The Sorting Sequence of the Peroxisomal Integral Membrane Protein PMP47 Is Contained within a Short Hydrophilic Loop

John M. Dyer, James A. McNew, and Joel M. Goodman

Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, Texas 75235-9041

Abstract. No targeting sequence for peroxisomal integral membrane proteins has yet been identified. We have previously shown that a region of 67 amino acids is necessary to target Pmp47, a protein that spans the membrane six times, to peroxisomes. This region comprises two membrane spans and the intervening loop. We now demonstrate that the 20 amino acid loop, which is predicted to face the matrix, is both necessary and sufficient for peroxisomal targeting. Sufficiency was demonstrated with both chloramphenicol acetyltransferase and green fluorescent protein as carriers. There is a cluster of basic amino acids in the middle of the loop that we predict protrudes from the membrane surface into the matrix by a flanking stem structure. We show that the targeting signal is composed of this basic cluster and a block of amino acids immediately downstream from it.

Most organellar proteins are synthesized in the cytoplasm and are directed to their destination by *cis*-acting sequences. These sequences are recognized either by cytoplasmic factors (such as the signal recognition particle) or by receptors upon arrival at the target membrane. The translocation apparatus is then engaged. The protein will either assemble into the membrane or will be fully translocated into the internal compartment. This bifurcation of import pathways is usually not a function of targeting signals but is rather dictated by the presence of hydrophobic domains that serve as stop-transfer sequences (Blobel et al., 1979; Yost et al., 1983).

The details of these steps that occur at the surface of the ER (Gorlich and Rapoport, 1993; Jungnickel et al., 1994), mitochondria (Hannavy et al., 1993; Schwarz and Neupert, 1994), and inner membrane of *Escherichia coli* (Bassford et al., 1991; Economu and Wickner, 1994) are becoming well understood due to intensive research. The corresponding events at the peroxisomal membrane are less clear and may be different in significant ways.

Peroxisomes comprise a matrix compartment surrounded by a single membrane. Peroxisomal matrix proteins are synthesized on free polysomes and are posttranslationally imported into the organelle (Lazarow and Fujiki, 1985). About half of peroxisomal matrix proteins contain the carboxy-terminal tripeptide sequence Ser-Lys-Leu, or variations of this sequence, which was the first peroxisomal targeting sequence, termed PTS1, to be defined (Gould et al., 1989). A second peroxisomal targeting sequence (PTS2) has been found on the NH_2 terminus of 3-ketoacyl-CoA thiolase and three other proteins (Swinkels et al., 1991). This sequence is 16–26 amino acids in length and may or may not be cleaved upon import. Perlp from *Hansenula polymorpha* is unique in that it has both a PTS1 and a PTS2 sequence (Waterham et al., 1994). Other peroxisomal proteins do not contain either sequence. Recent evidence is accumulating that demonstrates that both PTS1- and PTS2-containing proteins can be imported as oligomeric complexes (Glover et al., 1994; McNew and Goodman, 1994). Furthermore, the import of highly cross-linked substrates and even 9-nm gold particles into peroxisomes suggests that the oligomeric proteins do not have to be disassembled and unfolded on the surface before import (Walton et al., 1995).

Peroxisomal membrane proteins (Pmp)[1] are also synthesized on free polysomes (with one possible exception) (Bodnar and Rachubinski, 1991) and are posttranslationally targeted to the organelle (Lazarow and Fujiki, 1985). However, much less is known about their targeting and assembly. In vitro studies indicate that membrane insertion is temperature dependent but does not require ATP hydrolysis (Diestelkötter and Just, 1993). With one exception (Liu et al., 1995), integral membrane proteins do not contain a recognizable PTS1 or PTS2 sequence. Previous

Address all correspondence to Joel M. Goodman, Department of Pharmacology, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75235-9041. Tel.: (214) 648-2359. Fax: (214) 648-2994. e-mail: goodma02@utsw.swmed.edu.

J.A. McNew's present address is Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, 1275 York Ave., New York, NY 10021.

1. *Abbreviations used in this paper*: CAT, chloramphenicol acetyltransferase; GFP, green fluorescent protein; HA, hemagglutinin; Pmp, peroxisomal membrane protein.

studies using an integral Pmp from *Candida boidinii*, Pmp47, demonstrated that the targeting signal was contained within an internal region of the protein sequence (McCammon et al., 1994). Pmp47 has homology to members of the mitochondrial family of solute transporters, which span the membrane six times (Kuan and Saier, 1993). The important targeting information on this protein was found within the region consisting of spans 4 and 5 and the intervening loop. Here we show that this loop sequence of Pmp47, which is predicted to face the matrix based on protease mapping (McCammon et al., 1994), is both necessary and sufficient for sorting to peroxisomes. This is the first report of a sorting sequence, termed mPTS, for a peroxisomal integral membrane protein.

Materials and Methods

Strains and Culturing Conditions

Standard recombinant techniques (Sambrook et al., 1989) were performed using *E. coli* strains XL1-Blue (*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F'proAB lacI^{q}ZDM15*, Tn*10 (tet^r)]*) and TG-1 (F' *traD36 laclq Δ[lacZ]M15 proA+B+/supE Δ[hsdM-mcrB]5* [$r_k^-m_k^-$McrB$^-$] *thi Δ(lac-proAB)*). *E. coli* strain CJ236 (F' *cat[*=pCJ105; M13^s Cm^r]/*dut ung1 thi-1 rel A1 spoT1 mcrA*) was used for site-directed mutagenesis (Kunkel et al., 1987). Yeast expression experiments were performed using *Saccharomyces cerevisiae* strain MMYO11 (*MATα ade 2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 can1-100* Ole$^+$) (McCammon et al., 1994) or *pas10-1 (MATα leu2 his4 trp1::URA3)* (van der Leij et al., 1992). Yeast transformants were maintained on synthetic dextrose plates (2% glucose and 0.67% yeast nitrogen base without amino acids [Difco Laboratories, Detroit, MI]) supplemented with appropriate amino acids and bases. Liquid culturing for the induction of peroxisomes and subsequent expression of galactose-inducible vectors were performed in MMYO11 essentially as described (McNew and Goodman, 1994). Briefly, yeast were inoculated into 250 ml of SGd medium (3% glycerol, 0.1% glucose, 0.67% yeast nitrogen base, and appropriate auxotrophic substrates) and grown for 2 d at 30°C in an air shaker at 300 rpm. 250 OD_{600} U were transferred to a new flask and "boosted" with 1 × YP (1% yeast extract, 2% peptone) for 4 h. The cells were harvested by centrifugation at 7,000 *g* for 10 min and resuspended in 250 ml of semisynthetic oleate medium (0.05% yeast extract and 0.1% oleic acid) (McCammon et al., 1990) containing appropriate amino acids and bases. Cells were grown for another 15 h, and galactose was added to a final concentration of 0.1%. The cells were harvested 24 h later by centrifugation at 7,000 *g* for 10 min. For peroxisomal induction of cells carrying oleate-inducible plasmids, the same protocol was followed as described above, with the following changes: 500 ml cultures were typically used, cells were harvested 20 h after growth in oleic acid–containing medium, and no galactose was added.

Expression of oleate-inducible plasmids in *pas10-1* was performed by culturing the cells in 1 liter SGd to an OD_{600} of 0.1 to 0.6 (as indicated in the legend of Fig. 4). Oleic acid was then added to a final concentration of 0.1%, and the cells were cultured for an additional 18–22 h. Cells were harvested and organelles were isolated as described below.

Preparation of Spheroplasts and Organelle Fractionation

Cells from either the galactose or oleic acid induction protocol were processed by the same fractionation procedure. The harvested cells were washed with 20 ml of distilled water and spun at 7,000 *g* for 10 min. The cell pellet was washed once more with 20 ml of water, resuspended in 20 ml of 0.1 M Tris·SO_4, pH 9.2, 10 mM DTT, and incubated for 15 min at 30°C, at 300 rpm. The cells were pelleted at 7,000 *g* for 10 min, washed once in 20 ml of 1 M sorbitol, and brought up in 20 ml of 1 M sorbitol, 20 mM $KHPO_4$, pH 7.5. Zymolyase 100T (ICN Biomedicals, Irvine, CA) was added at 0.4 μg/OD_{600} U of cells. Incubations were performed for ~1 h at 30°C, at 300 rpm. Spheroplasting was microscopically verified by osmotic lysis. Spheroplasts were collected by centrifugation at 7,000 *g* for 10 min at 4°C and were resuspended in 2 ml of 1 M SMA (1 M sorbitol, 5 mM 2-[*N*-morpholinoethanesulfonate], pH 5.5, and 0.2 mM [4-{2-aminoethyl} benzenesulfonylfluoride HCl]). Spheroplasts were osmotically lysed by adding 6 ml of 0.25 M SMA (equivalent to 1 M SMA except the concentration of sorbitol is 0.25 M) and then reequilibrated with 6 ml of 1.75 M SMA.

Unbroken spheroplasts, nuclei, and cell debris were cleared by centrifugation at 1,000 *g* for 5 min at 4°C. The postnuclear supernatant was transferred to a new tube, and the remaining cell pellet was reextracted with 2 ml of 1 M SMA. After spinning, the supernatants were combined and spun once more at low speed to ensure removal of particulates. The supernatant was then spun at 25,000 *g* for 25 min at 4°C to obtain a pellet consisting of mainly peroxisomes and mitochondria. The supernatant of this spin was saved, and the pellet was redissolved in 200 μl of 1 M SMA. For SDS-PAGE analysis, 2% of the total volume of supernatant and pellet fractions was precipitated in 10% TCA, and 0.4% of each was analyzed by SDS-PAGE.

Peroxisomes and mitochondria were resolved on a 15–45% continuous Nycodenz gradient (McNew and Goodman, 1994). 200 μl of the 25,000 *g* crude organellar fraction were loaded onto the gradient, and the samples were spun at 100,000 *g* for 55 min at 3°C in a vertical rotor (VTi65; Beckman Instruments, Inc., Palo Alto, CA). The gradients were fractionated from the bottom into 12 samples. 50 μl of each sample was dissolved in 20 μl of 4× sample buffer, and 30 μl was analyzed by SDS-PAGE.

For the experiment shown in Fig. 1, whole cell lysates were prepared and assayed for protein as previously described (Veenhuis and Goodman, 1990).

Extraction of Nycodenz Gradient Fractions and Flotation of Membranes

50 μl of each Nycodenz gradient fraction (prepared from cells expressing chloramphenicol acetyltransferase [CAT]-loop-hemagglutinin [HA]) were diluted 1:5 with either 1 M SMA, 1 M SMA containing 1 M NaCl, or 200 mM Na_2CO_3. The samples were vortexed for 1 min at 4°C, kept on ice for 1 h, vortexed again for 1 min at 4°C, and then spun at 100,000 *g* for 15 min at 4°C. For the flotation experiment, a crude organellar pellet was resuspended in 60% wt/wt sucrose containing 30 mM Tris, pH 8.9. Layers of 40 and 15% sucrose were added, and the samples were centrifuged for 18 h at 55,000 rpm in a rotor (SW60; Beckman Instruments, Inc.). Peroxisomal membranes were harvested in and above the 15–40% sucrose interface.

SDS-PAGE and Immunoblotting

All polyacrylamide gel electrophoreses were performed by the Laemmli system (Laemmli, 1970), with the stacking gel at pH 6.8 and resolving gel at pH 9.2. Samples were generally analyzed using 4% stacking and 9% resolving gels. Prestained and unstained molecular weight markers were purchased from Bio-Rad Laboratories (Hercules, CA).

Proteins were transferred to nitrocellulose for immunoblotting (Towbin et al., 1979). Antibodies used in these studies included the anti-CAT polyclonal antibody (1:1,000) (5 Prime-3; Prime Inc., Denver, CO), the anti-thiolase polyclonal antibody (1:40,000) (kind gift of Jon Rothblatt, Dartmouth Medical School, Hanover, NH), the anti-Pmp27 polyclonal antibody (1:1,000) (Marshall et al., 1995), the anti-acyl CoA oxidase polyclonal antibody (1:500) (McNew et al., 1993), the anti-HA mAb 12CA5 (1:1,000; 1:200) (Boehringer Mannheim Biochemicals, Indianapolis, IN), and the anti-myc mAb 9E10 (kind gift of Richard Anderson, University of Texas Southwestern, Dallas, TX).

Vector Constructions: General Methods and Reagents

Restriction enzymes and other DNA modifying enzymes were purchased from New England Biolabs (Beverly, MA) or Boehringer Mannheim Biochemicals and were used as recommended by the suppliers. All mutagenesis reactions were performed using site-directed mutagenesis techniques (Kunkel et al., 1987) or annealed synthetic oligonucleotides (synthesized by Molecular Cardiology, University of Texas Southwestern, Dallas, TX). All DNA manipulations resulting in changes in coding sequence were verified using di-deoxy sequencing methods and the Sequenase Reaction Kit (United States Biochemical Corp., Cleveland, OH).

Construction of COOH-terminal Deletions of Pmp47 Fused to the HA Epitope Tag

An EcoRI/XbaI fragment, encoding amino acids 1–267 of Pmp47 plus a short region of 5′ untranslated sequence, served as the starting point for the construction of progressive COOH-terminal deletions. The EcoRI–

XbaI fragment was isolated from pRS47/EP (McCammon et al., 1990) and subcloned into pBluescript KS− (Stratagene, La Jolla, CA) to generate the plasmid pEX. A silent PstI site was introduced at nucleotide position 1281 (codon 150) to facilitate subcloning of a smaller piece of Pmp47 for further mutagenesis. The site was created using the PstI oligo (Table I) to generate plasmid pEPX. The 350-bp PstI–XbaI fragment was subcloned into pKS− to generate plasmid pPX.

SpeI sites were inserted into pPX after nucleotide positions 1,436 (amino acid 200), 1,508 (amino acid 224), and 1,568 (amino acid 244) using oligos Spe-1436, Spe-1508, and Spe-1568, respectively. The resulting plasmids were called pPSX-1436, pPSX-1508, and pPSX-1568, respectively. These sites delimit the beginning of transmembrane span 4 (amino acid 200), the end of span 4 (amino acid 224), and the end of the loop sequence (amino acid 244). The end of transmembrane span 5 is delimited by the endogenous XbaI site (amino acid 267) (McCammon et al., 1994). The SpeI sites were introduced in the same reading frame as the endogenous XbaI site. They also provided compatible cohesive ends with XbaI upon restriction digestion.

To fuse the truncation mutants of Pmp47 with the HA epitope tag, it was necessary to modify the polylinker of pBluescript KS− to create a NotI site that was in frame with the XbaI site of Pmp47. This was accomplished by subcloning a synthetic double-stranded linker (annealed oligos Recon-1 and Recon-2) into the XbaI–SacI sites of pKS−. The resulting vector was called pKS-recon. The synthetic fragment contains an in-frame NotI site (for subcloning the HA sequence) just downstream of the XbaI site, includes a stop codon after the NotI site, and provides a BglII site for cloning the fragment into yeast expression vectors. Three copies of the HA epitope sequence were cloned into this vector by transferring the NotI fragment from plasmid pSM491 (kindly provided by Carol Berkower and Susan Michaelis, Johns Hopkins University, Baltimore, MD) into the NotI site to generate plasmid pKS^r-HA. The EcoRI–XbaI fragment of Pmp47 was transferred from vector pEPX into the EcoRI–XbaI sites of pKS^r-HA to generate plasmid pKS-267PHA. The truncation mutants were constructed by replacing the wild-type PstI–XbaI fragment of pKS-267HA with the PstI–SpeI fragments of the pPSX series described above. The resulting plasmids were called pKS-200PHA, pKS-224PHA, and pKS-244PHA, respectively.

The truncation mutants were transferred to a yeast expression vector by subcloning the EcoRI–BglII fragments of pKS-267PHA, pKS-244PHA, pKS-224PHA, and pKS-200PHA into the EcoRI–BamHI sites of pgC20p (McNew and Goodman, 1994) to generate plasmids pY267PHA, pY244PHA, pY224PHA, and pY200PHA, respectively (*GAL1-10, ARS1-CEN4, URA3*).

Construction of an Oleate-inducible Expression Cassette Using the Pmp27 Promoter and Terminator

The ClaI–BamHI fragment of pRS313-PMP27, containing the entire open reading frame of Pmp27 (Marshall et al., 1995) and 650 bp of 5′ and 200 bp of 3′ sequence, was transferred to the ClaI–BamHI sites of pKS− to generate the plasmid pKS-27. Site-directed mutagenesis was used to introduce HindIII, EcoRI, and SmaI sites immediately before the endogenous ATG start codon (primer 5′-HES). A second primer (3′-SPB) was used to introduce SmaI, PstI, and BglII sites immediately upstream of the endogenous stop codon. Cleavage of the resulting plasmid with SmaI, followed by religation of the vector piece, effectively replaced the Pmp27 coding re-

Table I. Sequence of Synthetic Oligonucleotides

Oligo name	Sequence (5′–3′)
3′-SPB	GACATGTGGAAAGCTCCCGGGCTGCAGAGATCTTAGCTTTCTTTTCATC
5′-HES	GTAATAGTATAATCAATAAGCTTGAATTCCCCGGGGTCTGTGATACACTG
244-Bsp	CAATACACAATTTTTTCCGGAGAACAATTAAAATCA
Block-1	CAATACACAATTTTTGGAGGTGGCGGAGCCGGCGGTGGAAAAATTAAAAAGAGAAAT
Block-2	CAATTAAAATCATTTATTGTTGGAGGTGCCGGCGGTAATATTACACCTGTTGATGC
Block-3	GTTAAAATTAAAAAGAGAGGAGGTGGTGCCGGCGGAGGTTTATTATTAGGTGCTTTTGG
CAT-seq	CCGCTGGCGATTCAGTGGTT
Flag-bottom	GATCTCTGCAGGGATCCTTATTTATCGTCATCATCTTTATAATCAGCGC
Flag-top	GGCCGCGCTGATTATAAAGATGATGACGATAAATAAGGATCCCTGCAGA
GAKL-bottom	GAGATCTTTATAATTTAGCACCT
GAKL-top	CCGGAGGTGCTAAATTATAAAGATCTCTGCA
GMyc-bottom	GATCTCTGCAGTTAATTTAAATCTTCTTCTGAAATTAATTTTTGTTCCATT
GMyc-top	CCGGAATGGAACAAAAATTAATTTCAGAAGAAGATTTAAATTAACTGCAGA
Gseq	GCCCTTTCGAAAGATCCC
Gstop-bottom	GAGATCTTTAACCT
Gstop-top	CCGGAGGTTAAAGATCTCTGCA
Linker-1	CCGGAGAACAATTAAAATCATTTATTGTTAAAATTAAAAAGAGAAATATTACACCTGTTGATGCTG
Linker-2	CCGGCAGCATCAACAGGTGTAATATTTCTCTTTTTAATTTTAACAATAAATGATTTTAATTGTTCT
Myc-bottom	GATCTCTGCAGGGATCCTTAATTTAAATCTTCTTCTGAAATTAATTTTTGTTCCATACCGC
Myc-top	GGCCGCGGTATGGAACAAAAATTAATTTCAGAAGAAGATTTAAATTAAGGATCCCTGCAGA
Poly-1	CCGGCACTAGTAGCGGCCGCTCTAGATAGCTGCAGTAGCTAACTAGAGATCTG
Poly-2	GATCCAGATCTCTAGTTAGCTACTGCAGCTATCTAGAGCGGCCGCTACTAGTG
PstI oligo	GAGTATGGCTGCAGGTGCAGTTGCT
Recon-1	CTAGAAGCGGCCGCTAAAGATCTGAGCT
Recon-2	CAGATCTTTAGCGGCCGCTT
SB1-bottom	GGCCGCCACCAACAATAAATGATTTTAATTGTTCCCCACCG
SB1-top	CCGGCGGTGGGGAACAATTAAAATCATTTATTGTTGGTGGC
SB2-bottom	GGCCGCCACCTCTCTTTTTAATTTTCCCACCG
SB2-top	CCGGCGGTGGGAAAATTAAAAAGAGAGGTGGC
SB23-bottom	GGCCGCCACCAGCATCAACAGGTGTAATATTTCTCTTTTTAATTTTCCCACCG
SB23-top	CCGGCGGTGGGAAAATTAAAAAGAGAAATATTACACCTGTTGATGCTGGTGGC
SB3-bottom	GGCCGCCACCAGCATCAACAGGTGTAATATTCCCACCG
SB3-top	CCGGCGGTGGGAATATTACACCTGTTGATGCTGGTGGC
Spe-1436	GAAAAATGAAGGCTGGACTAGTCAAAAATTGTTTACTG
Spe-1508	CAATACACAATTTTTACTAGTGAACAATTAAAATCA
Spe-1568	ACACCTGTTGATGCTACTAGTTTATTATTAGGTGCT
Spe-repair	GGTGCCGGCGGAGGTACTAGTTTATTATTAGGTGCT
SR-bottom	CAGATCTCCTAG
SR-top	GATCCTAGGAGATCTGGTAC

gion with a polylinker consisting of HindIII, EcoRI, SmaI, PstI, and BglII restriction sites. This plasmid is called pKS-27poly.

Construction of Glycine Block Replacements in the Loop of 1-244

pPSX-1568, which is the parent plasmid for the 1-244 deletion mutant (described above), was used for the introduction of the glycine block mutations. The glycine blocks were introduced using site-directed mutagenesis with primers Block-1, Block-2, Block-3, and Spe-repair (Table I). The glycine blocks replaced the first eight, middle five, and last seven amino acids of the loop sequence (amino acids 225–244). The Spe-repair primer was needed to reintroduce the SpeI site of the Block-3 mutant. The glycine primers included a NgoMI site for diagnostic purposes, which resulted in the presence of a single alanine within the block of glycine residues (see Fig. 5). The resulting plasmids were called pPSX-1568-BI, pPSX-1568-BII, and pPSX-1568-BIII.

The glycine block mutants were returned to the 1-244 context by replacing the wild-type PstI–XbaI fragment of pKS-267PHA with the PstI–SpeI fragments of the pPSX series described above. The resulting plasmids were called pKS244-BI-HA, pKS244-BII-HA, and pKS244-BIII-HA, respectively. The coding sequence of each block mutant was transferred to the oleate-inducible expression cassette by subcloning an EcoRI–BglII fragment into the EcoRI–BglII sites of pKS-27poly. The 1-244 deletion mutant (without glycine blocks) was also transferred from plasmid pKS244-PHA to serve as a positive control for sorting experiments. The subsequent plasmids were called pKS-27-B0, pKS-27-BI, pKS-27-BII, and pKS-27-BIII. The gene cassettes were transferred to a yeast shuttle vector by subcloning a KpnI–XbaI fragment into the KpnI–XbaI sites of pRS316 (Sikorski and Hieter, 1989). The final constructs were called p27U-B0, p27U-BI, p27U-BII, and p27U-BIII (*CEN6*, *URA3*). The KpnI–SacI fragment of pKS-27-B0 was also transferred to pRS314 for expression in *pas10-1* (*CEN6*, *TRP1*).

Construction of CAT-loop, CAT-loop-HA, and CAT-HA

The CAT coding sequence was transferred to pBluescript KS− by subcloning a HindIII–BamHI fragment from pgC20p (McNew and Goodman, 1994) into the HindIII–BamHI sites of pKS−. The loop sequence of Pmp47 was fused to the COOH terminus of CAT by subcloning synthetic, annealed oligonucleotides (Linker-1 and -2) into the NgoMI site of CAT, which is located immediately upstream of the stop codon. Ligation of the loop sequence knocked out the 5′ NgoMI site, which was done for diagnostic and further cloning purposes (see below). The resulting plasmid was called pKS-CATL. The CAT-loop fusion was transferred back to the pgC20p yeast expression vector as a HindIII–BamHI fragment to generate clone pYCATL (*CEN6*, *URA3*).

The HindIII–BamHI fragment of pgC20p was also transferred to pUC18 (New England Biolabs) to generate plasmid pCAT-18. To facilitate the fusion of other sequences to the COOH terminus of CAT, the 3′ end of the CAT gene was reconstructed by subcloning a synthetic annealed fragment (oligos poly-1 and poly-2) into the NgoMI–BamHI sites of pCAT-18. The synthetic linker introduced SpeI, NotI, XbaI, stop codon, PstI, three frame-shifted stop codons, and BglII sequences immediately downstream of the CAT NgoMI site. The mutation was verified using the CAT-seq sequencing primer. The resulting plasmid was called pCAT18-poly.

CAT-loop-HA and CAT-HA were initially constructed by subcloning three copies of the HA epitope (taken as a NotI fragment from pSM491) into the NotI site of pCAT18-poly to generate the plasmid pCAT18-HA. The CAT-loop sequence was fused to HA by replacing the HindIII–NgoMI fragment of pCAT-18HA with the HindIII–NgoMI fragment of pKS-CATL. The resulting plasmid was called pCAT18-LHA. The HA sequences of both pCAT18-LHA and pCAT18-HA were truncated to two copies of the HA epitope by replacing their BamHI–KpnI fragments with a synthetic annealed fragment (oligos SR-top and SR-bottom) that introduced a stop codon immediately after the BamHI site of the HA sequence. A StyI site was included for diagnostic purposes. A BglII site was also included to provide a site for liberation of the coding sequence. The subsequent plasmids were called pCAT18-LHAΔ and pCAT18-HAΔ.

The CAT sequences were transferred to the oleate-inducible cassette by subcloning the HindIII–BglII fragments of pCAT18-LHAΔ and pCAT18-HAΔ into the HindIII–BglII sites of pKS-27poly. The resulting plasmids were called pKS27-LHA and pKS27-HA, respectively. Finally, the KpnI–XbaI fragment of pKS27-LHA was transferred to the KpnI–XbaI sites of YEplac112 (Gietz and Sugino, 1988) to generate YEpCAT-LHA, and the KpnI–SacII fragment of pKS27-HA was transferred to the KpnI–SacII sites of pRS314 (Sikorski and Hieter, 1989) to generate p27T-CAT-HA (*CEN6*, *TRP1*).

Construction of CAT-sufficiency Block Mutants

The sufficiency block mutants were constructed by subcloning synthetic, annealed oligos into the NgoMI–NotI sites of pCAT18-HAΔ (oligos SB1-top and -bottom, SB2-top and -bottom, SB3-top and -bottom, and SB23-top and -bottom). Each sufficiency block sequence was flanked on each side by three glycine residues to provide distance and flexibility between the block, CAT, and HA. The CAT-sufficiency mutants were inserted into the pKS-27poly vector (using HindIII–BglII sites) to pick up the oleate-inducible *cis* elements and then transferred to YEplac112 using KpnI–XbaI sites. The final plasmids were called YEp-SB1, YEp-SB2, YEp-SB3, and YEp-SB23 (2 μm, *TRP1*).

Replacement of HA with the FLAG or Myc Sequences in the Key Pmp47 Sorting Constructs

The HA tag was replaced in the 1-244 deletion mutant by digestion of pKS-244PHA with NotI and BglII and subcloning synthetic annealed oligos encoding for FLAG or myc sequences (oligos Flag-top and -bottom, Myc-top and -bottom, respectively). The FLAG sequence encoded an alanine followed by the eight-residue epitope D Y K D D D D K, a stop codon, and BamHI and PstI restriction sites (to be used for subsequent cloning steps; see below). The myc sequence encoded a glycine followed by the 12-residue epitope M E Q K L I S E E D L N (Evan et al., 1985), a stop codon, and BamHI and PstI restriction sites. The sequences were verified by di-deoxy sequencing, using Spe-1508 as a primer. The corresponding fusions between 1-224, Flag, and myc were generated by swapping the EcoRI–NotI fragment from pKS-224PHA into the 1-244 fusions. All fusions were then transferred to a yeast expression vector by subcloning an EcoRI–BglII fragment into the EcoRI–BamHI sites of pgC20p (McNew and Goodman, 1994). The resulting vectors (pY244-Myc, pY244-Flag, pY224-Myc, and pY224-Flag) allow galactose-inducible expression in *S. cerevisiae* (*GAL1-10*, *CEN4*, *URA3*).

Fusion of the Pmp47 Loop to the Green Fluorescent Protein

The plasmid pGFP-C1, which is a mammalian expression vector that carries the gene for the green fluorescent protein (GFP), was purchased from Clontech (Palo Alto, CA). To facilitate fusion of the Pmp47 loop to GFP, a BspEI site was introduced immediately upstream of the loop sequence in plasmid pKS-244-myc (primer 244-Bsp). The mutagenesis was verified by sequencing from the BspEI to the PstI site, just downstream of the myc tag (sequencing primer Spe-1436). The loop-myc sequence was fused to GFP by transferring the BspEI–BglII fragment to pGFP-C1. The Eco47III–BamHI fragment (which encodes the GFP-loop-myc fusion) was transferred to the SmaI–BglII sites of pKS-27poly to pick up the oleate-inducible Pmp27 promoter/terminator elements. This expression cassette was transferred to the high-copy yeast expression vector pRS424 (2 μm, *TRP1*) by subcloning a KpnI–SacI fragment. The final vector was called p27T-GFP-LM.

Several control constructs were created by adding myc, AKL (a PTS1), or a stop codon to the end of the GFP sequence. The AKL and stop codon sequences were fused to the end of GFP by digestion of pGFP-C1 with BspEI–PstI and subcloning synthetic annealed oligos (GAKL-top and -bottom, Gstop-top and -bottom, respectively). The AKL oligos encoded a spacer of two glycines, the alanine, lysine, and leucine motif, a stop codon, and then a BglII site. The stop codon sequences encoded just the two glycines, the stop codon, and the BglII site. The mutations were verified by sequencing (primer Gseq). The fusion sequences were transferred to pKS-27poly by subcloning an Eco47III–BglII fragment into the SmaI–BglII sites. The plasmids were called pKS27-GAKL and pKS27-Gstop. The gene cassettes were then transferred to the yeast expression vector pRS424 by subcloning a KpnI–SacI fragment. The resulting vectors were called p27T-GFP-AKL and p27T-GFP-stop (2 μm, *TRP1*).

The myc epitope was fused to GFP by cloning overlapping synthetic oligos into the BspEI–BglII sites of pKS27-GAKL (oligos Gmyc-top and -bottom). The primers encoded a glycine residue, the myc epitope sequence (described above), a stop codon, and a PstI site. After verification

by sequencing, the GFP–myc fusion was transferred to yeast vector pRS424 by subcloning a KpnI–SacI fragment. The final plasmid was called p27T-GFP-myc.

Results

Loop Sequence between Membrane Spans 4 and 5 Is Essential for Sorting to Peroxisomes

To define the region of sorting information within Pmp47 more precisely, we constructed a series of carboxy-terminal deletion mutants that were truncated at the junctions of span 4, loop, or span 5 (Fig. 1 *A*). Three copies of the HA epitope tag were fused to the carboxy termini to allow detection, and the constructs were expressed in *S. cerevisiae* under control of the *GAL1-10* promoter. The size of the fusion proteins on immunoblots suggested that each of the truncation mutants was synthesized correctly (Fig. 1 *B*), although the 1-200 mutant was much less stable than the others and often could not be detected.

We next determined the subcellular location of the mutants. A crude organellar pellet consisting of mostly peroxisomes and mitochondria was first prepared from cells expressing the constructs by differential centrifugation, and the peroxisomes and mitochondria were then separated on a 15–45% continuous Nycodenz gradient. Coomassie staining of the Nycodenz fractions (Fig. 2 *A*) revealed that most of the protein mass was in the mitochondrial fractions (the four fractions at the top of the gradient) compared to the peroxisomal fractions near the bottom of the gradient (Fig. 2 *B*), as we have seen previously (McNew and Goodman, 1994).

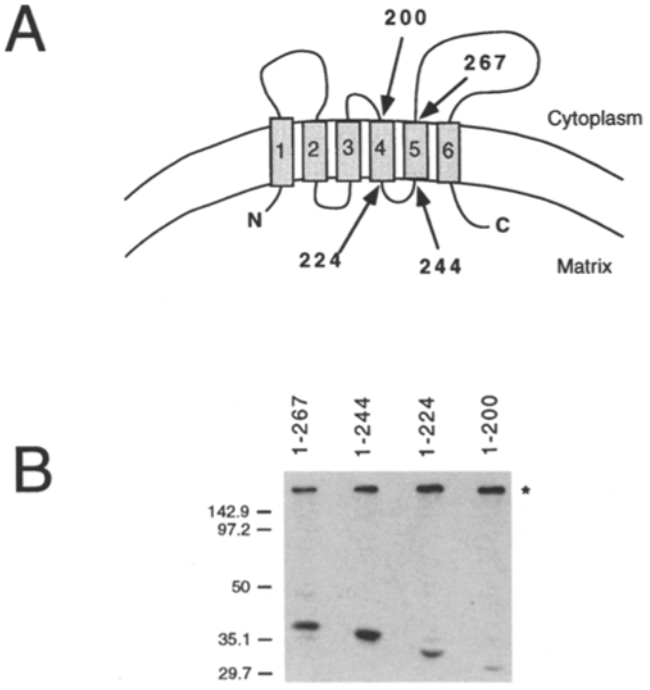

Figure 1. Expression of COOH-terminal deletions of Pmp47. (*A*) Topology map of Pmp47. Amino acid numbers at probable membrane interfaces of spans 4 and 5 are indicated. (*B*) Expression of HA fusions. Whole cell lysates from cells expressing the deletions, with the HA epitope fused to their carboxy termini, were subjected to SDS-PAGE and immunoblotting with the anti-HA antibody. Numbers above the lanes, the region of Pmp47 present; numbers to the left, the masses of standard proteins. The 1-200-HA construct was unstable and sometimes not detectable. The cross-reactive high molecular weight band serves as an internal loading control.

Both the Pmp47 1-267–HA and 1-244–HA fusions targeted to peroxisomes (Fig. 2 *C*), demonstrating that membrane span 5 (amino acids 245–267) was not necessary for peroxisomal targeting. While the distribution of fusion proteins between organellar supernatants and pellets was somewhat variable from experiment to experiment, the 1-244–HA construct always sorted better to the pellet than the 1-267–HA construct within any experiment. In contrast, removal of the loop sequence (to generate 1-224–HA) always obliterated peroxisomal targeting. The instability of the 1-200–HA protein made it difficult to reproducibly detect after Nycodenz fractionation, but long exposures of Western blots confirmed that it also did not sort to peroxisomes (data not shown), confirming a previous result with a 1-199–DHFR construct (McCammon et al., 1994). These results show that the hydrophilic loop between spans 4 and 5 is necessary for sorting of Pmp47 to peroxisomes.

Loop Can Target Chloramphenicol Acetyltransferase to Peroxisomes

Since our deletion experiments suggested that only the loop was important for targeting Pmp47, we then determined whether this region alone was sufficient for this function. Initially, we fused the loop to the COOH terminus of chloramphenicol acetyltransferase and expressed the fusion on a low-copy plasmid in *S. cerevisiae* under control of the *GAL1-10* promoter. This construct, however, was extremely unstable (confirmed by pulse–chase experiments; not shown), and we could not confirm its subcellular localization. To improve stability and detection, we expressed the fusion on a high-copy plasmid and added two copies of the HA epitope to the carboxy terminus, generating CAT-loop-HA. We also replaced the *GAL1-10* promoter with the *PMP27* oleate-responsive promoter (Marshall et al., 1995). The use of the *PMP27* promoter allowed synthesis of the fusions during peroxisomal proliferation (galactose inhibits this process) and detection of sorting even before proliferation was completed. Although protein staining of Nycodenz gradients revealed that peroxisomes and mitochondria assayed at this time (16–20 h after addition of oleic acid) were still resolved, blotting of peroxisomal marker proteins showed more extensive trailing toward the top of the gradient (Fig. 3). The lighter organelles in the trail probably consist of peroxisomal tubules and young maturing peroxisomes (Erdmann and Blobel, 1995). This is supported by the observation that the marker for peroxisomal membranes (Pmp27) is more enriched at higher fractions toward the top of the gradient than matrix proteins.

The CAT-loop-HA fusion sorted to peroxisomes (Fig. 3), indicating that the 20-amino acid loop was sufficient for sorting in the context of CAT and HA. It had a very similar blotting pattern in a Nycodenz gradient as the peroxisomal markers Pmp27 and thiolase. The negative control construct, CAT-HA, was almost exclusively cytoplasmic,

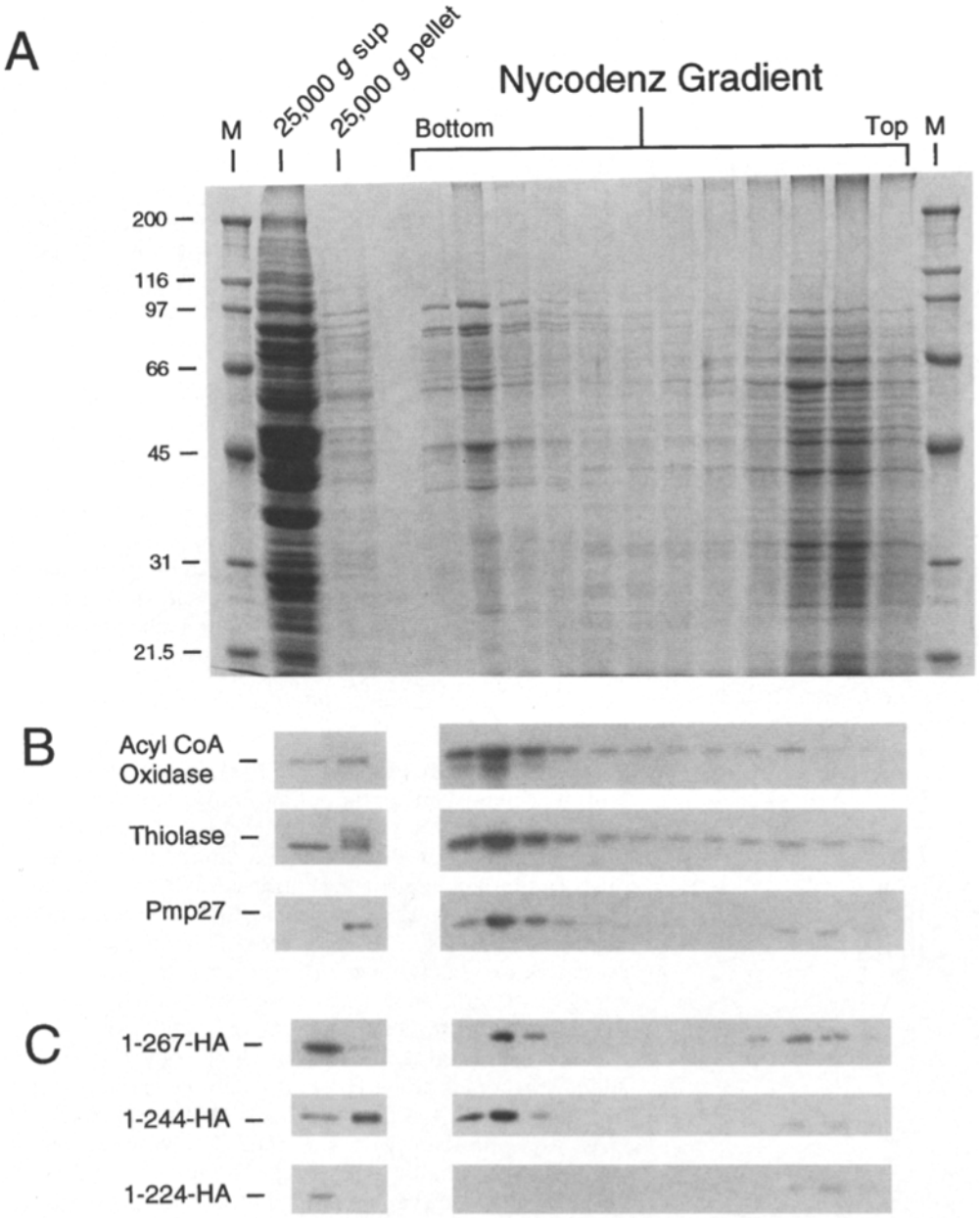

Figure 2. The loop between spans 4 and 5 is necessary for peroxisomal targeting. (*A*) The indicated fractions from cells expressing 1-244-HA were subjected to SDS-PAGE and Coomassie blue staining to show total protein and distribution of peroxisomes and mitochondria at the bottom and top of the gradient, respectively. Analysis of strains expressing the other fusions gave indistinguishable results. (*B*) Identical fractions to those shown in *A* were immunoblotted with antibodies to the indicated peroxisomal enzymes. (*C*) Sorting of deletion constructs. Analysis was similar to that shown in *B* except the anti-HA antibody was used. Constructs were driven by the *GAL1-10* promoter.

and none of this protein was seen in peroxisomal fractions. We term the targeting information within the loop mPTS.

To confirm that the loop was indeed targeting CAT to peroxisomes and not simply aggregating or sorting to organelles that comigrated with peroxisomes, we expressed CAT-loop-HA in cells containing a mutation in *PAS10* (van der Leij et al., 1992). This strain has abnormal peroxisomal membranes and fails to sort PTS1-containing proteins; we reasoned that its peroxisomes would migrate aberrantly in Nycodenz gradients. Furthermore, we allowed these cells to grow in oleic acid under conditions that promoted either a small or large extent of peroxisomal proliferation. These two conditions greatly influenced the migration pattern of peroxisomal marker proteins (Fig. 4, *A* and *B*). It can be seen in the figure that most of CAT-loop-HA comigrated with peroxisomal markers under both of these conditions. The nature of the minor fraction of CAT-loop-HA at the very bottom of the gradient shown in Fig. 4 *B* is unknown. As expected, CAT-HA failed to sort (data not shown). We also expressed the Pmp47 1-244–HA fusion in the *pas10-1* strain, and again this construct comigrated with peroxisomal markers (Fig. 4 *C*). Therefore, we conclude that these loop-containing constructs are mainly associated with peroxisomes and not in protein aggregates or targeting to comigrating organelles. It should be noted that in the *pas10-1* strain there were lower levels of peroxisomal marker proteins detected by Coomassie blue staining of organellar fractions, compared to those from the control strain. This explains the very weak signals of thiolase and Pmp27 in the supernatant and pellet fractions.

To probe the nature of the interaction between CAT-loop-HA and peroxisomes, a series of extractions was performed (data not shown). All of CAT-loop-HA in the Nycodenz gradient was particulate and spun down with membranes under low salt conditions. Much of CAT-loop-HA that was associated with peroxisomes was removed by

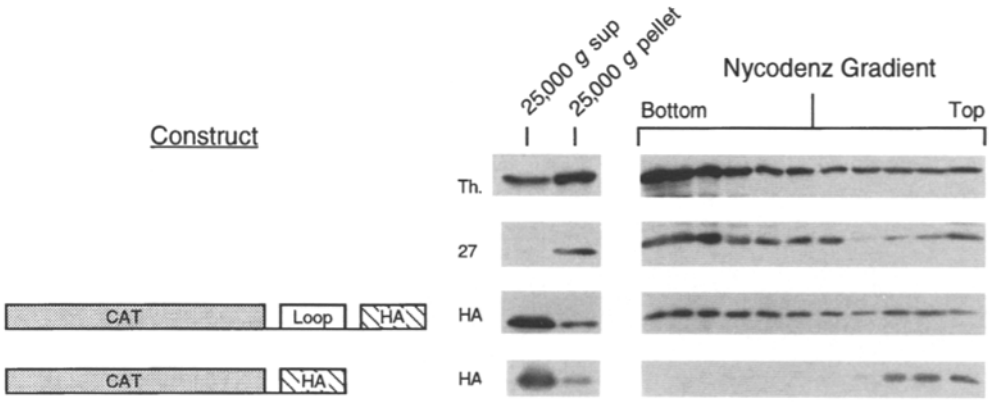

Figure 3. The loop is sufficient for peroxisomal targeting. Cells expressing the CAT-loop-HA construct were fractionationed and analyzed by immunoblots as in Fig. 2 for thiolase (*Th.*), Pmp27 (*27*), or fusion protein (*HA*). CAT-HA, the negative control, did not fractionate with peroxisomes. Constructs were driven by the oleate-inducible *PMP27* promoter.

the addition of 1 M NaCl and also did not remain with peroxisomal membranes when they were floated on a sucrose gradient. About half of CAT-loop-HA could not be removed from peroxisomes purified on Nycodenz, even with 0.1 M Na_2CO_3, pH 11. Clearly, the relationship of CAT-loop-HA with the peroxisomal membrane is complex, with a fraction bound peripherally and the rest bound more tightly, perhaps denatured within the membrane.

No Single Group of Amino Acids within the Loop Is Essential for Targeting

The sequence of the loop is E Q L K S F I V K I K K R N I T P V D A. An obvious feature of the loop is a block of basic residues (KIKKR) in the middle. To determine the relative importance of this basic cluster (termed Block II) and the flanking regions of the loop (Blocks I and III) for targeting, two complementary strategies were taken. First, in a loss-of-function approach, we modified the Pmp47 1-244–HA fusion such that glycines (with one alanine for construction purposes) were substituted for each block. Glycine residues were chosen instead of alanines since repetitive stretches of alanines might favor the formation of a helical structure (Chou and Fasman, 1978), which in turn might disrupt the nucleation of important secondary or tertiary interactions in the loop (e.g., a hairpin-like structure in which the first and last regions of the loop sequence interact).

Analysis of these glycine block replacement mutants driven by the *PMP27* promoter demonstrated that they all cofractionated, although to different extents, with the peroxisomal membrane marker (Fig. 5). Replacement of Block I with glycines (in the construct ΔBI) had the least effect; targeting of this construct most closely resembled

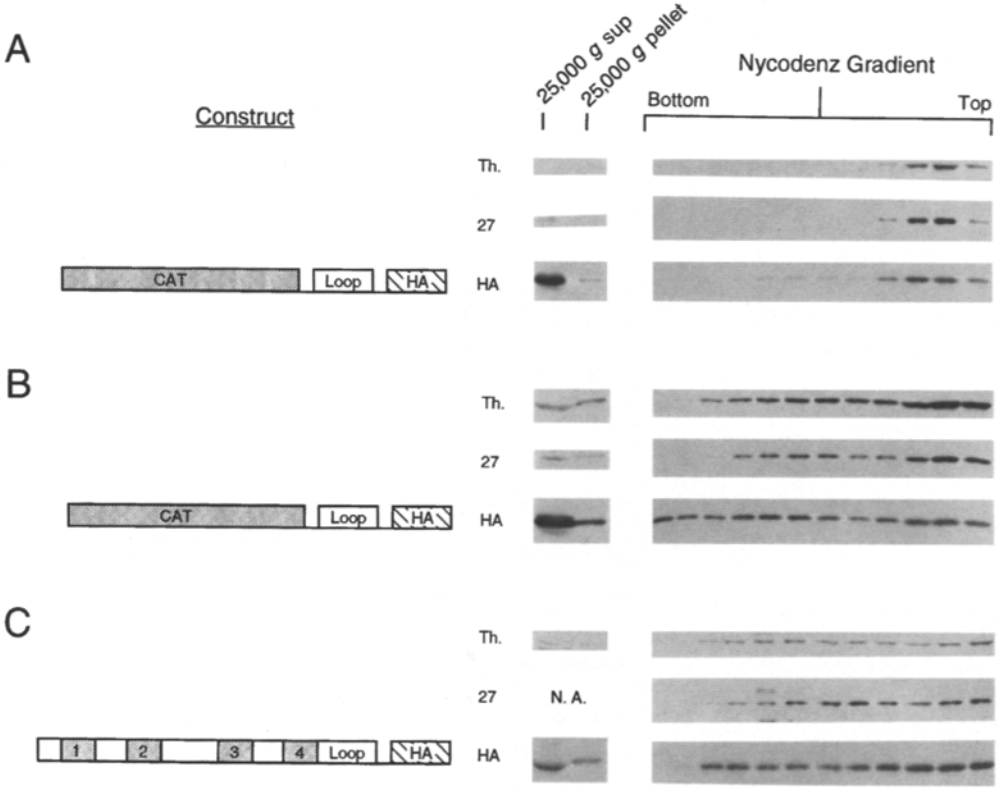

Figure 4. The loop targets in *pas10-1* cells. (*A*) CAT-loop-HA was expressed in the *pas10-1* strain. Cells were cultured as in Materials and Methods. Oleic acid was added when the cells were at OD_{600} of 0.1 and cultured for an additional 18 h. (*B*) Same as *A* except oleic acid was added at OD_{600} 0.6 and cultured for an additional 22 h. (*C*) Cells expressing Pmp47 1-244-HA were cultured to an OD_{600} of 0.4. Oleic acid was added, and cells were harvested after an additional 20 h. N.A., not available.

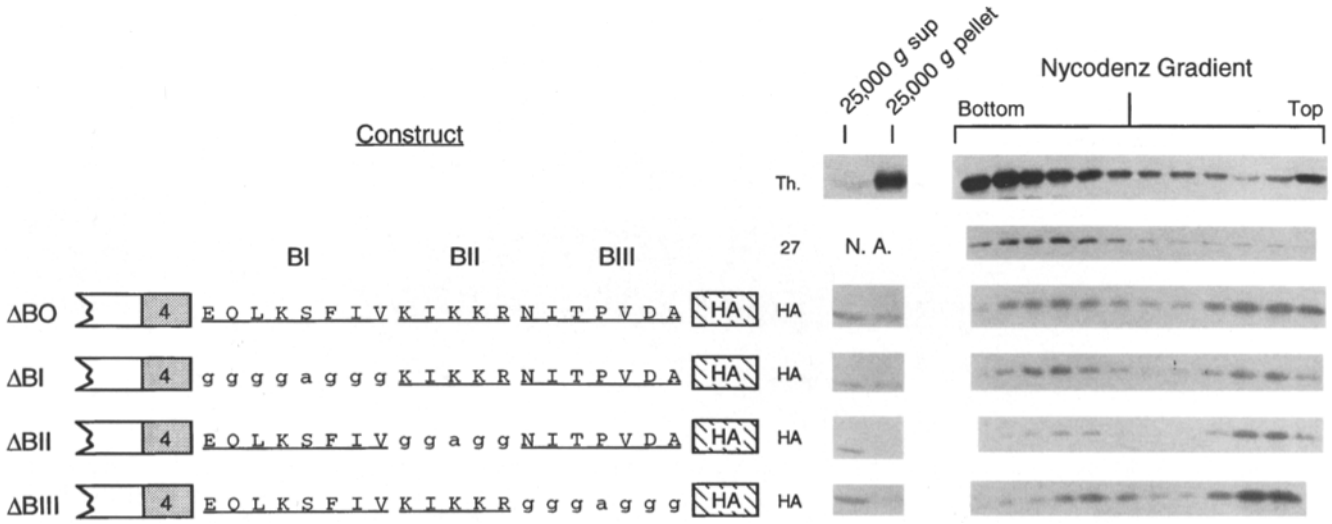

Figure 5. No single block of amino acids within the loop is required for targeting. Amino acids within the loop were grouped into Blocks I, II, or III, as indicated in the first construct, here termed *ΔB0*, which is identical to 1-244-HA. Constructs were made in which glycines and alanines (*g* and *a*) substituted for each block, as indicated. Their sorting was analyzed by SDS-PAGE and immunoblotting, as in previous figures. The thiolase and Pmp27 markers corresponded to the strain expressing 1-244-HA (*ΔB0*). Constructs were driven by the oleate-inducible *PMP27* promoter.

the positive control, 1-244–HA (shown in the figure as *ΔB0*). In contrast, substitution of Block II or Block III significantly reduced the efficiency of targeting (compare the ratio of signal in the peroxisomal vs mitochondrial region), although neither substitution totally abolished targeting. Replacement of Block II usually resulted in weaker sorting than replacement of Block III. The colocalization of these constructs with mitochondria may reflect a competition between recognition of degenerate mitochondrial signals in Pmp47 (Pmp47 is likely to contain some residual mitochondrial targeting information; it mislocalizes to this organelle in the absence of peroxisomal proliferation [McCammon et al., 1990; Sulter et al., 1993]), or nonspecific binding, since mitochondria comprise most of the protein mass within the Nycodenz gradient.

We conclude from the block substitution experiments that no single region of the loop is absolutely required for sorting to peroxisomes, provided that the remaining parts of the loop are still present. Since substitution of Block I with glycines had the least effect on targeting, the simplest interpretation of these data is that the targeting information is contained within Blocks II and III, and that Block II is most important for this function.

Block II or Block III Is Sufficient for Sorting CAT to Peroxisomes

To determine whether Blocks II or III were sufficient for peroxisomal targeting, each block, or both in combination, were placed between CAT and HA, and their ability to sort to peroxisomes was tested. The results from this gain-of-function approach are presented in Fig. 6. They demonstrate that both Blocks II and III contain information that is sufficient for targeting, although the targeting of Block II was stronger. Both blocks together sort with higher efficiency than either alone. In contrast, Block I is least able to target CAT-HA to peroxisomes. Although the Block I–containing fusion protein can be detected in the gradient, the blotting pattern is very different than that of the peroxisomal membrane marker. These results are completely consistent with the behavior of the glycine-block substitution mutations shown in Fig. 5.

Expression of CAT-Block III-HA resulted in two bands that were detected by Western blotting. The unexpected additional band migrates as a species 3 kD higher than the expected molecular weight of the construct. The sequence of Block III, N I T P V D A, contains the canonical N X T

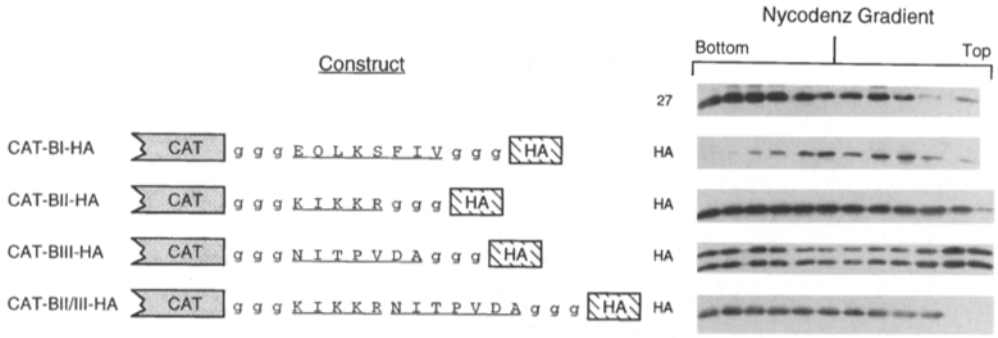

Figure 6. Block II or III is sufficient for peroxisomal targeting. Cells expressing the indicated constructs were fractionated and analyzed on Nycodenz gradients. The marker Pmp27 was analyzed for the CAT-BI-HA strain. Constructs were driven by the oleate-inducible *PMP27* promoter.

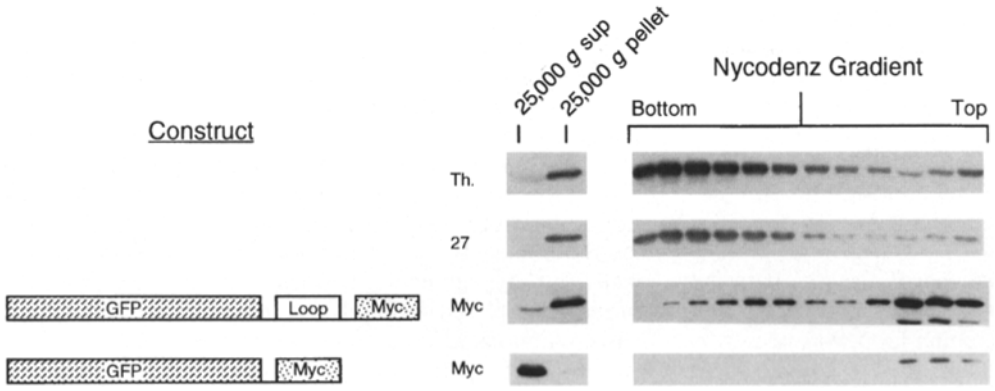

Figure 7. Loop can target a different carrier protein to peroxisomes. Strains containing GFP-loop-myc and GFP-myc were analyzed for targeting. Thiolase and Pmp27 were analyzed from the GFP-loop-myc containing strain.

sequence for N-linked glycosylation, and the molecular weight difference between the two bands is very close to that expected for a core glycosylation unit ($GlucNac_2$-Man_9-Glu_3). Digestion of the sample with N-glycosidase F, however, had no effect on the higher molecular weight species (data not shown). Therefore, the nature of the covalent modification remains unknown.

Peroxisomal Targeting by the mPTS Does Not Require CAT or HA

To test whether the epitope tag HA contributed in any way to the sorting of the Pmp47 truncation mutants, we substituted both FLAG and myc epitope tags for HA in the context of Pmp47 1-244. All of these proteins targeted with equal efficiency (not shown), demonstrating that the HA epitope did not contribute to the sorting signal.

To show that the targeting of the loop was not dependent on CAT, we placed it in a different context, substituting GFP for CAT and myc for HA. Fig. 7 shows that the loop was necessary and sufficient for peroxisomal targeting in this different context. Organellar GFP-loop-myc could not be visualized by fluorescence in yeast, in contrast to GFP-PTS1, which gave a clear punctate pattern (data not shown). Membrane flotation experiments indicated that GFP-loop-myc strongly interacted with the membrane. This association may have led to denaturation of the protein or thermal deexcitation, causing quenching of fluorescence.

Discussion

Sorting Sequence of Pmp47 Is Located within a Hydrophilic Loop

The mitochondrial carrier family of proteins, of which Pmp47 is a member, is predicted to span the membrane six times (Jank et al., 1993; Kuan and Saier, 1993). Internal stretches of homology suggest that this family arose by gene triplication, in which transmembrane spans 1, 2, and the intervening sequence are the repeating unit. Sorting analysis of the ADP/ATP translocator suggests that the mitochondrial sorting information may be present in each

Table II. Sequence Comparison of the Pmp47 Loop with Other Peroxisomal Membrane Proteins

	Preceding hydrophobic region					
	Number of residues	Sequence position	Average hydropathy index*	Block I	Block II	Block III
Pmp47A‡	19	206–224	1.7	EQLKSFIV	KIKKR	NITPVDA
Pmp47B‡	19	206–224	1.7	EQLKSFIV	KIKKR	NVTPVDA
Human 47 homologue§	‖	–	–	EGLKRQ	LLKKR	MKLSSLDV
S.c. 47 homologue¶	19	186–204	1.3	DTLKQRKL	RRKRE	NGLDIHLTNLET
Per8p**	12	182–193	2.5	KNVA	NLRKL	WGATKTVQDS
Per9p‡‡	20	16–35	1.4	KLAELNERLKEENFAKE	QIKRR	FKQTQNDC
Pas3p§§	22	18–39	2.2	KRWLYKQQLRITEQHFIKE	QIKRR	FEQTQEDS
Car1p‖‖	19	307–325	1.9	NRWRRWLARTWR	KTKKI	MSTTGGEG
Pmp70¶¶	19	221–239	1.1	SGLFLT	RLRRP	IGKMTIMEQ
Consensus	12–20	–	1.7	(4-17 a.a.) x x (K/R) (K/R) (3-7 a.a.) (T/S) x x (D/E) x		

*Kyte and Doolittle, 1982.
‡Moreno et al., 1994.
§Possible homologue from human EST library. GenBank accession number: R54274. Sequence corresponds to aa 220-209 of Pmp47. Sequence identity (similarity) = 26% (54%).
‖Not sequenced.
¶Possible homologue from Yeast Genome Sequencing project. ID: YIL134W. Sequence identity (similarity) = 27% (52%).
**Tan et al., 1995.
‡‡Baerends et al., manuscript submitted for publication.
§§Höhfeld et al., 1991.
‖‖Berteaux-Lecellier et al., 1995.
¶¶Kamijo et al., 1990.

of the repeated regions (Adrian et al., 1986; Pfanner et al., 1987; Smagula and Douglas, 1988). Overexpression of Pmp47 in the absence of activated peroxisomes leads to targeting to mitochondria, suggesting that Pmp47 may retain a low affinity for the mitochondrial sorting machinery (McCammon et al., 1990; Sulter et al., 1993). However, in addition to the loss of strong mitochondrial targeting sequences, Pmp47 must have gained a targeting sequence that caused it to be directed to peroxisomes.

Previous results demonstrated that the region between amino acids 199–267 of Pmp47, which comprises membrane spans 4, 5, and the intervening loop, was necessary for sorting to peroxisomes (McCammon et al., 1994). In the present study, we have shown that the loop itself is sufficient for targeting. Furthermore, we have shown by both loss-of-function and gain-of-function experiments that the sorting information resides in the last 12 amino acids of the loop.

Characteristics of the mPTS and a Hypothetical Structural Model

Since the peroxisomal targeting information within the loop is likely to have resulted from a gain-of-function mutation over the course of evolution, the nature of the targeting signal might be gleaned by noting which properties of the loop sequence are most different from the other mitochondrial carrier family proteins. Analysis of the alignment by Kuan and Saier (1993) reveals three unique features of the Pmp47 loop sequence (*underlined residues* in Table II). The first and most striking difference of the Pmp47 loop is the presence of a cluster of basic residues in the middle of the loop sequence (Block II). Second, this cluster is flanked on each side by a short stretch of hydrophobic residues that are enriched in branched-β amino acids (Ile and Thr in Blocks II and III, respectively). Third, Block III is terminated by a unique aspartic acid residue.

These motifs are also found in several other peroxisomal membrane proteins (*bold residues* in Table II). Block I appears to be least conserved in sequence content and length among these sequences, suggesting that it contains little if any sorting information. However, it may contribute to structural features of the signal conformation (see below). Blocks II and III are more highly conserved, especially in those elements that are most different from the mitochondrial carrier family sequences. There are at least two contiguous basic residues in each basic cluster, and the spacing between the threonine and aspartic acid residues is well conserved. Another feature that became apparent in analyzing these sequences is that the loop regions were always preceded by an extensive stretch of hydrophobic residues, most of which are predicted to be transmembrane spans. Although this stretch of residues is not necessary for targeting to peroxisomes, it may play an important role in the assembly of the protein into the lipid bilayer (see below).

Based on minimalist rules for de novo design (DeGrado et al., 1989) and secondary structure predictions (Chou and Fasman, 1978; Leszcynski and Rose, 1986), we suggest that the loop sequence itself may favor the formation of a stem-loop structure (Fig. 8 *A*). This model suggests that Block I begins as a short, hydrophilic, helical extension of transmembrane span 4. The rest of Block I and beginning of Block III form a short anti-parallel β-sheet. Block II is the intervening loop, which generates a bulb of positive charge. In this conformation, the sorting signal would be extended out from the body of the Pmp47 protein, thus promoting its accessibility to the sorting machinery.

Although Blocks II or III probably cannot adopt these structural properties when fused singly behind CAT, it

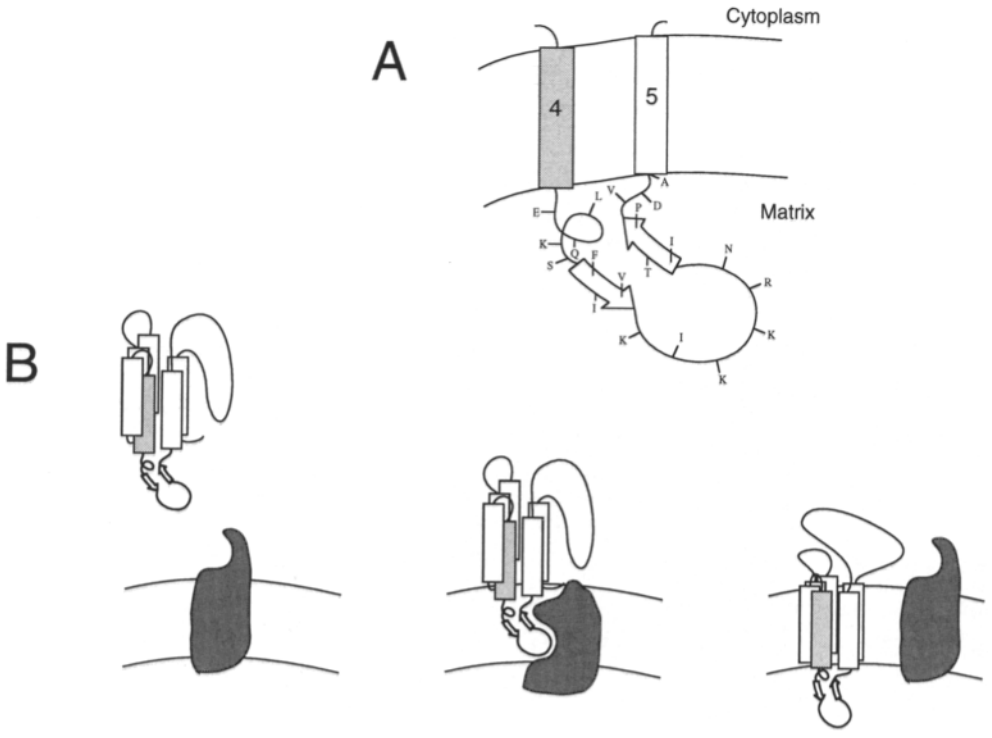

Figure 8. (*A*) Proposed structure of the Pmp47 targeting loop (see text for details). (*B*) A hypothetical model of the assembly of Pmp47 into the peroxisomal membrane. In the cytoplasm, span 4, the most hydrophobic span (*light shading*), is shielded from the aqueous environment by the other more amphipathic spans. Pmp47 binds to its receptor/assembly complex (*darkly shaded object*), which induces a conformational change, forcing Pmp47 into the membrane. This may promote closer interaction of span 4 with the bilayer and lateral diffusion of the carrier away from the receptor complex.

may not be necessary to generate these structures to expose the sorting information within the context of this carrier protein. The formal alternative is that the loop sequence may be less organized than we propose, such that each region of the loop operates independently of the other.

A Model for Peroxisomal Membrane Protein Assembly

In vitro studies using Pmp22 led Diestekötter and Just (1993) to propose a two-step mechanism for the assembly of peroxisomal membrane proteins. The first step was the binding of the protein to the membrane, which could occur at low temperatures but was strongly reduced by pretreatment of peroxisomes with protease. This step was followed by insertion, which was blocked at low temperatures and completely abolished by pretreatment with protease. These data imply that binding and subsequent insertion are mediated by a proteinaceous receptor.

The properties of mPTS described in this paper are fully consistent with this model. The loop, which contains the targeting information, is a hydrophilic sequence that likely requires a proteinaceous receptor for binding to the membrane. Blocks II and III may be recognized by a single receptor (Fig. 8 *B*), or the receptor may have several protein components. The presence of two targeting sequences (Block II and III) may be important for correctly orienting the incoming protein with respect to the assembly machinery. The interaction between the stem-loop and receptor may promote a conformational change that facilitates exposure of the lipid environment to the hydrophobic span that precedes the loop. This would in turn drive the protein into the bilayer.

This model may also explain our data with CAT– and GFP–loop fusions. We performed a series of extractions on peroxisomes isolated from cells expressing CAT-loop-HA to investigate the interaction between the loop sequence and the peroxisomal membrane. Since CAT normally exists as a stable trimeric molecule, and the CAT-loop-HA fusion protein lacked a hydrophobic span before the loop, we suspected that the presence of the bulky hydrophilic carrier protein might prevent proper assembly in the membrane.

Our extraction data suggest that CAT-loop-HA might be engaged with a surface-bound receptor or might have simply denatured on the surface of the organelle or within the bilayer after interacting with the membrane assembly machinery. We have tried, without success, to localize CAT-loop-HA by immunofluorescence and immunoelectron microscopy in yeast and mammalian cells. We speculate that the interaction of CAT-loop-HA with the membrane sterically hindered the accessibility of the CAT or HA antibody. In contrast, we were able to easily localize CAT-PTS1 or CAT-HA (when coimported with CAT-PTS1) constructs within the peroxisomal matrix (McNew and Goodman, 1994; McNew and Goodman, 1996).

We had similar difficulties in analyzing the GFP-loop-myc fusion protein. Biochemically, it is clear that GFP-loop-myc is found predominantly in the organellar pellet fraction (Fig. 7) and is observed in peroxisomal fractions. In addition, there is nearly the same amount of GFP-loop-myc protein present as the negative control construct GFP-myc, which is clearly cytoplasmic. However, fluorescence analysis of living cells expressing GFP-loop-myc showed very little fluorescence, while cells expressing GFP-myc showed high levels of cytoplasmic fluorescence (data not shown). Although it could be argued that the addition of the loop to GFP destabilizes the protein, it is more likely that the protein is entangled in the membrane, resulting in quenching of fluorescence.

Now that an mPTS has been identified, it may be easier to identify components of the machinery that catalyze the assembly of integral peroxisomal membrane proteins. Several cytoplasmic and peroxisomal factors have already been described that catalyze the targeting and import of matrix proteins. It will be interesting to determine whether the assembly of membrane proteins requires a completely separate machine, or whether a common import pathway exists for the two classes of proteins.

We thank Jon Rothblatt for the anti-thiolase antibody and Ben Distel for the *pas10-1* strain. Kimberly Campbell, Jehangir W. Gowani, and Yolanda Thomas provided excellent technical assistance.

This work was supported by National Institutes of Health (NIH) postdoctoral fellowship GM16974 (to J.M. Dyer), NIH Pharmacological Sciences training grant T32 GM-07062 (to J.A. McNew), and NIH grant GM31859 and The Robert A. Welch Foundation grant I-1085 (to J.M. Goodman).

Received for publication 19 October 1995 and in revised form 5 February 1996.

References

Adrian, G.S., M.T. McCammon, D.L. Montgomery, and M.G. Douglas. 1986. Sequences required for delivery and localization of the ADP/ATP translocator to the mitochondrial inner membrane. *Mol. Cell. Biol.* 6:626–634.

Bassford, P., J. Beckwith, K. Ito, C. Kumamoto, S. Mzushima, D. Oliver, L. Randall, T. Silhavy, P.C. Tai, and B. Wickner. 1991. The primary pathway of protein export in *E. coli*. *Cell.* 65:367–368.

Berteaux-Lecellier, V., M. Picard, C. Thompson-Coffe, D. Zickler, A. Panvier-Adoutte, and J.-M. Simonet. 1995. A nonmammalian homolog of the PAF1 gene (Zellweger syndrome) discovered as a gene involved in caryogamy in the fungus Podospora anserina. *Cell.* 81:1043–1051.

Blobel, G., P. Walter, C.N. Chang, B.M. Goldman, A.H. Erickson, and V.R. Lingappa. 1979. Translocation of proteins across membranes: the signal hypothesis and beyond. *Symp. Soc. Exp. Biol.* 33:9–36.

Bodnar, A., and R. Rachubinski. 1991. Characterization of the integral membrane polypeptides of rat liver peroxisomes isolated from untreated and clofibrate-treated rats. *Biochem. Cell Biol.* 69:499–508.

Chou, P.Y., and G.D. Fasman. 1978. Prediction of protein secondary structure. *Adv. Enzymol. Relat. Areas Mol. Biol.* 47:45–148.

DeGrado, W.F., Z.R. Wasserman, and J.D. Lear. 1989. Protein design, a minimalist approach. *Science (Wash. DC).* 243:622–628.

Diestelkötter, P., and W.W. Just. 1993. In vitro insertion of the 22-kD peroxisomal membrane protein into isolated rat liver peroxisomes. *J. Cell Biol.* 123: 1717–1725.

Economu, A., and W. Wickner. 1994. SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. *Cell.* 78:835–843.

Erdmann, R., and G. Blobel. 1995. Giant peroxisomes in oleic acid–induced *Saccharomyces cerevisiae* lacking the peroxisomal membrane protein Pmp27p. *J. Cell Biol.* 128:509-523.

Evan, G.I., G.K. Lewis, G. Ramsay, and J.M. Bishop. 1985. Isolation of monoclonal antibodies specific for human *c-myc* proto-oncogene product. *Mol. Cell. Biol.* 5:3610–3616.

Gietz, R.D., and A. Sugino. 1988. New yeast-*Eschericia coli* shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites. *Gene (Amst.).* 74:527–534.

Glover, J.R., D.W. Andrews, and R.A. Rachubinski. 1994. *Saccharomyces cerevisiae* peroxisomal thiolase is imported as a dimer. *Proc. Natl. Acad. Sci. USA.* 91:10541–10545.

Gorlich, D., and T.A. Rapoport. 1993. Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum. *Cell.* 75:615–630.

Gould, S.J., G.-A. Keller, N. Hosken, J. Wilkinson, and S. Subramani. 1989. A conserved tripeptide sorts proteins to peroxisomes. *J. Cell Biol.* 108:1657–1664.

Hannavy, K., S. Rospert, and G. Schatz. 1993. Protein import into mitochondria: a paradigm for the translocation of polypeptides across membranes. *Curr. Opin. Cell Biol.* 5:694–700.

Höhfeld, J., M. Veenhuis, and W. H. Kunau. 1991. PAS3, a *Saccharomyces cerevisiae* gene encoding a peroxisomal integral membrane protein essential for peroxisome biogenesis. *J. Cell Biol.* 114:1167–1178.

Jank, B., B. Habermann, R.J. Schweyen, and T.A. Link. 1993. PMP47, a peroxisomal homologue of mitochondrial solute carrier proteins. *Trends Biochem. Sci.* 18:427–428.

Jungnickel, B., T.A. Rapoport, and E. Hartmann. 1994. Protein translocation: common themes from bacteria to man. *FEBS Lett.* 346:73–77.

Kamijo, K., S. Taketani, S. Yokota, T. Osumi, and T. Hashimoto. 1990. The 70-kDa peroxisomal membrane protein is a member of the Mdr (P-glycoprotein)-related ATP-binding protein superfamily. *J. Biol. Chem.* 265:4534–4540.

Kuan, J., and M.H. Saier. 1993. The mitochondrial carrier family of transport proteins: structural, functional, and evolutionary relationships. *Crit. Rev. Biochem. Mol. Biol.* 28:209–233.

Kunkel, T.A., J.D. Roberts, and R.A. Zakour. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* 154: 367–382.

Kyte, J., and R.F. Doolittle. 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157:105–132.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond.).* 227:680–685.

Lazarow, P.B., and Y. Fujiki. 1985. Biogenesis of peroxisomes. *Annu. Rev. Cell Biol.* 1:489–530.

Leszcynski, J.F., and G.D. Rose. 1986. Loops in globular proteins: a novel category of secondary structure. *Science (Wash. DC).* 234:849–855.

Liu, H., X. Tan, K.A. Russell, M. Veenhuis, and J.M. Cregg. 1995. *PER3*, a gene required for peroxisome biogenesis in *Pichia pastoris*, encodes a peroxisomal membrane protein involved in protein import. *J. Biol. Chem.* 270: 10940–10951.

Marshall, P.A., Y.I. Krimkevich, R.H. Lark, J.M. Dyer, M. Veenhuis, and J.M. Goodman. 1995. Pmp27 promotes peroxisomal proliferation. *J. Cell Biol.* 129:345–355.

McCammon, M.T., C.A. Dowds, K. Orth, C.R. Moomaw, C.A. Slaughter, and J.M. Goodman. 1990. Sorting of peroxisomal membrane protein PMP47 from *Candida boidinii* into peroxisomal membranes of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 265:20098–20105.

McCammon, M.T., J.A. McNew, P.J. Willy, and J.M. Goodman. 1994. An internal region of the peroxisomal membrane protein PMP47 is essential for sorting to peroxisomes. *J. Cell Biol.* 124:915–925.

McNew, J.A., and J.M. Goodman. 1994. An oligomeric protein is imported into peroxisomes in vivo. *J. Cell Biol.* 127:1245–1257.

McNew, J.A., and J.M. Goodman. 1996. The targeting and assembly of peroxisomal proteins: some old rules do not apply. *Trends Biol. Sci.* In press.

McNew, J.A., K. Sykes, and J.M. Goodman. 1993. Specific cross-linking of the proline isomerase cyclophilin to a non-proline-containing peptide. *Mol. Biol. Cell.* 4:223–232.

Moreno, M., R. Lark, K.L. Campbell, and J.M. Goodman. 1994. The peroxisomal membrane proteins of *Candida boidinii*: gene isolation and expression. *Yeast.* 10:1447–1457.

Pfanner, N., P. Hoeben, M. Tropschug, and W. Neupert. 1987. The carboxyl-terminal two-thirds of the ADP/ATP carrier polypeptide contains sufficient information to direct translocation into mitochondria. *J. Biol. Chem.* 262: 14851–14854.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. 2nd edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Schwarz, E., and W. Neupert. 1994. Mitochondrial protein import: mechanisms, components and energetics. *Biochim. Biophys. Acta.* 1187:270–274.

Sikorski, R.S., and P. Hieter. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics.* 122:19–27.

Smagula, C.S., and M.G. Douglas. 1988. ADP-ATP carrier of *Saccharomyces cerevisiae* contains a mitochondrial import signal between amino acids 72 and 111. *J. Cell. Biochem.* 36:323–327.

Sulter, G., H. Waterham, E. Vrieling, J. Goodman, W. Harder, and M. Veenhuis. 1993. Expression and targeting of a 47 kDa integral peroxisomal membrane protein of *Candida boidinii* in wild type and a peroxisome-deficient mutant of *Hansenula polymorpha*. *FEBS Lett.* 315:211–216.

Swinkels, B.W., S.J. Gould, A.G. Bodnar, R.A. Rachubinski, and S. Subramani. 1991. A novel, cleavable peroxisomal targeting signal at the amino-terminus of the rat 3-ketoacyl-CoA thiolase. *EMBO (Eur. Mol. Biol. Organ.) J.* 10: 3255–3262.

Tan, X., H.R. Waterham, M. Veenhuis, and J.M. Cregg. 1995. The *Hansenula polymorpha PER8* gene encodes a novel peroxisomal integral membrane protein involved in proliferation. *J. Cell Biol.* 128:307–319.

Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA.* 76:4350–4354.

van der Leij, I., M. van den Berg, R. Boot, M. Franse, B. Distel, and H.F. Tabak. 1992. Isolation of peroxisome assembly mutants from *Saccharomyces cerevisiae* with different morphologies using a novel positive selection procedure. *J. Cell Biol.* 119:153–162.

Veenhuis, M., and J.M. Goodman. 1990. Peroxisomal assembly: membrane proliferation precedes the induction of the abundant matrix proteins in the methylotrophic yeast *Candida boidinii*. *J. Cell Sci.* 96:583–590.

Walton, P.A., P.E. Hill, and S. Subramani. 1995. Import of stably folded proteins into peroxisomes. *Mol. Biol. Cell.* 6:675–683.

Waterham, H.R., V.I. Titorenko, P. Haima, J.M. Cregg, W. Harder, and M. Veenhuis. 1994. The *Hansenula polymorpha PER1* gene is essential for peroxisome biogenesis and encodes a peroxisomal matrix protein with both carboxy- and amino-terminal targeting signals. *J. Cell Biol.* 127:737–749.

Yost, C.S., J. Hedgpeth, and V.R. Lingappa. 1983. A stop transfer sequence confers predictable transmembrane orientation to a previously secreted protein in cell-free systems. *Cell.* 34:759–766.