ELSEVIER

Gene 199 (1997) 39–48

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

Human syntaxin 7: a Pep12p/Vps6p homologue implicated in vesicle trafficking to lysosomes

Hui Wang, Laurence Frelin, Jonathan Pevsner *

Department of Molecular Neurobiology, Kennedy Krieger Institute and Department of Neuroscience, The Johns Hopkins University School of Medicine, 707 North Broadway, Baltimore, MD 21205, USA

Received 27 January 1997; accepted 19 May 1997; Received by J.L. Slightom

Abstract

The movement of hydrolases and other proteins to lysosomes is accomplished by vesicle trafficking. Specific vesicles are targeted from the *trans*-Golgi network via a prelysosomal compartment to lysosomes. The specificity of vesicle transport is thought to occur through the interaction of vesicle proteins with receptors on a particular target membrane. The syntaxins are a family of transmembrane proteins that have been implicated as vesicle receptors involved in vesicle docking and fusion. Syntaxins 1–4 are localized to the plasma membrane, and in particular, syntaxin 1a mediates synaptic vesicle docking in the nerve terminal. Syntaxins 5 and 6 have been localized to *cis*-Golgi and *trans*-Golgi network compartments, respectively. We now report the identification of syntaxin 7 from a human brain cDNA library. The syntaxin 7 gene is localized to human chromosome 6. By Northern analysis, the syntaxin RNA was found to be broadly distributed. Based on its homology to yeast and plant vacuolar syntaxins, we propose that syntaxin 7 has a role in vesicle trafficking between the Golgi complex and lysosomes. In vitro binding studies reveal that syntaxin 7 binds αSNAP, a key regulator of transport vesicle fusion at multiple stages of the secretory pathway. © 1997 Elsevier Science B.V.

Keywords: αSNAP; Vacuole; Intracellular transport; Golgi complex; Late endosome

1. Introduction

The lysosome is an acidic organelle that is enriched in hydrolases, and capable of degrading both internalized and endogenous macromolecules. The lysosome is the major site of catabolism in the cell. Its functional importance is highlighted by the occurrence of several dozen lysosomal storage disorders that cause neurological and other damage (Scriver et al., 1995).

Lysosomal proteins, including soluble hydrolases and transmembrane proteins, are transported to lysosomes selectively through the secretory pathway (Kornfeld and Mellman, 1989). Lysosomal hydrolases acquire phosphomannosyl residues in the Golgi complex that mediate their binding to mannose 6-phosphate receptors (Kornfeld, 1992). They are subsequently delivered to an acidified late endosomal (prelysosomal) compartment and then to lysosomes by vesicular transport. In addition to this biosynthetic pathway, macromolecules are targeted to lysosomes in vesicles from the cell surface via endosomes.

Relatively little is known of the proteins responsible for mediating vesicle trafficking to mammalian prelysosomal or lysosomal compartments. In contrast, studies of synaptic vesicle exocytosis in the presynaptic nerve terminal are relatively advanced (Südhof, 1995; Scheller, 1995). This is due in part to the abundance of synaptic vesicles, allowing a description of dozens of vesicle-associated proteins. In a model of synaptic vesicle docking and fusion (Rothman, 1994), two integral membrane

* Corresponding author, at: Kennedy Krieger Institute, 707 N. Broadway, Baltimore, MD 21205, USA. Tel.: +1 410 5508615; Fax: +1 410 5508003; e-mail: pevsner@welchlink.welch.jhu.edu

Abbreviations: aa, amino acid(s); BLAST, basic local alignment search tool; bp, base pair(s); cDNA, DNA complementary to RNA; GST, glutathione *S*-transferase; nt, nucleotide(s); NSF, *N*-ethylmaleimide sensitive factor; PAUP, phylogeny analysis using parsimony; PCR, polymerase chain reaction; SNAP, soluble *N*-ethylmaleimide sensitive factor attachment protein; SNAP-25, synaptosomal-associated protein of 25 kDa; VAM, vacuolar morphology; VAMP, vesicle-associated membrane protein; vps, vacuolar protein sorting.

PII S0378-1119(97)00343-0

proteins of the vesicle, VAMP (vesicle-associated membrane protein; also called synaptobrevin) and synaptotagmin, form a docking complex with two plasma membrane-associated proteins, syntaxin and SNAP-25 (synaptosomal-associated protein of 25 kDa). Upon dissociation of synaptotagmin, two additional proteins bind in a complex: the *N*-ethylmaleimide sensitive factor (NSF) and SNAPs (soluble NSF attachment proteins). These proteins form a complex that participates in the docking and/or fusion of synaptic vesicles with the plasma membrane. The specificity of vesicle docking and fusion is proposed to arise from the selective binding of particular combinations of vesicle proteins (v-SNAREs for vesicle SNAP receptors, such as VAMP) to acceptor membrane proteins (t-SNAREs for target membrane SNAP receptors, such as syntaxin).

Vesicle trafficking in multiple stages of the secretory pathway may be mediated by organelle- or compartment-specific v-SNAREs and t-SNAREs in mammals, in the yeast *Saccharomyces cerevisiae* and in other organisms (Bennett and Scheller, 1993). Syntaxin isoforms have been localized to the plasma membrane (syntaxins 1–4), *cis*-Golgi complex (syntaxin 5) and to the *trans*-Golgi complex(syntaxin 6) (Bennett et al., 1993; Bock et al., 1996). However, mammalian syntaxins and VAMP-like molecules have not been described in pathways of vesicle trafficking to or from lysosomes.

In *S. cerevisiae*, genes responsible for the proper targeting of vacuolar enzymes have been identified by genetic approaches rather than the biochemical and physiological approaches taken in mammalian systems. Over 40 complementation groups defining vacuolar protein sorting (*vps*) genes were identified (Bankaitis et al., 1986; Robinson et al., 1988; Rothman and Stevens, 1986). The *vps* mutants have been divided into six classes (A–E) based on the vacuolar morphology (Raymond et al., 1992). In a few cases no identifiable vacuolar structures are present, but in most mutants, vacuolar structure is normal, despite the occurrence of protein sorting defects.

One of the VPS genes encodes a syntaxin homologue that is essential for trafficking to vacuoles. Vps6p (also called Pep12p) is defined by a class D *vps* mutant required for sorting of soluble hydrolases from the Golgi complex to the vacuole (Jones, 1977; Becherer et al., 1996). αPep12, a Pep12p homologue from the plant *Arabidopsis thaliana*, was identified by functional complementation of a *pep12* mutant (Bassham et al., 1995). We now report the identification of a cDNA encoding a mammalian Pep12p homologue from human brain, syntaxin 7. The cDNA is also designated syntaxin 7. This syntaxin may mediate vesicle trafficking to a prelysosomal or lysosomal compartment. Its ability to bind αSNAP suggests that, like other trafficking pathways, vesicle trafficking to lysosomes may utilize NSF and αSNAP as well as lysosomal v-SNAREs and t-SNAREs.

2. Materials and methods

2.1. cDNA cloning and sequencing

Searches of the GenBank database (release 96.0) with the *S. cerevisiae* Pep12p and *A. thaliana* aPep12 protein sequences resulted in the identification of overlapping human expressed sequence tags (accession numbers H49700, T81994, T81999 and N42697) as well as a rat sequence (H33185). A human syntaxin 7 cDNA was cloned by synthesizing two oligonucleotides (5′ GCA TAG AAG CCA ATG TGG AAA ATG and 5′ GAC TGA TAA TCG CAA CTC CAA TGA C) as polymerase chain reaction (PCR) primers, using as a template a cDNA clone derived from human fetal liver or spleen that contained the H49700 sequence. The PCR was performed in a Minicycler (MJ Research) for 30 cycles with conditions of denaturing (94°C, 1 min), annealing (55°C, 1 min) and extension (55°C, 1 min). The resulting 148-bp product was gel-purified, radiolabeled with random hexamers, and used to probe 275 000 plaques of a human fetal brain cDNA library in λZAP (Stratagene, La Jolla, CA, USA). Positively hybridizing plaques were confirmed by secondary screening, and 11 cDNA clones, converted into bluescript SK⁻ plasmids, were partially sequenced and showed a 95% nucleotide overlap to portions of each other. The putative 5′ untranslated and coding regions of one of the longest of these clones was sequenced completely on both strands and designated syntaxin 7 (GenBank accession number U77942).

A cDNA encoding syntaxin 7 without its carboxyterminal hydrophobic domain was prepared by PCR, and subcloned into pGEX-KG. The PCR primers used were 5′ CGG AAT TCC CAT GTC TTA CAC TCC AGG AGT and 5′ ACG CGT CGA CGC ACA GGG TTT TTC TGG ATT. Subsequent transformation into *Escherichia coli* and induction were performed as previously described (Pevsner et al., 1994).

2.2. Computer analyses

DNA and protein sequences were analyzed with the Genetic Computer Group (Madison, WI, USA) software package. The basic local alignment search tool (BLAST) algorithm was used to search the GenBank database. Sequence comparisons were performed with the BESTFIT program. Statistical significance for the relatedness of two proteins was also determined with this program by generating *Z* scores. *Z* scores were obtained by measuring the quality score between two proteins, subtracting the mean quality score obtained from comparisons with 50 randomized shuffles of one protein, and dividing this value by the standard deviation of those 50 scores. *Z* scores above three are considered statistically significant. Multiple sequence

alignments were performed with the PILEUP program using default penalty values.

Coiled-coil motifs were analyzed with the COILS program, version 2.1 using the MTIDK matrix and a 21-aa residue scan (Lupas, 1996). The Phylogentic Analysis Using Parsimony (PAUP) program was used to generate phylogenetic trees (pre-release version 4d51, generously provided by Dr David Swafford, Smithsonian Institute). Confidence values for the clades were obtained by bootstrapping ($n=100$ replicates).

2.3. Northern blotting

^{32}P-labeled cDNA probe containing the full-length syntaxin 7 was used for the RNA distribution analysis. Hybridization on multiple tissue Northern blots of human and rat poly(A)$^+$ RNA (Clontech, Palo Alto, CA, USA) were performed exactly according to the manufacturer's protocol. The blot was washed at high stringency and exposed to Kodak XAR-5 film to generate an autoradiogram or to a phosphorimager (Molecular Dynamics, Sunnyvale, CA, USA) to quantitate the relative amounts of signal in different tissues.

2.4. Syntaxin 7 binding assay

GST/Syntaxin 7 and GST fusion proteins were immobilized on glutathione sepharose beads (Pharmacia, Uppsala, Sweden), and polyhistidine/αSNAP fusion protein was immobilized on TALON metal affinity resin (Clontech) prepared according to the manufacturers' protocols. The GST fusion proteins were eluted from the beads by incubation with 20 mM of glutathione for 30 min. A binding assay was performed by incubating 5 μg of immobilized αSNAP with 5 μg of syntaxin 7/GST or 2 μg GST in a total volume of 20 μl in 100 mM of PBS at 22°C for 1 h. After the beads were washed five times in PBS/1% Tween-20, proteins were eluted with sample buffer, electrophoresed on a 10% polyacrylamide gel, transferred to nitrocellulose, Ponceau S-stained, and Western-blotted with an anti-GST antiserum (Pharmacia) visualized by ECL (Amersham, Amersham, UK).

3. Results and discussion

3.1. Identification of a gene encoding a new mammalian syntaxin

BLAST analysis of GenBank DNA sequences revealed the presence of several human-derived expressed sequence tags (see Materials and Methods) predicted to encode a protein with homology to Pep12p/Vps6p from the yeast *S. cerevisiae* and aPep12 from the thale cress *A. thaliana*. We generated a 148-bp cDNA fragment from a human cDNA clone (GenBank accession number H49700) by PCR, and used the radiolabeled fragment as a probe to screen a human fetal brain cDNA library. One of the isolated cDNAs is predicted to encode a new Pep12p/Vps6p homologue that we named syntaxin 7 (Fig. 1). The cDNA is 1614 bp with a putative start Met at nt 80–82, a long open reading frame predicted to encode a protein of 261 aa and a molecular mass of 29 845 Da, and a stop codon at nt 863–865. The putative start Met is preceded by three in-frame stop codons in the 5′ untranslated region, suggesting that the cDNA encodes the amino terminus of the protein. Furthermore, the putative start Met is surrounded by a consensus sequence that is characteristic of eukaryotic ribosome binding sites used for the initiation of translation (Kozak, 1987). In particular, there is a purine at the position 3 nt upstream from the ATG codon. The stop codon at nt 863–865 is predicted to terminate the protein immediately following a hydrophic transmembrane domain (see below), consistent with the position of stop codons in all other known syntaxins. The 3′ untranslated region is 749 nt and lacks a poly(A) tail. This region also lacks consensus signals for polyadenylation such as AATAAA (Proudfoot, 1991). A BLAST database search with the most 3′ region of the human syntaxin 7 cDNA revealed >95% identity to six human expressed sequence tags (accession numbers AA197060, R73740, R82540, Z45822, T71551, and H49605), none of which contains a poly(A) tail.

The syntaxin 7 protein has a predicted isoelectric point of 5.3, consistent with the relatively low pI values of other syntaxins. It also contains a putative single transmembrane domain formed by the carboxy-terminal 20-aa residues. Based on the COILS algorithm, syntaxin 7 is predicted to contain three coiled-coil domains, also consistent with the predicted structure of other syntaxins. The protein is predicted to contain sites for phosphorylation by protein kinase C (five sites), calmodulin-dependent protein kinase II (two sites), and casein kinase II (five sites)(Pearson and Kemp, 1991) (Fig. 1).

Comparison of the human syntaxin 7 protein sequence with other proteins in the non-redundant protein database revealed a significant homology to aPep12 from *A. thaliana* (34% aa identity) and *S. cerevisiae* Pep12p/Vps6p (25% aa identity). Protein identity comparisons were also made between pairs of proteins with the BESTFIT program. These analyses generated Z scores that reflect the statistical significance of homologies between two proteins. In comparing two proteins, a quality score is generated and compared to the quality score obtained from 50 independent BESTFIT comparisons, using one of the protein sequences in a randomized order (maintaining an identical aa composition in each case). The Z scores for syntaxin 7 homology to Pep12 from *A. thaliana* ($Z=24$) and *S. cerevisiae* ($Z=13$) are highly significant (Table 1). Pep12p is required for the

```
GAGGGAGCCGTGGAGGTCCAGGTGACTGCTTAGAAAACTGCACAGCATCTGATGAAATTA   60
GCGAATAAGAACATCAACCATGTCTTACACTCCAGGAGTTGGTGGTGACCCCACCCAGTT  120
                    M  S  Y  T  P  G  V  G  G  D  P  T  Q  L    14

GGCCCAGAGGATCTCTTCTAACATCCAGAAGATCACACAGTGTTCTGTGGAAATACAAAG  180
 A  Q  R  I  S  S  N  I  Q  K  I  T  Q  C  S  V  E  I  Q  R    34

AACTCTGAATCAACTTGGAACACCTCAAGATTCACCTGAATTGAGGCAACAGTTGCAACA  240
 T  L  N  Q  L  G  T  P  Q  D  S  P  E  L  R  Q  Q  L  Q  Q    54

GAAGCAGCAGTATACTAACCAGCTTGCCAAAGAAACAGATAAGTACATTAAAGAGTTTGG  300
 K  Q  Q  Y  T  N  Q  L  A  K  E  T  D  K  Y  I  K  E  F  G    74

ATCTCTGCCCACCACCCCCAGTGAACAGCGTCAAAGGAAAATACAGAAGGATCGCTTAGT  360
 S  L  P  T  T  P  S  E  Q  R  Q  R  K  I  Q  K  D  R  L  V    94

GGCAGAGTTCACAACATCACTGACAAACTTCCAGAAGGTCCAGAGGCAGGCTGCTGAGCG  420
 A  E  F  T  T  S  L  T  N  F  Q  K  V  Q  R  Q  A  A  E  R   114

AGAGAAAGAGTTTGTTGCTCGAGTAAGAGCCAGTTCCAGAGTGTCTGGCAGTTTTCCTGA  480
 E  K  E  F  V  A  R  V  R  A  S  S  R  V  S  G  S  F  P  E   134

GGACAGCTCAAAAGAAAGGAATCTTGTATCCTGGGAAAGCCAAACTCAACCTCAAGTGCA  540
 D  S  S  K  E  R  N  L  V  S  W  E  S  Q  T  Q  P  Q  V  Q   154

GGTGCAGGATGAAGAAATTACAGAGGATGACCTCCGTCTTATTCATGAGAGAGAATCTTC  600
 V  Q  D  E  E  I  T  E  D  D  L  R  L  I  H  E  R  E  S  S   174

TATCAGGCAACTTGAAGCTGATATTATGGATATTAATGAAATATTTAAAGATTTGGGAAT  660
 I  R  Q  L  E  A  D  I  M  D  I  N  E  I  F  K  D  L  G  M   194

GATGATTCATGAACAAGGAGATGTAATAGATAGCATAGAAGCCAATGTGGAAAATGCAGA  720
 M  I  H  E  Q  G  D  V  I  D  S  I  E  A  N  V  E  N  A  E   214

GGTGCACGTTCAGCAAGCAAATCAGCAGCTGTCAAGGGCAGCAGATTATCAGCGCAAATC  780
 V  H  V  Q  Q  A  N  Q  Q  L  S  R  A  A  D  Y  Q  R  K  S   234

CAGAAAAACCCTGTGCATCATCATTCTTATCCTTGTCATTGGAGTTGCGATTATCAGTCT  840
 R  K  T  L  C  I  I  I  L  I  L  V  I  G  V  A  I  I  S  L   254

CATCATATGGGGATTGAACCACTGAAGTTATAAAGGAGCACACTGTCGCACTACATTGTC  900
 I  I  W  G  L  N  H  *                                        261

TAAATTATGTAGGAAGATTCCTGTAATCATGTTTTTTTAATTATTATTTTAAAGCTATTG  960
TATAAAGGATGGTTCCCATACTTTGTTATTTTTATTGGGGGGGTTGGGCGGGTTCCTTTG 1020
GATTAAATCTGATATTTTCTAATACTGAAAGATTTTCTAAATGTCACTGCTGACATAACT 1080
CCCTTGGTCTTCAATTTAATAGTTGTTAAGTTTTGGGCCACATTGCATATGCCTTTCATT 1140
TATAATTTATTTACCCTGCTTGACTTAGTTTGGGGAATTCGGAAATTTAAGGTGTGTGTA 1200
TTCTGTTGGGATCTCCCTGCCACGTGAACACACCAAGATGTGTGTTACTTCAAGTTAAAA 1260
CTCCCCAAAATTTAATTTTTGATTTGCTTCCACCAGGGGAAAATATTCTCCAATAATGTA 1320
AAATAATTAAGGTCCAATACATGGGTTGTATTTTTCTGGTTCACAACAGCACAAAGTGTC 1380
TTTCATTTTTTTGTTGGATTTCCTTTAAGATCTTTTTTACCCTGAAGTCGGTGAACACTT 1440
TTCTAGTTAATTTGATACTCTTTCTGTGTATATAATAAGCTTTTGCTGTAGATTGCCTAG 1500
TAAAATTACTAAGGATAGGTTGTTTTTACATATGGTCTATTTAAGTCTGATGTTTACGGG 1560
GGAGAGTGTAGTTACTAAAAATGTTTAACATAATTTGGAAGAAGAGTATGAACA       1614
```

Fig. 1. Nucleotide (nt) and predicted amino acid (aa) sequence of human syntaxin 7. The nt are numbered on the right, and aa (in standard single-letter code) are numbered in italics. In the nt sequence, four in-frame stop codons preceding the putative initiation ATG are underlined. In the protein sequence, a hydrophobic putative transmembrane domain is underlined (aa 238–259). There are three putative coiled coil domains (aa E47–E72, R166–L192, and G200–Y230). Aa that are potentially phosphorylated by protein kinase C are bold and underlined (T66, S125, S136, S174, S234). S or T residues potentially phosphorylated by calmodulin-dependent protein kinase II (S20, S174) or by casein kinase II (T41, T79, S131, S136, T161) are in bold.

Table 1
Identities between syntaxin 7, six yeast syntaxins, and aPep12 from *A. thaliana*

	Yeast Sso1	Yeast Sso2	Yeast Sed5	Yeast Vam3	Yeast hyp.	Yeast Pep12	A. thal. Pep12	Human syn 7
Sso1 (yeast)	—	74 (79)	18 (6)	21 (7)	23 (9)	21 (8)	22 (8)	19 (9)
Sso2 (yeast)		—	21 (6)	21 (8)	23 (10)	18 (9)	26 (16)	21 (8)
Sed5 (yeast)			—	23 (10)	21 (6)	23 (8)	22 (10)	24 (9)
Vam3 (yeast)				—	22 (8)	23 (10)	30 (14)	24 (13)
Hyp. (yeast)					—	25 (14)	23 (8)	20 (10)
Pep12 (yeast)						—	31 (13)	25 (13)
Pep12 (*A. thaliana*)							—	34 (24)
Syntaxin 7 (human)								—

Percentage aa identities are shown based on pairwise comparisons between each protein. Numbers in parentheses are Z scores reflecting the statistical significance of the homologies (see Materials and Methods).
Hyp., a hypothetical *S. cerevisiae* syntaxin (GenBank Accession No. Z74760). Other accession numbers are given in Fig. 3.

sorting of lumenal, but not integral membrane hydrolases to the lysosome-like vacuole (Becherer et al., 1996). Deletion or disruption of the *PEP12* gene causes the missorting of hydrolases, such as carboxypeptidase Y, and the accumulation of 40- to 50-nm transport vesicles that are trafficking from the late Golgi complex to a prevacuolar, endosomal compartment (Becherer et al., 1996). The *A. thaliana* aPep12 cDNA was isolated by functional complementation of a yeast pep12 mutant (Bassham et al., 1995). Based on its homology to Pep12p and aPep12, human syntaxin 7 may function in the transport of lysosomal enzymes from the Golgi complex to a prelysosomal (endosomal) compartment or to lysosomes.

Analysis of protein sequences in GenBank revealed that there are six syntaxin homologues in *S. cerevisiae*, including two that were not previously identified (Vam3p (Wada et al., 1992) and a hypothetical protein). Comparison of all six yeast syntaxins with *A. thaliana* aPep12 and human syntaxin 7 suggests that syntaxin 7 is a mammalian Pep12 homologue (Tables 1 and 2). Alignment of syntaxin 7 with Pep12 from *A. thaliana* and *S. cerevisiae* reveals that the three proteins share regions of identity throughout their lengths, but especially in the region preceding the putative carboxy terminal transmembrane domains (Fig. 2).

We also compared syntaxin 7 to mammalian syntaxins 1–6 (Table 2). Syntaxin 7 has the highest aa identity to syntaxin 3 (27% identity, $Z=11$). In contrast to syntaxin 6 (Bock et al., 1996), syntaxin 7 does not share any significant homology to SNAP-25 (20% aa identity, $Z=2.8$).

A BLAST search of the database of sequence tagged sites revealed that syntaxin 7 is 97% identical at the nt level to a cDNA clone (GenBank accession number G15368) mapped to human chromosome 6. Hybridization of radiolabeled syntaxin 7 cDNA to a monochromosomal panel of human chromosomal DNA by Southern blotting revealed a pattern consistent with the assignment of the syntaxin 7 gene to chromosome 6 (data not shown). Computer database searches of other human syntaxins revealed that the syntaxin 1a gene is localized to chromosome 7 (GenBank accession number G15465), whereas the syntaxin 3 gene maps to chromosome 11 (GenBank accession number G26641). The syntaxin 1b gene has been mapped to chromosome 16p11.2 by fluorescence in situ hybridization (Smirnova et al., 1996). The genes encoding these four human syntaxins are thus not clustered on a single chromosome.

3.2. Phylogenetic analyses

We constructed an inferred phylogenetic tree (cladogram) to analyze the evolutionary relationships between syntaxin 7 and 31 other syntaxins (Fig. 3). The 32 syntaxins represent operational taxonomic units that

Table 2
Identities between syntaxin 7 and other mammalian syntaxins

	Syn 1	Syn 2	Syn 3	Syn 4	Syn 5	Syn 6	Syn 7
Syntaxin 1 (human)	—	64 (51)	63 (51)	48 (40)	23 (9)	21 (4)	26 (11)
Syntaxin 2 (rat)		—	63 (51)	45 (38)	21 (8)	21 (4)	26 (12)
Syntaxin 3 (human)			—	41 (29)	22 (10)	24 (7)	27 (11)
Syntaxin 4 (human)				—	23 (6)	20 (4)	26 (13)
Syntaxin 5 (human)					—	19 (7)	23 (7)
Syntaxin 6 (rat)						—	17 (4)
Syntaxin 7 (human)							—

Percentage aa identities are shown based on pairwise comparisons between human or rat syntaxins 1–7. Numbers in parentheses are Z scores reflecting the statistical significance of the homologies (see Materials and Methods). Accession numbers are given in Fig. 3.

```
syntaxin 7     .......... .....MSYTP GVGGDP...T QLAQRISSNI QKTTQCSVEI  32
Pep12 (plant) MSFQDLEAGT RSPAPNRFTG GRQQRPSSRG DPSQEVAAGI FRTSTAVNSF
Pep12 (yeast) MSEDEFFGGD NEGV...WNG SRFSDSPEFQ TLKEEVAAEL FETNGQISTL

syntaxin 7     QRTLNQLGTP QDSPELRQQL QQ........ KQQYTNQLAK ETDKYIIEFG  74
Pep12 (plant) FRLVNSIGTP KDTLELRDKL QK........ TRLQISELVK NTSAKLLEAS
Pep12 (yeast) QQFTATLKSF IDRGDVSAKV VERINKRSVA KIEEIGGLIK KINTSVTKMD

syntaxin 7     SLPTTPSEQR QRKIQKDRLV AEFTTSLTNF QKVQRQAAER EKEFVARVRA 124
Pep12 (plant) EADLHGSASQ IKKIADAKLA KDFQSVLKEF QKAQRLAAER EITYTPVVTK
Pep12 (yeast) AIE.EASLDK TQIIAREKLV RDVSYSFQEF QGIQRQFTQV MKQVNERAKE

syntaxin 7     SSRVSGSFPE ......DSSK ERNLVSWESQ TQPQVQVQDE ETTEDDL... 165
Pep12 (plant) EIPTSYNAPE LDTESLRISQ QQALL.LQSR RQ.EVVFLDN ETTFNEA...
Pep12 (yeast) SLEASEMAND ..AALLDEEQ RQNSSKSTRI PGSQIVIERD PINNEEFAYQ

syntaxin 7     .RLIHERESS IRQLEADIMD INEIFKDLGM MIHEQGDVID SIEANVENAE 214
Pep12 (plant) ..IIEEREQG IREIEDQIRD VNGMFKDLAL MVNHQGNIVD DISSKLDNSH
Pep12 (yeast) QNLIEQRDQE ISNIERGITE LNEVFKDLGS VVQQQGVLVD NIEAKIYTTS

                                    transmembrane domain
syntaxin 7     VHVQQANQQL SRAADYQRKS RKTLCIIILI LVIGVAIISL IIWGLNH    261
Pep12 (plant) AATTQATVQL RKAAKTQRSN SSLTCLLILI FGIVLLIVII VVLV
Pep12 (yeast) DNTQLASDEL RKAMRYQKRT SRWRVYLLIV LLVMLLFIFL IMKL
```

Fig. 2. Aa alignment of syntaxin 7 and Pep12 from the yeast *S. cerevisiae* and from the thale cress *A. thaliana*. The single letter aa code is shown, and residues for syntaxin 7 are numbered on the right. Amino acid residues that are conserved in all three proteins are shaded.

were identified in computer searches of public DNA and protein databases; in addition to syntaxin 7, and the yeast syntaxins, we identified a hypothetical *C. elegans* protein as a novel syntaxin (GenBank accession number Z69661). To construct the tree, the 32 syntaxin sequences were multiply aligned, and character positions containing any gaps were eliminated. Parsimony analysis was used to construct a tree that required the minimal number of evolutionary changes to account for the differences among the 32 syntaxins at each aa position. The tree is unrooted because no ancestral syntaxin is known to define an outgroup. In an unrooted tree such as this, there is no root node, and branch lengths specify relationships among the syntaxins without defining a primordial evolutionary path.

To gain a statistical measure of confidence in the tree, we performed a bootstrap analysis. A total of 100 trees were generated from the initial data set, and the percentage of trees containing a particular clade was measured. (A clade is a group of syntaxins that contains a common ancestor that is not shared by any syntaxin outside the group.) Bootstrap values $>70\%$ are associated with statistical significance at the $P<0.05$ level (Hillis and Bull, 1993).

The cladogram indicates that syntaxin 7 is part of a clade consisting of five syntaxins. Three (Pep12p/Vps6p, aPep12 and Vam3p) have a presumed vacuolar or endosomal localization (Jones, 1977; Wada et al., 1992; Bassham et al., 1995; Becherer et al., 1996). The intracellular localization of the hypothetical yeast syntaxin (accession number Z74760) is not known, and it has not been functionally characterized. The phylogenetic analysis suggests that, in addition to syntaxin 7, a larger group of mammalian syntaxins may exist within this clade that are involved in intracellular trafficking to endosomes, lysosomes, or other intracellular organelles.

3.3. Tissue distribution in human and rat

To determine the tissue distribution of the syntaxin 7 messenger RNA, a human syntaxin 7 cDNA clone was radiolabeled and used to probe a human RNA blot (Northern blot). Major messenger RNA species of 3.6 kb and 1.8 kb were detected in all tissues tested (Fig. 4A). The highest mRNA levels of the 3.6-kb transcript were in placenta (with a relative signal based on measurement of pixel units by phosphorimaging of 100%), with intermediate levels in heart (82% relative to placenta), skeletal muscle (75%), kidney (72%) and brain (71%), and lowest levels in pancreas (33%), lung (24%) and liver (13%). As a positive control for the integrity of the RNA, the same blot was probed with 100 ng of radiolabeled β-actin (Fig. 4B). The broad tissue distribution of syntaxin 7 is consistent with a role for the protein in a constitutive vesicle trafficking process. This is in contrast to syntaxin 1a, which is selectively enriched in brain (Bennett et al., 1992). The presence of two transcripts in each lane of the Northern blot could be accounted for by differential polyadenylation, or by the detection of two genes by the probe. The smaller transcript (1.8 kb) is approximately the same size as the syntaxin cDNA clone that was sequenced.

The human syntaxin 7 cDNA was also used to probe a rat RNA blot with comparable results (data not

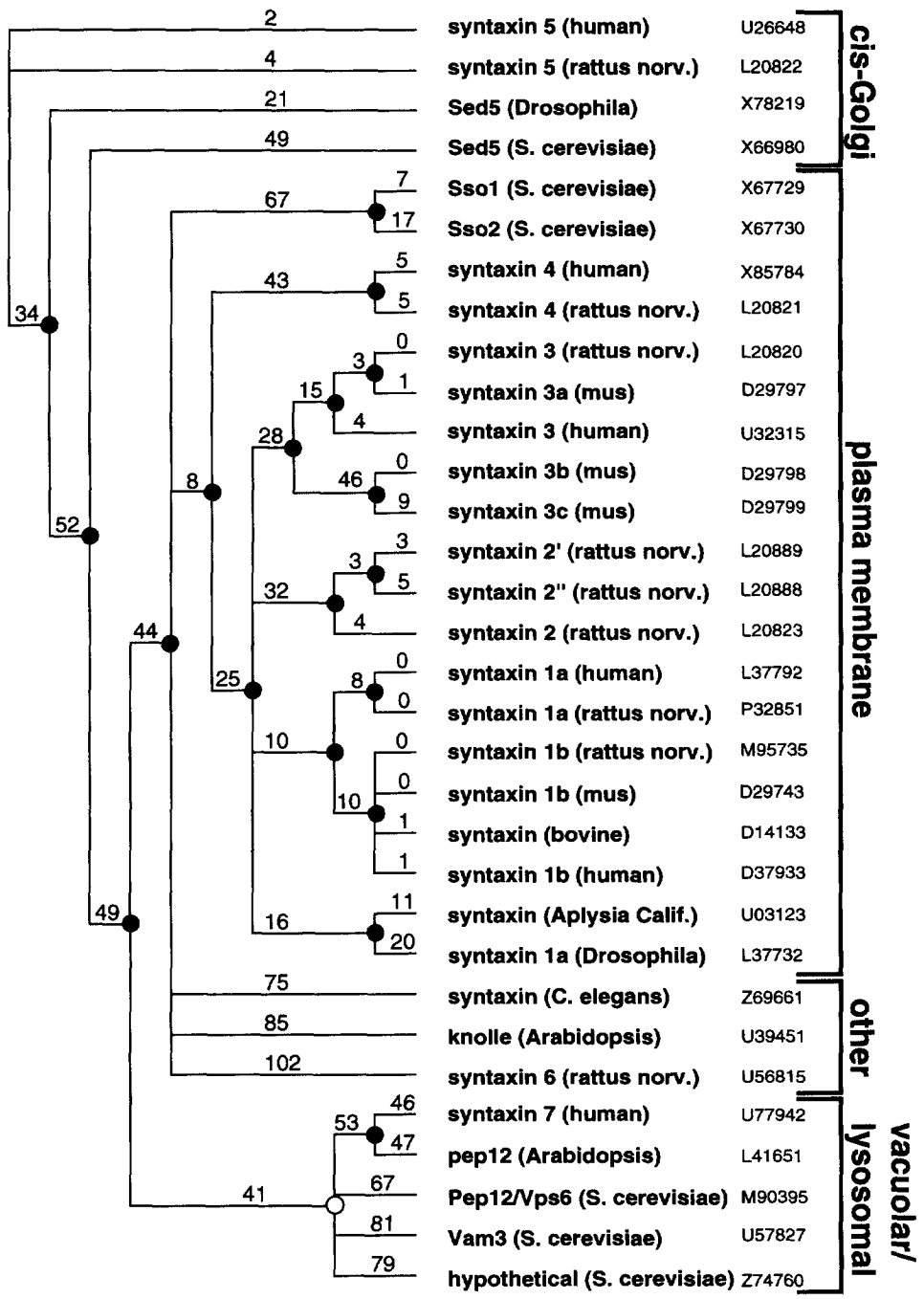

Fig. 3. Phylogenetic analysis of the syntaxin family of proteins. Thirty-two full-length syntaxins were identified in the GenBank DNA database (release 96.0, 10/96) that encoded full-length syntaxins. The tree represents an unrooted cladogram. Branch length values are indicated, and are additive. Bootstrap values ≥70% are indicated by closed circles; a value of 53% is indicated by an open circle. These percentages derive from sampling 100 trees to obtain confidence values for the groupings of particular clades. The names of the proteins are indicated with the species and the DNA accession numbers. Additional sequences that do not encode full-length proteins are not shown. The full names of the abbreviated species names given in the figure are *Drosophila melanogaster*, *Saccharomyces cerevisiae*, *Mus musculis*, *Rattus norvegicus*, *Aplysia californica* and *Caenorhabditis elegans*.

shown). A major messenger RNA species of 2.4 kb was detected in all tissues tested, with a minor band at 1.5 kb. Highest levels of expression were in kidney, with intermediate levels in lung, spleen, heart and brain, and lower levels in skeletal muscle, liver and testis.

3.4. Syntaxin 7 binds αSNAP

A cDNA clone encoding rat αSNAP with a polyhistidine tag was expressed in *E. coli* and immobilized on metal agarose beads. This αSNAP was electrophoresed

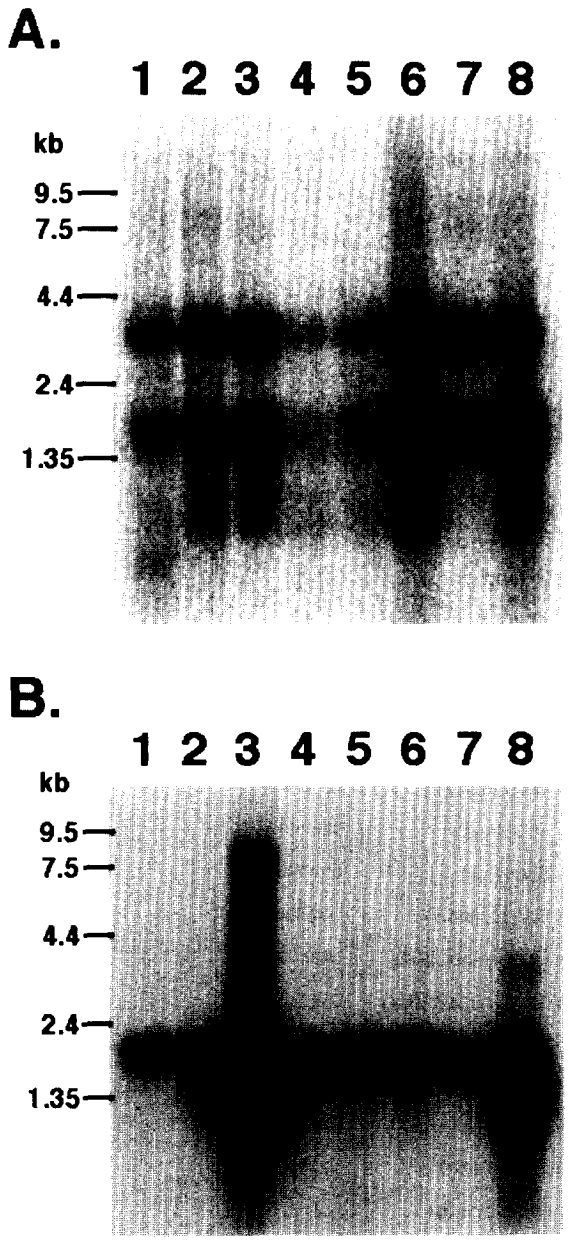

Fig. 4. RNA distribution (Northern) analysis of syntaxin 7. Each lane contained 2 μg of human poly(A)$^+$ RNA. Molecular size markers are indicated in kb. Blots were probed with 50 ng of random primer labeled syntaxin 7 cDNA (panel A) or β-actin (panel B). Lanes: 1, pancreas; 2, kidney; 3, skeletal muscle; 4, liver; 5, lung; 6, placenta; 7, brain; 8, heart. The radioactive signal was quantitated by phosphorimaging (Molecular Dynamics).

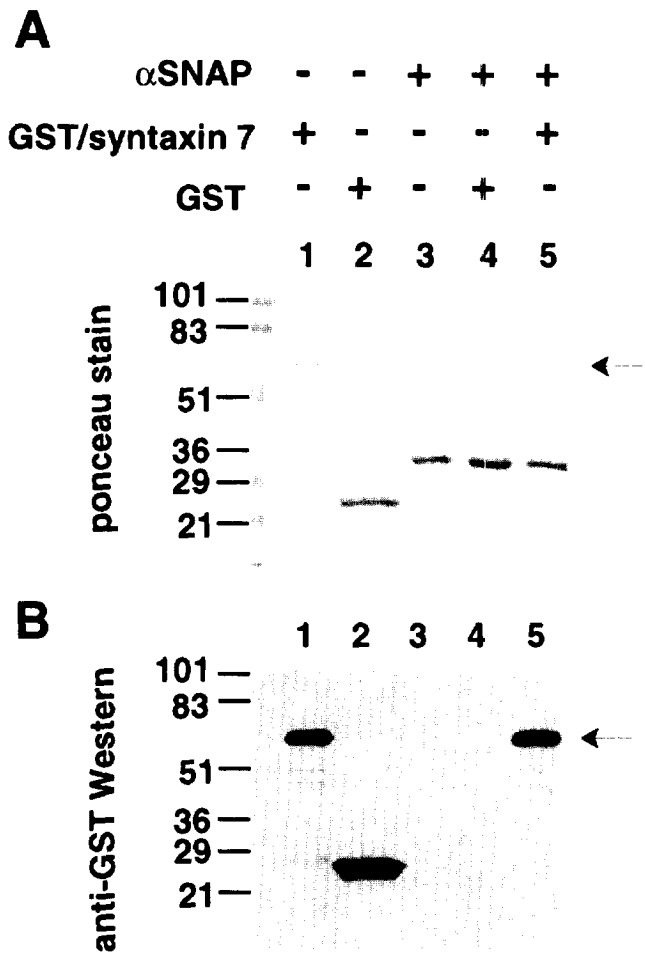

Fig. 5. Syntaxin 7 binds to αSNAP. Recombinant GST/syntaxin 7 (lanes 1 and 5) or GST (lanes 2 and 4) were incubated to equilibrium with recombinant αSNAP immobilized on metal agarose beads. Beads were washed, and bound material was detected by Ponceau S staining (panel A) and by an anti-GST Western blot (panel B) (lanes 3–5). In the presence of immobilized αSNAP, GST/syntaxin 7 binding was observed (lane 5, panels A and B, arrows). The binding assay was repeated four times with different batches of αSNAP, GST/syntaxin 7, and GST fusion proteins. Panels A and B are from the same experiment.

on a 10% polyacrylamide gel, transferred to nitrocellulose, and visualized by Ponceau S staining (Fig. 5A, lanes 3–5). Recombinant GST/syntaxin 7 was similarly visualized by Ponceau S staining (Fig. 5A, lane 1) and by Western blotting with an anti-GST polyclonal antiserum (Fig. 5B, lane 1). GST/syntaxin 7 bound specifically to αSNAP, as determined by Ponceau S staining (Fig. 5A, lane 5) and by Western blotting (Fig. 5B, lane 5). GST did not bind to αSNAP (Fig. 5A and B, lane 4), although the anti-GST did detect GST (Fig. 5A, lane 2). These controls confirm the specificity of αSNAP binding to syntaxin 7.

αSNAP is a functional homologue of Sec17p from *S. cerevisiae* (Griff et al., 1992). SEC17 is required in vivo for vesicle trafficking from endoplasmic reticulum to the Golgi complex, and from the Golgi complex to the plasma membrane. αSNAP may function in vesicle trafficking by binding to SNAREs (such as a complex of syntaxin 1a, SNAP-25, and VAMP in the nerve terminal) to recruit the binding of NSF (Söllner et al., 1993). In particular, αSNAP binds directly to syntaxin 1a (Hanson et al., 1995; Kee et al., 1995). ATP hydrolysis by NSF then may trigger membrane fusion, or possibly NSF participates in an earlier priming step

(Chamberlain et al., 1995). A role for αSNAP in trafficking to lysosomes has not been demonstrated, but the binding of αSNAP by syntaxin 7 is consistent with this hypothesis. Furthermore, Sec17p is required for the homotypic fusion of yeast vacuoles (Haas and Wickner, 1996; Mayer et al., 1996). Our results suggest that syntaxin 7 may serve as a binding partner for proteins related to VAMP, SNAP-25, sec1 proteins, or other proteins that mediate vesicle trafficking to prelysosomal compartment or lysosomes.

Syntaxins in general bind to soluble 65-kDa to 80-kDa proteins, known collectively as the sec1 family (reviewed in Pevsner (1996)). These interactions include the interaction of syntaxins 1–3 with n-sec1, a neural-enriched mammalian Sec1p homologue that is also referred to as Munc-18, RbSec1, mSec1 and p67 (Hata et al., 1993; see Pevsner, 1996). Two sec1-related proteins in yeast, Vps33p/Slp1p and Vps45p, have mammalian homologues r-vps33a, r-vps33b and h-vps45 (Pevsner et al., 1996). We investigated whether these three mammalian proteins bind to human syntaxin 7. A cDNA encoding syntaxin 7 was in vitro transcribed and translated in the presence of [^{35}S]methionine, then incubated with recombinant r-vps33a, r-vps33b, or h-vps45 immobilized on glutathione sepharose beads as GST fusion proteins. No specific binding of r-vps33a, r-vps33b, or h-vps45 was detected (data not shown).

Acknowledgement

We thank Dr. David Swafford of the Smithsonian Institute for providing access to PAUP, Dr. Richard Scheller of Stanford University for providing an αSNAP cDNA clone and N. Varg for manuscript preparation.

References

Bankaitis, V.A., Johnson, L.M., Emr, S.D., 1986. Isolation of yeast mutants defective in protein targeting to the vacuole. Proc. Natl. Acad. Sci. USA 83, 9075–9079.

Bassham, D.C., Gal, S., da Silva Conceicao, A., Raikhel, N.V., 1995. Proc. Natl. Acad. Sci. USA 92, 7262–7266.

Becherer, K.A., Rieder, S.E., Emr, S.D., Jones, E.W., 1996. Novel syntaxin homologue, Pep12p, required for the sorting of lumenal hydrolases to the lysosome-like vacuole in yeast. Mol. Biol. Cell 7, 579–594.

Bennett, M.K., Calakos, N., Scheller, R.H., 1992. Syntaxin: a synaptic protein implicated in docking of synaptic vesicles at presynaptic active zones. Science 257, 255–259.

Bennett, M.K., Garcia-Arraras, J.E., Elferink, L.A., Peterson, K., Fleming, A.M., Hazuka, C.D., Scheller, R.H., 1993. The syntaxin family of vesicular transport receptors. Cell 74, 863–873.

Bennett, M.K., Scheller, R.H., 1993. The molecular machinery for secretion is conserved from yeast to neurons. Proc. Natl. Acad. Sci. USA 90, 2559–2563.

Bock, J.B., Lin, R.C., Scheller, R.H., 1996. A new syntaxin family member implicated in targeting of intracellular transport vesicles. J. Biol. Chem. 271, 17961–17965.

Chamberlain, L.H., Roth, D., Morgan, A., Burgoyne, R.D., 1995. Distinct effects of αSNAP, 14–3–3 proteins, and calmodulin on priming and triggering of regulated exocytosis. J. Cell Biol. 130, 1063–1070.

Griff, I.C., Schekman, R., Rothman, J.E., Kaiser, C.A., 1992. The yeast *SEC17* gene product is functionally equivalent to mammalian α-SNAP protein. J. Biol. Chem. 267, 12106–12115.

Haas, A., Wickner, W., 1996. Homotypic vacuole fusion requires Sec17p (yeast αSNAP) and Sec18p (yeast NSF). EMBO J. 15, 3296–3305.

Hanson, P.I., Otto, H., Barton, N., Jahn, R., 1995. The *N*-ethlymaleimide-sensitive fusion protein and αSNAP induce a conformational change in syntaxin. J. Biol. Chem. 270, 16955–16961.

Hata, Y., Slaughter, C.A., Südhof, T.C., 1993. Synaptic vesicle fusion complex contains unc-18 homologue bound to syntaxin. Nature 366, 347–351.

Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42, 182–192.

Jones, E.W., 1977. Proteinase mutants of *Saccharomyces cerevisiae*. Genetics 85, 23–33.

Kee, Y., Lin, R.C., Hsu, S.-C., Scheller, R.H., 1995. Distinct domains of syntaxin are required for synaptic vesicle fusion complex formation and dissociation. Neuron 14, 991–998.

Kornfeld, S., 1992. Structure and function of mannose 6-phosphate/insulinlike growth factor II receptors. Annu. Rev. Biochem. 61, 307–330.

Kornfeld, S., Mellman, I., 1989. The biogenesis of lysosomes. Annu. Rev. Cell Biol. 5, 483–525.

Kozak, M., 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. 15, 8125–8148.

Lupas, A., 1996. Prediction and analysis of coiled-coil structures. Methods Enzymol. 266, 513–525.

Mayer, A., Wickner, W., Haas, A., 1996. Sec18p (NSF)-driven release of Sec17p (αSNAP) can precede docking and fusion of yeast vacuoles. Cell 85, 83–94.

Pearson, R.B., Kemp, B.E., 1991. Protein kinase phosphorylation site sequences and consensus specificity motifs: tabulations. Meths Enzymol. 200, 62–79.

Pevsner, J., Hsu, S.-C., Scheller, R.H., 1994. n-Sec1: a neural-specific syntaxin-binding protein. Proc. Natl. Acad. Sci. USA 91, 1445–1449.

Pevsner, J., 1996. The role of Sec1p-related proteins in vesicle trafficking in the nerve terminal. J. Neurosci. Res. 45, 89–95.

Pevsner, J., Hsu, S.-C., Hyde, P.S., Scheller, R.H., 1996. Mammalian homologues of yeast vacuolar protein sorting (*vps*) genes implicated in Golgi-to-lysosome trafficking. Gene 183, 7–14.

Proudfoot, N., 1991. Poly(A) signals. Cell 64, 671–674.

Raymond, C.K., Howald-Stevenson, I., Vater, C.A., Stevens, T.H., 1992. Morphological classification of the yeast vacuolar protein sorting mutants: evidence fro a prevacuolar compartment in class E *vps* mutants. Mol. Biol. Cell 3, 1389–1402.

Robinson, J.S., Klionsky, D.J., Banta, L.M., Emr, S.D., 1988. Protein sorting in *Saccharomyces cerevisiae*: isolation of mutants defective in the delivery and processing of multiple vacuolar hydrolases. Mol. Cell. Biol. 8, 4936–4948.

Rothman, J.E., 1994. Mechanisms of intracellular protein transport. Nature 372, 55–63.

Rothman, J.H., Stevens, T.H., 1986. Protein sorting in yeast: mutants defective in vacuole biogenesis mislocalize vacuolar proteins into the late secretory pathway. Cell 47, 1041–1051.

Scheller, R.H., 1995. Membrane trafficking in the presynaptic nerve terminal. Neuron 14, 893–897.

Scriver, C.R., Beaudet, A.L., Sly, W.S., Valle, D., 1995. The Metabolic and Molecular Bases of Inherited Disease. McGraw-Hill, New York, NY.

Smirnova, T., Miniou, P., Viegas-Pequignot, E., Mallet, J., 1996. Assignment of the human syntaxin 1B gene (STX) to chromosome 16p11.2 by fluorescence in situ hybridization. Genomics 36, 551–553.

Söllner, T., Whiteheart, S.W., Brunner, M., Erdjument-Bromage, H., Geromanos, S., Tempst, P., Rothman, J.E., 1993. SNAP receptors implicated in vesicle targeting and fusion. Nature 362, 318–324.

Südhof, T.C., 1995. The synaptic vesicle cycle: a cascade of protein–protein interactions. Nature 375, 645–653.

Wada, Y., Ohsumi, Y., Anraku, Y., 1992. Genes for directing vacuolar morphogenesis in *Saccharomyces cerevisiae*. J. Biol. Chem. 267, 18665–18670.