu. Rev. Biochem.* 71:405–34
42. Oleinikov AV, Zhao J, Makker SP. 2000. *Biochem. J.* 347(Part 3):613–21
43. Gotthardt M, Trommsdorff M, Nevitt MF, Shelton J, Richardson JA, et al. 2000. *J. Biol. Chem.* 275:25616–24
44. Morris SM, Tallquist MD, Rock CO, Cooper JA. 2002. *EMBO J.* 21:1555–64
45. Nykjaer A, Dragun D, Walther D, Vorum H, Jacobsen C, et al. 1999. *Cell* 96:507–15
46. Verroust PJ, Kozyraki R. 2001. *Curr. Opin. Nephrol. Hypertens.* 10:33–38
47. Garcia CK, Wilund K, Arca M, Zuliani G, Fellin R, et al. 2001. *Science* 292: 1394–98
48. Al-Kateb H, Bahring S, Hoffmann K, Strauch K, Busjahn A, et al. 2002. *Circ. Res.* 90:951–58
49. Norman D, Sun XM, Bourbon M, Knight BL, Naoumova RP, Soutar AK. 1999. *J. Clin. Invest.* 104:619–28
50. Berdnik D, Torok T, Gonzalez-Gaitan M, Knoblich J. 2002. *Dev. Cell* 3:221–31
51. Santolini E, Puri C, Salcini AE, Gagliani MC, Pelicci PG, et al. 2000. *J. Cell Biol.* 151:1345–52
52. Marks MS, Roche PA, van Donselaar E, Woodruff L, Peters PJ, Bonifacino JS. 1995. *J. Cell Biol.* 131:351–69
53. Gough NR, Zweifel ME, Martinez-Augustin O, Aguilar RC, Bonifacino JS, Fambrough DM. 1999. *J. Cell Sci.* 112: 4257–69
54. Rous BA, Reaves BJ, Ihrke G, Briggs JA, Gray SR, et al. 2002. *Mol. Biol. Cell* 13:1071–82
55. Rohrer J, Schweizer A, Russell D, Kornfeld S. 1996. *J. Cell Biol.* 132:565–76
56. Ohno H, Stewart J, Fournier MC, Bosshart H, Rhee I, et al. 1995. *Science* 269:1872–75
57. Boll W, Ohno H, Zhou SY, Rapoport I, Cantley LC, et al. 1996. *EMBO J.* 15:5789–95
58. Höning S, Griffith J, Geuze HJ, Hunziker W. 1996. *EMBO J.* 15:5230–39
59. Ohno H, Fournier MC, Poy G, Bonifacino JS. 1996. *J. Biol. Chem.* 271: 29009–15
60. Hirst J, Bright NA, Rous B, Robinson MS. 1999. *Mol. Biol. Cell* 10:2787–802
61. Stephens DJ, Crump CM, Clarke AR, Banting G. 1997. *J. Biol. Chem.* 272: 14104–9
62. Rapoport I, Miyazaki M, Boll W, Duckworth B, Cantley LC, et al. 1997. *EMBO J.* 16:2240–50
63. Ohno H, Aguilar RC, Yeh D, Taura D, Saito T, Bonifacino JS. 1998. *J. Biol. Chem.* 273:25915–21
64. Aguilar RC, Boehm M, Gorshkova I, Crouch RJ, Tomita K, et al. 2001. *J. Biol. Chem.* 276:13145–52
65. Ohno H, Tomemori T, Nakatsu F, Okazaki Y, Aguilar RC, et al. 1999. *FEBS Lett.* 449:215–20
66. Boehm M, Bonifacino JS. 2002. *Gene* 286:175–86
67. Aguilar RC, Ohno H, Roche KC, Bonifacino JS. 1997. *J. Biol. Chem.* 272: 2760–66
68. Owen DJ, Evans PR. 1998. *Science* 282: 1327–32
69. Collins BM, McCoy AJ, Kent HM, Evans PR, Owen DJ. 2002. *Cell* 109: 523–35
70. Andrews J, Smith M, Merakovsky J, Coulson M, Hannan F, Kelly LE. 1996. *Genetics* 143:1699–711
71. Martina JA, Bonangelino CJ, Aguilar RC, Bonifacino JS. 2001. *J. Cell Biol.* 28:1111–20
72. Walther K, Krauss M, Diril MK, Lemke S, Ricotta D, et al. 2001. *EMBO Rep.* 2:634–40
73. Owen DJ, Setiadi H, Evans PR, McEver RP, Green SA. 2001. *Traffic* 2:105–10

74. Nesterov A, Carter RE, Sorkina T, Gill GN, Sorkin A. 1999. *EMBO J.* 18:2489–99
75. Rohde G, Wenzel D, Haucke V. 2002. *J. Cell Biol.* 158:209–14
76. Pauloin A, Thurieau C. 1993. *Biochem. J.* 296:409–15
77. Olusanya O, Andrews PD, Swedlow JR, Smythe E. 2001. *Curr. Biol.* 11:896–900
78. Conner SD, Schmid SL. 2002. *J. Cell Biol.* 156:921–29
79. Ricotta D, Conner SD, Schmid SL, von Figura K, Höning S. 2002. *J. Cell Biol.* 156:791–95
80. Gaidarov I, Smith ME, Domin J, Keen JH. 2001. *Mol. Cell.* 7:443–49
81. Haffner C, Di Paolo G, Rosenthal JA, de Camilli P. 2000. *Curr. Biol.* 10:471–74
82. Gaidarov I, Chen Q, Falck JR, Reddy KK, Keen JH. 1996. *J. Biol. Chem.* 271: 20922–29
83. Haucke V, De Camilli P. 1999. *Science* 285:1268–71
84. Shiratori T, Miyatake S, Ohno H, Nakaseko C, Isono K, et al. 1997. *Immunity* 6:583–89
85. Zhang Y, Allison JP. 1997. *Proc. Natl. Acad. Sci. USA* 94:9273–78
86. Chuang E, Alegre ML, Duckett CS, Noel PJ, Vander Heiden MG, Thompson CB. 1997. *J. Immunol.* 159:144–51
87. Schaefer AW, Kamei Y, Kamiguchi H, Wong EV, Rapoport I, et al. 2002. *J. Cell Biol.* 157:1223–32
88. Madrid R, Le Maout S, Barrault MB, Janvier K, Benichou S, Merot J. 2001. *EMBO J.* 20:7008–21
89. Letourneur F, Klausner RD. 1992. *Cell* 69:1143–57
90. Johnson KF, Kornfeld S. 1992. *J. Biol. Chem.* 267:17110–15
91. Johnson KF, Kornfeld S. 1992. *J. Cell Biol.* 119:249–57
92. Chen HJ, Yuan J, Lobel P. 1997. *J. Biol. Chem.* 272:7003–12
93. Dietrich J, Hou X, Wegener AM, Geisler C. 1994. *EMBO J.* 13:2156–66
94. von Essen M, Menne C, Nielsen BL, Lauritsen JP, Dietrich J, et al. 2002. *J. Immunol.* 168:4519–23
95. Pitcher C, Höning S, Fingerhut A, Bowers K, Marsh M. 1999. *Mol. Biol. Cell* 10:677–91
96. Bresnahan PA, Yonemoto W, Ferrell S, Williams-Herman D, Geleziunas R, Greene WC. 1998. *Curr. Biol.* 8:1235–38
97. Greenberg M, DeTulleo L, Rapoport I, Skowronski J, Kirchhausen T. 1998. *Curr. Biol.* 8:1239–42
98. Vowels JJ, Payne GS. 1998. *EMBO J.* 17:2482–93
99. Matter K, Yamamoto EM, Mellman I. 1994. *J. Cell Biol.* 126:991–1004
100. Miranda KC, Khromykh T, Christy P, Le TL, Gottardi CJ, et al. 2001. *J. Biol. Chem.* 276:22565–72
101. Pond L, Kuhn LA, Teyton L, Schutze MP, Tainer JA, et al. 1995. *J. Biol. Chem.* 270:19989–97
102. Sandoval IV, Martinez-Arca S, Valdueza J, Palacios S, Holman GD. 2000. *J. Biol. Chem.* 275:39874–85
103. Geisler C, Dietrich J, Nielsen BL, Kastrup J, Lauritsen JP, et al. 1998. *J. Biol. Chem.* 273:21316–23
104. Heilker R, Manning-Krieg U, Zuber JF, Spiess M. 1996. *EMBO J.* 15:2893–99
105. Dietrich J, Kastrup J, Nielsen BL, Odum N, Geisler C. 1997. *J. Cell Biol.* 138: 271–81
106. Höning S, Sandoval IV, von Figura K. 1998. *EMBO J.* 17:1304–14
107. Fujita H, Saeki M, Yasunaga K, Ueda T, Imoto T, Himeno M. 1999. *Biochem. Biophys. Res. Commun.* 255:54–58
108. Hofmann MW, Höning S, Rodionov D, Dobberstein B, von Figura K, Bakke O. 1999. *J. Biol. Chem.* 274:36153–58
109. Peden AA, Park GY, Scheller RH. 2001. *J. Biol. Chem.* 276:49183–87
110. Le Borgne R, Alconada A, Bauer U, Hoflack B. 1998. *J. Biol. Chem.* 273: 29451–61
111. Huizing M, Sarangarajan R, Strovel E,

Zhao Y, Gahl WA, Boissy RE. 2001. *Mol. Biol. Cell* 12:2075–85
112. Caplan S, Dell'Angelica EC, Gahl WA, Bonifacino JS. 2000. *Immunol. Lett.* 72:113–17
113. Kongsvik TL, Höning S, Bakke O, Rodionov DG. 2002. *J. Biol. Chem.* 277: 16484–88
114. Marks MS, Woodruff L, Ohno H, Bonifacino JS. 1996. *J. Cell Biol.* 135:341–54
115. Rodionov DG, Bakke O. 1998. *J. Biol. Chem.* 273:6005–8
116. Deleted in proof
117. Craig HM, Reddy TR, Riggs NL, Dao PP, Guatelli JC. 2000. *Virology* 271:9–17
118. Bremnes T, Lauvrak V, Lindqvist B, Bakke O. 1998. *J. Biol. Chem.* 273: 8638–45
119. Rapoport I, Chen YC, Cupers P, Shoelson SE, Kirchhausen T. 1998. *EMBO J.* 17:2148–55
120. Puertollano R, Aguilar RC, Gorshkova I, Crouch RJ, Bonifacino JS. 2001. *Science* 292:1712–16
121. Doray B, Ghosh P, Griffith J, Geuze HJ, Kornfeld S. 2002. *Science.* 6:1700–3
122. Puertollano R, van der Wel NN, Green LE, Eisenberg E, Peters PJ, Bonifacino JS. 2003. *Mol. Biol. Cell* 14:1545–57
123. Schweizer A, Kornfeld S, Rohrer J. 1996. *J. Cell Biol.*132:577–84
124. Höning S, Sosa M, Hille-Rehfeld A, von Figura K. 1997. *J. Biol. Chem.* 272: 19884–90
125. Zhu Y, Doray B, Poussu A, Lehto VP, Kornfeld S. 2001. *Science* 292:1716–18
126. Nielsen MS, Madsen P, Christensen EI, Nykjaer A, Gliemann J, et al. 2001. *EMBO J.* 20:2180–90
127. Takatsu H, Katoh Y, Shiba Y, Nakayama K. 2001. *J. Biol. Chem.* 276:28541–45
128. Doray B, Bruns K, Ghosh P, Kornfeld S. 2002. *J. Biol. Chem.* 277:18477–82
129. Jacobsen L, Madsen P, Nielsen MS, Geraerts WP, Gliemann J, et al. 2002. *FEBS Lett.* 511:155–58
130. He X, Chang WP, Koelsch G, Tang J. 2002. *FEBS Lett.* 524:183–87
131. Boman AL, Zhang CJ, Zhu X, Kahn RA. 2000. *Mol. Biol. Cell.* 11:1241–55
132. Dell'Angelica EC, Puertollano R, Mullins C, Aguilar RC, Vargas JD, et al. 2000. *J. Cell Biol.* 149:81–94
133. Hirst J, Lui WW, Bright NA, Totty N, Seaman MN, Robinson MS. 2000. *J. Cell Biol.* 149:67–80
134. Poussu A, Lohi O, Lehto VP. 2000. *J. Biol. Chem.* 275:7176–83
135. Misra S, Puertollano R, Kato Y, Bonifacino JS, Hurley JH. 2002. *Nature.* 415: 933–37
136. Kato Y, Misra S, Puertollano R, Hurley JH, Bonifacino JS. 2002. *Nat. Struct. Biol.* 9:532–36
137. Zhdankina O, Strand NL, Redmond JM, Boman AL. 2001. *Yeast* 18:1–18
138. Costaguta G, Stefan CJ, Bensen ES, Emr SD, Payne GS. 2001. *Mol. Biol. Cell* 12:1885–96
139. Mullins C, Bonifacino JS. 2001. *Mol. Cell. Biol.* 21:7981–94
140. Shiba T, Takatsu H, Nogi T, Matsugaki N, Kawasaki M, et al. 2002. *Nature* 415: 937–41
141. Meresse S, Ludwig T, Frank R, Hoflack B. 1990. *J. Biol. Chem.* 265:18833–42
142. Meresse S, Hoflack B. 1993. *J. Cell Biol.* 120:67–75
143. Rosorius O, Mieskes G, Issinger OG, Korner C, Schmidt B, et al. 1993. *Biochem. J.* 292:833–38
144. Korner C, Herzog A, Weber B, Rosorius O, Hemer F, et al. 1994. *J. Biol. Chem.* 269:16529–32
145. Le Borgne R, Schmidt A, Mauxion F, Griffiths G, Hoflack B. 1993. *J. Biol. Chem.* 268:22552–56
146. Doray B, Bruns K, Ghosh P, Kornfeld SA. 2002. *Proc. Natl. Acad. Sci. USA* 99:8072–77
147. Jones BG, Thomas L, Molloy SS, Thulin CD, Fry MD, et al. 1995. *EMBO J.* 14:5869–83

148. Wan L, Molloy SS, Thomas L, Liu GP, Xiang Y, et al. 1998. *Cell* 94:205–16
149. Crump CM, Xiang Y, Thomas L, Gu F, Austin C, et al. 2001. *EMBO J.* 20:2191–201
150. Olson LJ, Zhang J, Dahms NM, Kim JJ. 2002. *J. Biol. Chem.* 277:10156–61
151. Borden LA, Einstein R, Gabel CA, Maxfield FR. 1990. *J. Biol. Chem.* 265:8497–504
152. Griffiths G, Hoflack B, Simons K, Mellman I, Kornfeld S. 1988. *Cell* 52: 329–41
153. Rohrer J, Schweizer A, Johnson KF, Kornfeld S. 1995. *J. Cell Biol.* 130: 1297–306
154. Schweizer A, Kornfeld S, Rohrer J. 1997. *Proc. Natl. Acad. Sci. USA* 94:14471–76
155. Schweizer A, Stahl PD, Rohrer J. 2000. *J. Biol. Chem.* 275:29694–700
156. Diaz E, Pfeffer SR. 1998. *Cell* 93:433–43
157. Krise JP, Sincock PM, Orsel JG, Pfeffer SR. 2000. *J. Biol. Chem.* 275:25188–93
158. Carroll KS, Hanna J, Simon I, Krise J, Barbero P, Pfeffer SR. 2001. *Science* 292:1373–76
159. Miura S, Gan JW, Brzostowski J, Parisi MJ, Schultz CJ, et al. 2002. *J. Biol. Chem.* 277:32253–57
160. Tan PK, Howard JP, Payne GS. 1996. *J. Cell Biol.* 135:1789–800
161. Dunn R, Hicke L. 2001. *J. Biol. Chem.* 276:25974–81
162. Howard JP, Hutton JL, Olson JM, Payne GS. 2002. *J. Cell Biol.* 157:315–26
163. Tang HY, Xu J, Cai M. 2000. *Mol. Cell. Biol.* 20:12–25
164. Munn AL. 2001. *Biochim. Biophys. Acta* 1535:236–57
165. Pickart CM. 2001. *Annu. Rev. Biochem.* 70:503–33
166. Thrower JS, Hoffman L, Rechsteiner M, Pickart CM. 2000. *EMBO J.* 19:94–102
167. Hicke L, Riezman H. 1996. *Cell* 84:277–87
168. Roth AF, Davis NG. 1996. *J. Cell Biol.* 134:661–74
169. Bonifacino JS, Weissman AM. 1998. *Annu. Rev. Cell Dev. Biol.* 14:19–57
170. Hicke L. 1999. *Trends Cell Biol.* 9:107–12
171. Rotin D, Staub O, Haguenauer-Tsapis R. 2000. *J. Membr. Biol.* 176:1–17
172. Strous GJ, van Kerkhof P, Govers R, Ciechanover A, Schwartz AL. 1996. *EMBO J.* 15:3806–12
173. Strous GJ, van Kerkhof P, Govers R, Rotwein P, Schwartz AL. 1997. *J. Biol. Chem.* 272:40–43
174. van Kerkhof P, Govers R, dos Santos CMA, Strous GJ. 2000. *J. Biol. Chem.* 275:1575–80
175. van Kerkhof P, dos Santos CMA, Sachse M, Klumperman J, Bu GJ, Strous GJ. 2001. *Mol. Biol. Cell* 12:2556–66
176. van Kerkhof P, Sachse M, Klumperman J, Strous GJ. 2001. *J. Biol. Chem.* 276: 3778–84
177. Levkowitz G, Waterman H, Ettenberg SA, Katz M, Tsygankov AY, et al. 1999. *Mol. Cell* 4:1029–40
178. Joazeiro CA, Wing SS, Huang H, Leverson JD, Hunter T, Liu YC. 1999. *Science* 286:309–12
179. Waterman H, Katz M, Rubin C, Shtiegman K, Lavi S, et al. 2002. *EMBO J.* 21:303–13
180. Hammond DE, Urbe S, Vande Woude GF, Clague MJ. 2001. *Oncogene* 20:2761–70
181. Lee PS, Wang Y, Dominguez MG, Yeung YG, Murphy MA, et al. 1999. *EMBO J.* 18:3616–28
182. Wang Y, Yeung YG, Stanley ER. 1999. *J. Cell. Biochem.* 72:119–34
183. Booth JW, Kim MK, Jankowski A, Schreiber AD, Grinstein S. 2002. *EMBO J.* 21:251–58
184. Deblandre GA, Lai EC, Kintner C. 2001. *Dev. Cell* 1:795–806
185. Lai EC, Deblandre GA, Kintner C, Rubin GM. 2001. *Dev. Cell* 1:783–94
186. Yeh E, Dermer M, Commisso C, Zhou L, McGlade CJ, Boulianne GL. 2001. *Curr. Biol.* 11:1675–79

187. Buttner C, Sadtler S, Leyendecker A, Laube B, Griffon N, et al. 2001. *J. Biol. Chem.* 276:42978–85
188. Fujita Y, Krause G, Scheffner M, Zechner D, Leddy HE, et al. 2002. *Nat. Cell Biol.* 4:222–31
189. Panigada M, Porcellini S, Barbier E, Hoeflinger S, Cazenave PA, et al. 2002. *J. Exp. Med.* 195:1585–97
190. Burbea M, Dreier L, Dittman JS, Grunwald ME, Kaplan JM. 2002. *Neuron* 35:107–20
191. Hewitt EW, Duncan L, Mufti D, Baker J, Stevenson PG, Lehner PJ. 2002. *EMBO J.* 21:2418–29
192. Lorenzo ME, Jung JU, Ploegh HL. 2002. *J. Virol.* 76:5522–31
193. Wilkinson KD. 2000. *Semin. Cell Dev. Biol.* 11:141–48
194. Amerik AY, Nowak J, Swaminathan S, Hochstrasser M. 2000. *Mol. Biol. Cell* 11:3365–80
195. Losko S, Kopp F, Kranz A, Kolling R. 2001. *Mol. Biol. Cell* 12:1047–59
196. Reggiori F, Pelham HR. 2001. *EMBO J.* 20:5176–86
197. Terrell J, Shih S, Dunn R, Hicke L. 1998. *Mol. Cell.* 1:193–202
198. Dupre S, Haguenauer-Tsapis R. 2001. *Mol. Cell. Biol.* 21:4482–94
199. Horak J, Wolf DH. 2001. *J. Bacteriol.* 183:3083–88
200. Katzmann DJ, Babst M, Emr SD. 2001. *Cell* 106:145–55
201. Urbanowski JL, Piper RC. 2001. *Traffic* 2:622–30
202. Felder S, Miller K, Moehren G, Ullrich A, Schlessinger J, Hopkins CR. 1990. *Cell* 61:623–34
203. Maeda K, Kato Y, Sugiyama Y. 2002. *J. Control Release* 82:71–82
204. Longva KE, Blystad FD, Stang E, Larsen AM, Johannessen LE, Madshus IH. 2002. *J. Cell Biol.* 156:843–54
205. Bishop N, Horman A, Woodman P. 2002. *J. Cell Biol.* 157:91–101
206. Marchese A, Benovic JL. 2001. *J. Biol. Chem.* 276:45509–12
207. Rocca A, Lamaze C, Subtil A, Dautry-Varsat A. 2001. *Mol. Biol. Cell* 12:1293–301
208. Yu A, Malek TR. 2001. *J. Biol. Chem.* 276:381–85
209. Roberg KJ, Rowley N, Kaiser CA. 1997. *J. Cell Biol.* 137:1469–82
210. Beck T, Schmidt A, Hall MN. 1999. *J. Cell Biol.* 146:1227–38
211. Helliwell SB, Losko S, Kaiser CA. 2001. *J. Cell Biol.* 153:649–62
212. Soetens O, De Craene JO, Andre B. 2001. *J. Biol. Chem.* 276:43949–57
213. Roth AF, Davis NG. 2000. *J. Biol. Chem.* 275:8143–53
214. Shih SC, Sloper-Mould KE, Hicke L. 2000. *EMBO J.* 19:187–98
215. Nakatsu F, Sakuma M, Matsuo Y, Arase H, Yamasaki S, et al. 2000. *J. Biol. Chem.* 275:26213–19
216. Hicke L. 2001. *Nat. Rev. Mol. Cell. Biol.* 2:195–201
217. Marchal C, Haguenauer-Tsapis R, Urban-Grimal D. 2000. *J. Biol. Chem.* 275:23608–14
218. Galan JM, Haguenauer-Tsapis R. 1997. *EMBO J.* 16:5847–54
219. Beal R, Deveraux Q, Xia G, Rechsteiner M, Pickart C. 1996. *Proc. Natl. Acad. Sci. USA* 93:861–66
220. Govers R, ten Broeke T, van Kerkhof P, Schwartz AL, Strous GJ. 1999. *EMBO J.* 18:28–36
221. Petrelli A, Gilestro GF, Lanzardo S, Comoglio PM, Migone N, Giordano S. 2002. *Nature* 416:187–90
222. Soubeyran P, Kowanetz K, Szymkiewicz I, Langdon WY, Dikic I. 2002. *Nature* 416:183–87
223. Jackson PK, Eldridge AG, Freed E, Furstenthal L, Hsu JY, et al. 2000. *Trends Cell Biol.* 10:429–39
224. Kay BK, Williamson MP, Sudol M. 2000. *FASEB J.* 14:231–41
225. Dunn R, Hicke L. 2001. *Mol. Biol. Cell* 12:421–35
226. Gajewska B, Kaminska J, Jesionowska

A, Martin NC, Hopper AK, Zoladek T. 2001. *Genetics* 157:91–101
227. Kölling R, Losko S. 1997. *EMBO J.* 16:2251–61
228. Miller WE, Lefkowitz RJ. 2001. *Curr. Opin. Cell Biol.* 13:139–45
229. Shenoy SK, McDonald PH, Kohout TA, Lefkowitz RJ. 2001. *Science* 294:1307–13
230. Alarcon-Vargas D, Ronai Z. 2002. *Carcinogenesis* 23:541–47
231. Staub O, Abriel H, Plant P, Ishikawa T, Kanelis V, et al. 2000. *Kidney Int.* 57:809–15
232. Asher C, Chigaev A, Garty H. 2001. *Biochem. Biophys. Res. Commun.* 286: 1228–31
233. Lott JS, Coddington-Lawson SJ, Teesdale-Spittle PH, McDonald FJ. 2002. *Biochem. J.* 361:481–88
234. Dinudom A, Harvey KF, Komwatana P, Jolliffe CN, Young JA, et al. 2001. *J. Biol. Chem.* 276:13744–49
235. Cook DI, Dinudom A, Komwatana P, Kumar S, Young JA. 2002. *Cell Biochem. Biophys.* 36:105–13
236. Shimkets RA, Warnock DG, Bositis CM, Nelson-Williams C, Hansson JH, et al. 1994. *Cell* 79:407–14
237. Hansson JH, Nelson-Williams C, Suzuki H, Schild L, Shimkets R, et al. 1995. *Nat. Genet.* 11:76–82
238. Shi HK, Asher C, Chigaev A, Yung Y, Reuveny E, et al. 2002. *J. Biol. Chem.* 277:13539–47
239. Myat A, Henry P, McCabe V, Flintoft L, Rotin D, Tear G. 2002. *Neuron* 35: 447–59
240. Schwake M, Friedrich T, Jentsch TJ. 2001. *J. Biol. Chem.* 276:12049–54
241. Qiu L, Joazeiro C, Fang N, Wang HY, Elly C, et al. 2000. *J. Biol. Chem.* 275: 35734–37
242. De Wit R, Makkinje M, Boonstra J, Verkleij AJ, Post JA. 2001. *FASEB J.* 15:306–8
243. Stang E, Johannessen LE, Knardal SL, Madshus IH. 2000. *J. Biol. Chem.* 275: 13940–47
244. Burke P, Schooler K, Wiley HS. 2001. *Mol. Biol. Cell* 12:1897–910
245. de Melker AA, van der Horst G, Calafat J, Jansen H, Borst J. 2001. *J. Cell Sci.* 114:2167–78
246. Miyake S, Mullane-Robinson KP, Lill NL, Douillard P, Band H. 1999. *J. Biol. Chem.* 274:16619–28
247. Jehn BM, Dittert I, Beyer S, von der Mark K, Bielke W. 2002. *J. Biol. Chem.* 277:8033–40
248. Wang G, McCaffery JM, Wendland B, Dupre S, Haguenauer-Tsapis R, Huibregtse JM. 2001. *Mol. Cell. Biol.* 21:3564–75
249. Reggiori F, Pelham HR. 2002. *Nat. Cell Biol.* 4:117–23
250. Reggiori F, Black MW, Pelham HR. 2000. *Mol. Biol. Cell* 11:3737–49
251. Young P, Deveraux Q, Beal RE, Pickart CM, Rechsteiner M. 1998. *J. Biol. Chem.* 273:5461–67
252. Hofmann K, Falquet L. 2001. *Trends Biochem. Sci.* 26:347–50
253. Bilodeau PS, Urbanowski JL, Winistorfer SC, Piper RC. 2002. *Nat. Cell Biol.* 4:534–39
254. Lloyd TE, Atkinson R, Wu MN, Zhou Y, Pennetta G, Bellen HJ. 2002. *Cell* 108: 261–69
255. Polo S, Sigismund S, Faretta M, Guidi M, Capua MR, et al. 2002. *Nature* 416: 451–55
256. Shih SC, Katzmann DJ, Schnell JD, Sutanto M, Emr SD, Hicke L. 2002. *Nat. Cell Biol.* 4:389–93
257. Brett TJ, Traub LM, Fremont DH. 2002. *Structure* 10:797–809
258. Kalthoff C, Alves J, Urbanke C, Knorr R, Ungewickell EJ. 2001. *J. Biol. Chem.* 277:8209–16
259. Wilkinson CR, Seeger M, Hartmann-Petersen R, Stone M, Wallace M, et al. 2001. *Nat. Cell Biol.* 3:939–43
260. Futter CE, Collinson LM, Backer JM, Hopkins CR. 2001. *J. Cell Biol.* 155: 1251–64

261. Raiborg C, Bache KG, Mehlum A, Stang E, Stenmark H. 2001. *EMBO J.* 20:5008–21
262. Chin LS, Raynor MC, Wei X, Chen HQ, Li L. 2001. *J. Biol. Chem.* 276:7069–78
263. Komada M, Soriano P. 1999. *Genes Dev.* 13:1475–85
264. Garrus JE, von Schwedler UK, Pornillos OW, Morham SG, Zavitz KH, et al. 2001. *Cell* 107:55–65
265. Babst M, Odorizzi G, Estepa EJ, Emr SD. 2000. *Traffic* 1:248–58
266. Patnaik A, Chau V, Wills JW. 2000. *Proc. Natl. Acad. Sci. USA* 97:13069–74
267. Schubert U, Ott DE, Chertova EN, Welker R, Tessmer U, et al. 2000. *Proc. Natl. Acad. Sci. USA* 97:13057–62
268. Yasuda J, Hunter E, Nakao M, Shida H. 2002. *EMBO Rep.* 3:636–40
269. Strack B, Calistri A, Gottlinger HG. 2002. *J. Virol.* 76:5472–79
270. Pornillos O, Alam SL, Rich RL, Myszka DG, Davis DR, Sundquist WI. 2002. *EMBO J.* 21:2397–406
271. Klapisz E, Sorokina I, Lemeer S, Pijnenburg M, Verkleij AJ, van Bergen en Henegouwen PMP. 2002. *J. Biol. Chem.* 277: 30746–53
272. Oldham CE, Mohney RP, Miller SL, Hanes RN, O'Bryan JP. 2002. *Curr. Biol.* 12:1112–16
273. Katz M, Shtiegman K, Tal-Or P, Yakir L, Mosesson Y, et al. 2002. *Traffic* 3:740–51
274. Babst M, Katzmann D, Snyder W, Wendland B, Emr S. 2002. *Dev. Cell* 3:283–89
275. Bean AJ, Davanger S, Chou MF, Gerhardt B, Tsujimoto S, Chang Y. 2000. *J. Biol. Chem.* 275:15271–78
276. Cadavid AL, Ginzel A, Fischer JA. 2000. *Development* 127:1727–36
277. Chen X, Zhang B, Fischer JA. 2002. *Genes Dev.* 16:289–94