THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 14, Issue of April 6, pp. 10897–10905, 2001

Printed in U.S.A.

Discrete Targeting Signals Direct Pmp47 to Oleate-induced Peroxisomes in *Saccharomyces cerevisiae**

Received for publication, December 1, 2000, and in revised form, December 28, 2000
Published, JBC Papers in Press, January 16, 2001, DOI 10.1074/jbc.M010883200

Xiaodong Wang, Michael J. Unruh, and Joel M. Goodman‡

From the Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9041

Pmp47 is a peroxisomal membrane protein consisting of six transmembrane domains (TMDs). We previously showed that the second matrix loop containing a basic cluster of amino acids is important for peroxisomal targeting, and similar basic targeting motifs have been found in other peroxisomal membrane proteins. However, this basic cluster by itself targets to peroxisomes very poorly. We have developed a sensitive quantitative localization assay based on the targeting of Pmp47-GFP fusion proteins to identify the important elements of the basic cluster and to search for other targeting information on Pmp47. Our data suggest that side-chain structure and position as well as charge are important for targeting by the basic cluster. Analysis of other regions of Pmp47 indicates that all TMDs except TMD2 can be eliminated or substituted without significant loss of targeting. TMD2 plus an adjacent cytoplasmic-oriented sequence is crucial for targeting. Cytoplasmic-oriented sequences from two other peroxisomal membrane proteins, *Sc*Pex15p and *Sc*Pmp22, could partially substitute for the analogous sequence in Pmp47. Targeting with high fidelity to oleate-induced peroxisomes required the following elements: the cytoplasmic-oriented sequence and TMD2, a short matrix loop containing a basic cluster, and a membrane-anchoring TMD.

Peroxisomes consist of a dense matrix and surrounding membrane. Proteins are targeted to the peroxisomal matrix from the cytosol by virtue of peroxisomal targeting signals (PTSs)[1] (1). Most matrix proteins contain a PTS1 sequence at their extreme carboxyl terminus consisting of three amino acids identical or related to the tripeptide SKL. A few matrix proteins contain a PTS2 sequence close to the amino terminus, which is often cleaved upon entering the peroxisomal assembly pathway. Receptors for PTS1 and PTS2 have been identified, as have docking proteins for the PTS receptors on the peroxisomal surface (2, 3). Receptor and cargo may translocate through the membrane before receptors are recycled to the cytosol (4, 5).

Integral membrane proteins lack these PTS sequences and must be targeted by alternative sequences. We have used Pmp47 from the methylotrophic yeast *Candida boidinii* as a model protein to understand targeting to the peroxisome membrane. Pmp47 shares homology with members of the mitochondrial carrier protein family (6). These proteins span the membrane six times, and sequence similarity within them suggests that the family was derived from an ancient gene triplication. Pmp47 is important in the metabolism of medium-chain fatty acids and may transport ATP (7), and it is also critical for the import of dihydroxyacetone synthase, an abundant matrix protein (8). Pmp47 contains two loops of amino acids on the matrix side and three on the cytoplasmic side of the membrane. We previously identified a targeting motif contained within the second matrix loop that we termed the mPTS (9). We showed that the most important attribute of the targeting loop is a basic cluster of amino acids of the sequence KIKKR. Matrix-oriented basic clusters have since been shown to be an important targeting element in several other peroxisomal membrane proteins (10–13).

The Pmp47 second matrix loop fused to a carrier protein targets to peroxisomes very weakly (9), as determined by organelle fractionation. Similarly, a basic cluster from the cotton seed ascorbate peroxidase (APX), fused to a synthetic membrane span and carrier protein, targets to punctate and reticular structures, although peroxisomal localization was not demonstrated (13). In contrast, a cytosol-oriented region of 22 amino acids in rat Pmp22 can target a different fragment of the protein, which by itself is not sufficient for targeting, to the peroxisomal membrane (14). Because some peroxisomal membrane proteins integrate into peroxisomal precursor organelles before arriving at peroxisomes (15), a more complex targeting signal may be required to accommodate such a pathway.

We now describe experiments to indicate that Pmp47 requires discrete regions for effective targeting. To gain evidence for the relative contribution of each region we have established a sensitive and quantitative *in vivo* fluorescence localization assay in *Saccharomyces cerevisiae*. Results from the assay suggest that charge is not the sole factor governing targeting by the basic cluster; side-chain structure and residue position may also be important. We show that targeting at low efficiency occurs even when the basic cluster is eliminated, indicating other targeting elements must exist within Pmp47. We have identified the other important signal as consisting of a cytosol-oriented charged domain within cytoplasmic loop 1 and the adjacent transmembrane domain, TMD2. TMD2 by itself is poorly hydrophobic. Efficient targeting with high fidelity is only achieved by combining these discrete motifs (including the matrix basic cluster) with a strong transmembrane anchoring domain. Thus, effective targeting of Pmp47 requires discrete regions of the protein on both sides, and within, the membrane.

* This work was generously supported by grants from The Welch Foundation (I-1085), The National Institutes of Health (GM31859), and the Department of Pharmacology, University of Texas Southwestern. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Pharmacology, University of Texas Southwestern Medical School, 5323 Harry Hines Blvd., Dallas, TX 75390-9041. Tel.: 214-648-2359; Fax: 214-648-2994; E-mail: Joel.Goodman@UTSouthwestern.edu.

[1] The abbreviations used are: PTS, peroxisomal targeting signal; mPTS, peroxisomal membrane targeting signal; GFP, green fluorescence protein; PCR, polymerase chain reactions; PGK, phosphoglycerate kinase; SOEing, splicing by overlap extension; thiolase, peroxisomal 3-ketoacyl-CoA thiolase; TMD, transmembrane domain; kb, kilobase(s); APX, ascorbate peroxidase.

EXPERIMENTAL PROCEDURES

Strains and Culturing Conditions

Escherichia coli strain TG-1 (F′ *traD36 laclq* Δ*[lacZ]M15 proA+B+/ supE* Δ *[hsdM-mcrB]*5 [$r_k^-m_k^-$McrB$^-$] *thi* Δ*(lac-proAB)*) was used for plasmid amplification. Yeast *S. cerevisiae* strain MMYO11α was used in all the expression experiments (16). MMYO11α was transformed by the lithium acetate method (17). Yeast transformants were maintained on synthetic complete plates (2% glucose and 0.67% yeast nitrogen base without amino acids (Difco Laboratories, Detroit, MI)) supplemented with appropriate amino acids and bases. For glucose liquid culture, MMYO11α was grown in synthetic dextrose medium (2% glucose and 0.67% yeast nitrogen base) supplemented with appropriate amino acids and bases. To induce peroxisome proliferation in MMYO11α, cells were cultured in SGd medium (3% glycerol, 0.1% glucose, 0.67% yeast nitrogen base, supplemented with appropriate amino acids and bases) to an A_{600} of 1 to 2. They were then "boosted" by adding 0.1 volume of 10% yeast extract, 20% peptone and cultured for 4 h. Cells were harvested by centrifugation, re-inoculated to a final A_{600} of 1 into oleate medium (a semisynthetic medium (18) supplemented with 0.1% (v/v) oleic acid, 0.2% (v/v) Tween 40 and appropriate amino acids and bases), and cultured for 16–20 h.

Recombinant DNA Procedures and Reagents

Restriction enzymes and other DNA-modifying enzymes were purchased from New England BioLabs (Beverly, MA). Isolation of DNA fragments and plasmids was carried out using the QIAEXII gel extraction kit and QIAprep Spin Miniprep kit from Qiagen, Inc. (Valencia, CA). Polymerase chain reactions (PCR) were performed with a PTC-100 programmable thermal controller from MJ Research, Inc. (Watertown, MA) using *Pfu* DNA polymerase. Standard recombinant DNA techniques were performed (19).

Sources for Antibodies and Plasmids

Antibodies used in this study include anti-3-ketoacyl-CoA thiolase (anti-thiolase) polyclonal antibodies (kind gift of Jonathan Rothblatt), and Texas Red goat anti-rabbit IgG (H+L) conjugate (Molecular Probes, Eugene, OR).

Plasmid constructs used in this study include pEMBLy3012, containing the yeast phosphoglycerate kinase (PGK) promoter and terminator (kind gift of Michael White, University of Texas Southwestern Medical Center, Dallas, TX), pRS315 and pRS316 (20), and vectors pGFP-C1 and pDsRed (CLONTECH Laboratories, Inc., Palo Alto, CA).

Construction of Plasmids

Pmp47-(1–267)-GFP—The 1.8-kb PGK promoter-terminator was excised from pEMBLy3012 by *Hin*dIII and subcloned into the unique *Hin*dIII site of yeast centromeric plasmid pRS315. The resulting construct was named pRS315-PGK. A DNA fragment containing amino acids 1–267 of Pmp47 was isolated from *Candida boidinii* genomic DNA by PCR using upper primer W5U (5′-GAA GAT CTA TGT CTA CAA GAG AGT ACG AC-3′), which introduced a *Bgl*II site at the amino terminus (restriction sites introduced in this oligonucleotide and those described subsequently are *underlined*), and lower primer W5L (5′-AGG TAA AGA TTG AAC GCT ATC-3′). A unique endogenous *Xba*I site is present at amino acid 267. The GFP coding region was amplified from pGFP-C1 by PCR using upper primer W4U (5′-GCT CTA GAA TGG GTA AAG GAG AAG AAC TT-3′), which introduced an in-frame *Xba*I site to its amino terminus, and lower primer W4L (5′-GAA GAT CTT TAC TTG TAT AGT TCA TCC ATG CC-3′), which introduced a stop codon followed by a *Bgl*II site to its carboxyl terminus. The Pmp47-(1–267) fragment and the GFP coding region were both digested with *Bgl*II and *Xba*I, and were then cloned into the unique *Bgl*II site between the PGK promoter and terminator of the pRS315-PGK to yield pPmp47-(1–267)-GFP.

Pmp47-GFP Mutant Constructs—To facilitate the construction of mutant constructs from pPmp47-(1–267)-GFP, the *Bgl*II site between GFP and PGK terminator was disrupted by PCR mutagenesis to produce the construct D004. D004 has a unique *Bgl*II site preceding the start codon of the fusion protein and a unique *Nco*I site within the GFP coding region (at amino acid 55 of GFP). The DNA fragment encoding Pmp47-(1–267)-GFP (1–55) was used as a cassette for constructing the mutant Pmp47-GFP fusion proteins. The mutated forms of this cassette were cloned back into D004, replacing the original sequence.

To construct the amino-terminal truncation mutants, the coding regions for the amino-terminally truncated Pmp47-(*x*-267)-GFP (1–55) fragments were isolated from D004 by PCR using specific upper primers with a *Bgl*II site and start codon included at the 5′-end and a common lower primer W20L (5′-ACA CCA TAA GAG AAA GTA GTG A-3′) that is located 3′ of the endogenous *Nco*I site.

The carboxyl-terminal truncations, replacements, and internal deletions were constructed using PCR SOEing (Splicing by Overlap Extension) (21, 22). Upper primers starting at specific amino-terminal regions of Pmp47 with an upstream *Bgl*II site and a start codon were used as outer upper primers, and W20L was used as the common outer lower primer. A pair of inner SOEing primers, each of which contains the mutated sequences in its 5′-overlapping region, was used for each replacement construct. Briefly, the outer upper primer and the lower inner primer were used to isolate the amino-terminal flanking region of the replacement. The inner upper primer and the outer lower primer W20L were used to generate the carboxyl-terminal flanking region. Then the two fragments, which contained the mutated sequences in their overlapping region, were denatured, annealed to each other, and finally fused together by overlap extension. The carboxyl-terminal truncations and internal deletions were generated in a similar fashion, where the overlapping regions served solely as fusion junctions without introducing any replacement amino acids.

Block I and III deletion constructs were made from pPmp47-(1–243)-GGG-GFP (see Table I). The TMD4 substitution constructs were made from pPmp47-(1–267)-GFP, the TMD3 substitution construct was made from pPmp47-(70–267)-GFP, and the TMD2 exchange constructs were made from pPmp47-(53–267)-GFP (see Table III).

The Pmp47-derived sequences of all fusion constructs were confirmed by DNA sequencing.

DsRed-AKL—First, the PGK promoter-terminator from pEMBLy3012 was subcloned into the unique *Hin*dIII site of the yeast centromeric plasmid pRS316 to yield pRS316-PGK. The coding region of the red fluorescence protein DsRed was amplified from the plasmid pDsRed by two rounds of PCR. The upper primer W71U (5′-AAT GAA GAT CTA TGA GGT CTT CCA AGA ATG TTA-3′), which incorporated an upstream *Bgl*II site, and the lower primer W71L-1 (5′-**TTT AGC ACC TCC TCC** AAG GAA CAG ATG GTG GCG TC-3′) were used for the first round, and W71U and W71L-2 (5′-AAT GAA GAT CT**T TAC AAT TTA GCA CCT CCT CC**A AGG-3′) were used for the second round. The two lower primers incorporated three glycines, the peroxisomal targeting signal type I (PTS1), alanine-lysine-leucine, and the stop codon (all shown in *boldface*). Another *Bgl*II site was also added downstream of the stop codon. The amplified fragment was then cloned into the unique *Bgl*II site of the pRS316-PGK to yield pDsRed-AKL, which encodes a red fluorescence protein that targets well to peroxisomes.

Indirect Immunofluorescence, Fluorescence, and Confocal Microscopy

Immunofluorescence of *S. cerevisiae* was performed as described previously (23). Fluorescence in both living and fixed cells was visualized with a Zeiss Axioplan fluorescence microscope (Carl Zeiss, Inc., Thornwood, NY) and an MRC 600 confocal imaging system (Bio-Rad Laboratories, Hercules, CA). The GFP was excited at 488 nm, and the DsRed and the Texas Red fluorophore were excited at 568 nm. The images shown in the figures are single representative Z-sections through cells that were fixed, converted to spheroplasts, and permeabilized (23) to eliminate vacuolar autofluorescence.

Quantification of Peroxisomal Localization—Yeast transformants were cultured in SGd (synthetic medium containing 3% glycerol and 0.1% dextrose) and induced with oleic acid as described above. Living cells were harvested, spread in a monolayer on microscope slides, and observed by confocal microscopy. Five randomly chosen fields of at least 40 cells per field were examined for each culture. For each field a 19-section Z-series, 0.3 μm between optical slices, was recorded. The Z-series recordings were animated by IMAGE (version 1.60, National Institutes of Health). The percentage of the cells showing punctate fluorescence was determined and normalized against the positive and negative controls (Pmp47-(1–267)-GFP and untransformed MMY011α, respectively) in each experiment. The raw percentages of cells showing punctate fluorescence in the controls were typically 70% and 3%, respectively. The normalized value was termed the localization score, such that scores of 100 and 0 correspond to the positive and negative controls. A mean and standard error were determined from the localization scores of the five Z-series for each fusion protein. For the cluster modifications KIKQQ, KIQQQ, QIQQQ, AAKKA, AAAKA, and AAAAA, the data represent the average and standard error of ten Z-series from two independent experiments.

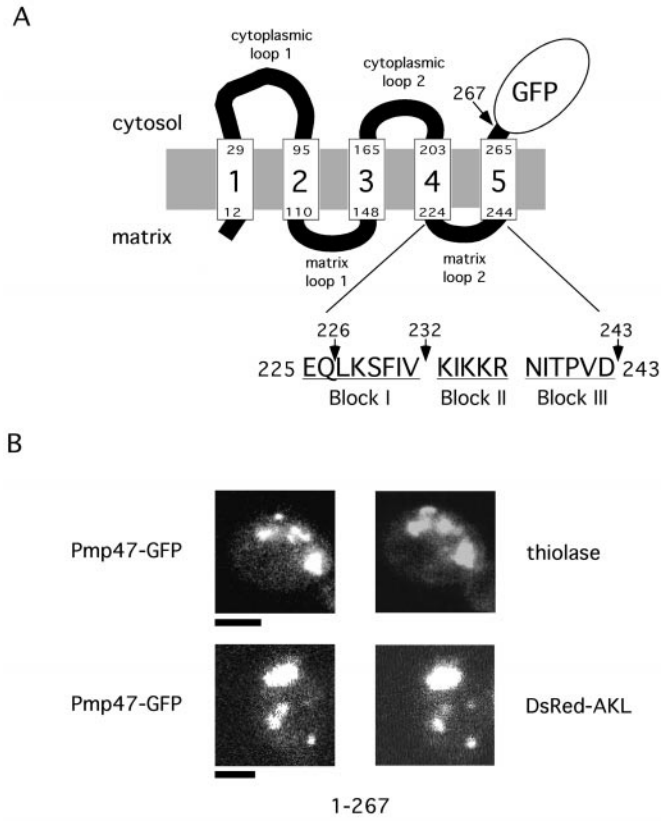

FIG. 1. **Pmp47-(1–267)-GFP sorts to oleate-induced peroxisomes in *S. cerevisiae*.** *A*, a topological map of Pmp47-(1–267)-GFP. *White rectangles* denote the first five (of six) TMDs and contain the numbers of their first and last amino acids. The amino acids of matrix loop 2 are shown, divided into three blocks so that block II consists of the basic cluster. The fusion junction is indicated by an *arrow*. *B*, representative pairs of fluorescence images of single cells that were transformed with Pmp47-(1–267)-GFP (here labeled *1–267*) or doubly transformed with Pmp47-(1–267)-GFP and DsRed-AKL and cultured in oleic acid to induce peroxisomes. Cells were prepared for indirect immunofluorescence to detect thiolase (*first row*) or permeabilized to remove vacuolar autofluorescence (*second row* and subsequent figures). In both sets GFP was detected by fluorescence. *Scale bar*, 2 μm.

RESULTS

A Quantitative Assay for Peroxisomal Localization—Fig. 1*A* illustrates the Pmp47-GFP fusion protein used as the basis of deletion and mutation analysis in this study. The Pmp47 section consists of the first 267 of a total of 423 amino acids of the full-length protein. This region comprises the amino-terminal matrix fragment, the first two cytosolic and two matrix-oriented loops, and the first five of the six transmembrane domains (TMDs). The sequence of matrix loop 2, containing the KIKKR basic cluster previously shown to be important for targeting, is also shown. Amino acids 1–267 have been shown previously to be sufficient for efficient targeting to peroxisomes in *S. cerevisiae* (9, 24).

Correct targeting of Pmp47-(1–267)-GFP to peroxisomes was evident either with 3-ketoacyl-CoA thiolase, a PTS2 protein used in previous studies (24), or DsRed-AKL, a PTS1-tagged red fluorescence protein (Fig. 1*B*), as observed by confocal microscopy of cells grown in oleic acid to induce peroxisomal proliferation. Peroxisomes of most cells were clustered by Pmp47-(1–267)-GFP, as apparent in the figure. The clustering was dependent on an intact Pmp47 amino terminus (data not shown) and is of unknown physiological relevance.

In preliminary experiments with several Pmp47-GFP fusion proteins we observed negligible mistargeting to organelles other than peroxisomes in living cells. If any fluorescence was seen, it was punctate and peroxisomal (*i.e.* colocalized with peroxisomal markers), suggesting that mistargeted Pmp47 fusion proteins were unstable in general; the few exceptions will be noted below. We also observed a correlation between the intensity of the fluorescence pattern and the percentage of cells displaying punctate fluorescence. Fusion proteins with relatively intense punctate fluorescence displayed this pattern in about 70% of cells. Weakly fluorescent punctate patterns were seen in 10–20% of cells (the remainder displayed no visible GFP fluorescence). We therefore developed a "localization score" as an indicator of peroxisomal targeting and retention of Pmp47-GFP fusion proteins. The localization score is based on the percentage of the cells displaying a peroxisome-like punctate pattern normalized to positive and negative controls (see "Experimental Procedures" for details). We separately verified the identity of the green fluorescent organelles of key constructs by assaying colocalization with a peroxisomal marker. The localization score discriminates between fusion proteins that are efficiently sorted to and retained in peroxisomes and those that are deficient in stability or targeting.

TABLE I
Dissecting the second matrix-oriented loop

GFP fusion protein	Localization score	Peroxisomal localization[a]
C-terminal deletions		
1–267	100 ± 4	Complete
1–243 (ΔTMD5)	90 ± 6	Complete
1–232 (ΔTMD5 and blocks II and III)	No fluorescence	
Block replacement		
Block I → GGGGAGGG	90 ± 5	NA
Block II → AAAAA	18 ± 2	Partial
Block III → GAAGGG	108 ± 4	NA
Block deletions		
1–243-GGG	17 ± 4	Complete
Δblock I in 1–243-GGG	25 ± 6	Complete
Δblock III in 1–243-GGG	56 ± 4	Complete
Isoleucine positions		
KIKKR-(1–267)	100 ± 4	Complete
IKKKR	95 ± 6	NA
KKIKR	102 ± 1	NA
RKKIK	96 ± 2	NA
Basicity titrations		
KIKKR-(1–267)	100 ± 4	Complete
KIKKQ	87 ± 4	Complete
KIKQQ	70 ± 7	Complete
KIQQQ	22 ± 2	Complete
QIQQQ	33 ± 6	Complete
AAKKR	60 ± 4	Complete
AAAKR	68 ± 10	Complete
AAKKA	26 ± 3	Complete/partial[b]
AAAKA	8 ± 3	Complete
AAAAR	27 ± 3	Complete
AAAAA (block II → AAAAA)	18 ± 2	Partial

[a] NA, not analyzed; complete, every GFP fluorescent spot colocalizes with peroxisomal marker; partial, some GFP fluorescent spots do not colocalize with peroxisomal marker.
[b] Most cells displayed complete colocalization.

Basicity, Position, and Side-chain Structure Are Important Elements of the Matrix Basic Cluster—An initial set of GFP fusion proteins were expressed to confirm the importance of the basic cluster contained within matrix loop 2 and to understand the important elements of the cluster (Table I). Both the localization scores and results of colocalization assays with thiolase or DsRed-AKL are shown. Removal of TMD5 had little effect on localization, indicating that it contained no essential targeting information. Further removal of blocks II (the basic cluster)

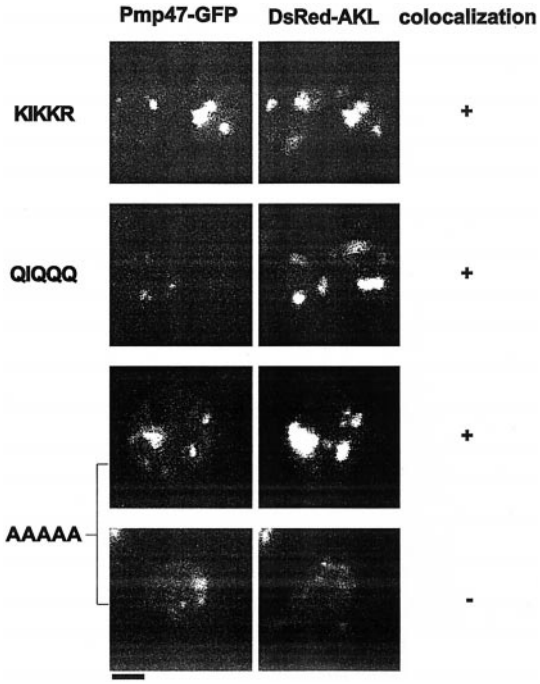

FIG. 2. **Residual peroxisomal targeting occurs even in the absence of a basic cluster.** Representative pairs of fluorescence images from cells doubly transformed with block II replacement mutants (in the context of 1–267) and DsRed-AKL. Colocalization is assigned when all the GFP stained structures are also stained with DsRed-AKL. Pmp47-(1–267)-GFP is denoted as *KIKKR* here. *Scale bar*, 2 μm.

and III of the second matrix loop completely obliterated localization, suggesting that this region contains important targeting information. This conclusion was supported first by replacement of each block of the loop with glycines or alanines. Only replacement of block II resulted in a low localization score. The colocalization assay revealed that, although correct targeting to peroxisomes occurred in some of the few cells that showed punctate fluorescence in the block II substitution, many of the green fluorescent organelles were not peroxisomes (Fig. 2), yielding partial colocalization. Finally, the importance of spacing of block II to the membrane was assessed with block deletions. Removal of block I resulted in localization score of 25. This may be an underestimation of the ability of the basic cluster to function close to a membrane span, because the matched control, Pmp47-(1–243-GGG)-GFP, had a localization score of only 17. The low score may reflect destabilization of the fusion protein by the triglycine spacer. Regardless, the basic cluster is functional to some extent when placed directly behind TMD4 or close to GFP (score of 56). Despite the low localization scores, these fusion proteins completely colocalize with peroxisomal markers (Table I).

We then designed GFP fusion proteins to understand the critical characteristics of the KIKKR cluster. First we altered the placement of the isoleucine residue within the cluster (in the context of Pmp47-(1–267)-GFP) and found no significant decrease in localization efficiency (Table I).

We then performed two titrations of basicity. In the first we substituted glutamines for basic residues beginning on the carboxyl side of the cluster. It can be seen that removal of two of the four basic residues (KIKQQ) still resulted in a localization score of 70 (Table I). However, the score dropped precipitously with one (KIQQQ) or no basic charge (QIQQQ) in the cluster, similar to the behavior of Pex3p (10). We were surprised that the QIQQQ-containing fusion protein still localized with a score of 33. The QIQQQ substitution completely colocalized with DsRed-AKL (Fig. 2).

Similar data were obtained with the second basic titration in which alanines replaced basic residues from the amino terminus of the cluster. AAAKR targeted with a score of 68, whereas the score dropped to 27 for AAAAR. Within the substituted clusters containing one or two basic charges, significant differences were seen: AAAKR localized much better than AAKKA (scores of 68 and 26, respectively), and AAAAR localized better than AAAKA (27 and 8). These data suggest that not all basic residues within the cluster are equivalent. Furthermore, the difference in targeting between the two neutral substitutions QIQQQ and AAAAA both in localization score (scores of 33 and 18, respectively) and colocalization with DsRed-AKL (complete *versus* partial colocalization) suggests that side-chain structure is important in basic cluster targeting. More definitive conclusions must await an *in vitro* targeting assay where stability can be controlled and correct folding more easily measured.

TMD2 and Its Amino-terminal Adjacent Region Are Required for Peroxisomal Localization—We suspected that important targeting information was contained in the first half of Pmp47, because targeting was difficult to demonstrate without it (24). To localize the putative targeting information a series of amino-terminal deletions of Pmp47-(1–267)-GFP were constructed (Table II). TMD1 was dispensable as well as most of the first cytosolic loop, because Pmp47-(70–267)-GFP targeted well. Excellent colocalization of a fusion protein beginning at amino acid 75, 20 residues from TMD2, can be seen in Fig. 3. A fusion protein that lacked the first 78 amino acids did not display any peroxisomal fluorescence. However, peroxisomal localization reappeared in a fusion protein that lacked the first 86 amino acids, although the localization score dropped to 30 and mistargeting to other organelles was observed (Fig. 3), suggesting that the distal region of the first cytoplasmic loop may be important for peroxisomal localization. Targeting to peroxisomes in oleate-grown cells was never observed in fusion proteins with amino termini beyond TMD2. Unlike most of the Pmp47-GFP proteins discussed earlier, a few of these fusion proteins stably mislocalized to nonperoxisomal structures of the oleate-grown cells, either punctate or reticular (Table II). Our results suggest that targeting information is contained within TMD2 or amino acids directly adjacent to it.

To further test whether TMD2 or the other remaining membrane-spanning domains have essential targeting information, we substituted these TMDs with spans from membrane proteins of other organelles. Membrane spans from Sec61p (a membrane protein of the endoplasmic reticulum) or a mitochondrial solute transporter were able to substitute for TMD4 or TMD3, respectively, indicating that neither Pmp47 span contained essential targeting information (Table III). However, we were unable to demonstrate peroxisomal targeting when Pmp47 TMD2 was substituted with the homologous TMDs from two different mitochondrial solute transporters; they have similar hydrophobicity to the Pmp47 TMD2. The substituted fusion proteins targeted instead to an unknown punctate organelle (Fig. 3). These experiments support a role of TMD2 in peroxisomal targeting.

Three Small Regions within Pmp47 Are Sufficient for Effective Targeting—Experiments described so far suggest that important peroxisomal membrane targeting information is contained in the basic cluster within matrix loop 2 as well as in TMD2 and an adjacent region of cytosolic loop 1. To determine whether these domains are sufficient for peroxisomal targeting, amino acids 70–110 (carboxyl end of cytoplasmic loop 1 and TMD2) were fused to matrix loop 2. Hydrophobicity analysis indicates that TMD2 is a weak transmembrane span, and previous data show that TMD2 is not sufficient to anchor the protein onto the membrane bilayer (24); therefore, we added

TABLE II
Targeting of amino-terminal truncation mutants to oleate-induced peroxisomes

Amino terminus of Pmp47-(*x*-267)-GFP	Localization score or fluorescence intensity	Localization
1	100 ± 6	Peroxisomes
Loop 1		
39	120 ± 5	Peroxisomes
53	114 ± 5	Peroxisomes
70	106 ± 4	Peroxisomes
75	74 ± 4	Peroxisomes
79	None	
87	30 ± 3	Peroxisomes + non-peroxisomal punctate structures
Loop 2		
114	None	
125	Strong	Non-peroxisomal punctate structures
141	Medium	Non-peroxisomal punctate structures
148	None	
Loop 3		
168	None	
176	None	
187	None	
190	None	
196	Strong	Reticular structures
199	Strong	Reticular structures
203	None	
Loop 4		
226	None	

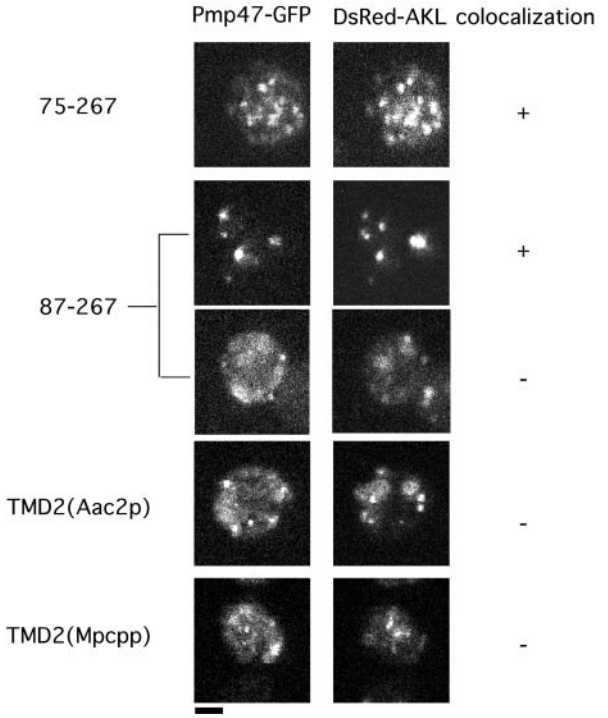

FIG. 3. **The importance of amino acids 70–110 for localization to oleate-induced peroxisomes.** Colocalization of the 75–267 GFP fusion is excellent, although its localization score is lower. Colocalization of 87–267 is much weaker but still can be seen. Replacement of TMD2 yields punctate patterns that do not correspond to peroxisomes. Shown are representative pairs of fluorescence images from cells doubly transformed with Pmp47-GFP fusions and DsRed-AKL. *75–267*, Pmp47-(75–267)-GFP; *87–267*, Pmp47-(87–267)-GFP; *TMD2(Aac2p)* and *TMD2(Mpcpp)*, Pmp47-(53–267)-GFP with the TMD2 replaced with the TMD2s from *Sc*Aac2p and *Sc*Mpcpp, respectively (see Table III). *Scale bar*, 2 μm.

TMD5, a strongly hydrophobic domain without essential targeting information, to allow membrane integration of the fusion protein. This fusion protein targeted to peroxisomes with a localization score of 48 and yielded a perfect colocalization pattern with DsRed-AKL, indicating that these small regions of Pmp47 were sufficient for targeting (Fig. 4).

The first matrix loop contains three pairs of basic residues at amino acids 123–124, 134–135, and 139–140. Because two basic residues within matrix loop 2 could replace the basic cluster, we asked whether the first matrix loop could substitute for the second one. The data in Fig. 4 show that both fusion proteins had similar localization scores and both colocalized well with DsRed-AKL. Therefore, these loops have redundant targeting information. However, TMD2 appears to be essential; substitution of TMD4 for TMD2 resulted in loss of fluorescence, suggesting an inability to target to peroxisomes (Fig. 4).

In addition to the demonstrated importance of matrix-oriented basic clusters in targeting to the peroxisomal membrane, a short cytosol-oriented region adjacent to TMD1 of rat Pmp22 has been shown to contain a peroxisomal targeting signal (14). Similarly, data suggest that a cytosolic region adjacent to the membrane span of Pex15p is required for peroxisomal localization (11). We asked whether either region could substitute for amino acids 70–94 (the region in cytoplasmic loop 1 adjacent to TMD2) of Pmp47. No peroxisomal targeting was seen with either substitution in yeast grown on oleic acid. However, we found that the Pex15p substitution yielded strong fluorescence and excellent colocalization in cells grown on glucose, a substrate that represses peroxisomal proliferation, although a few small peroxisomes remain (Fig. 5). Colocalization was also seen with the Pmp22 substitution under the same growth conditions, although fluorescence was weaker. In contrast, a cytosol-oriented region adjacent to the TMD1 of Sec61p yielded no fluorescence in cells grown in either medium. Although the effect of carbon source on peroxisomal membrane targeting is unclear at present, these data suggest that the cytoplasmic regions in Pmp47, Pex15p, and Pmp22 share a common targeting element. Comparison of the sequences (Fig. 5) failed to identify apparent sequence similarity, although all three sequences have amphipathic characteristics and net positive charge.

TABLE III
Targeting of the TMD substitution mutants to oleate-induced peroxisomes

Pmp47 TMD (GFP fusion)	Pmp47 TMD sequence and position	Replacement TMD position and sequence	Localization	Localization score
TMD 2 (53–267)	GLESALFGIAVTNFVY (95–110)	*Sc*Aac2p TMD2 (89–107) GNTANVIRYFPTQALNFAF	NPPS[a]	
		*Sc*Mpcpp TMD2 (75–91) GFGPTLLGYSIQGAFKF	NPPS	
TMD 3 (70–267)	MAAGAVAGTISRVATN (148–163)	*Sc*Aac2p TMD3 (131–146) LASGGAAGALSLLFVY	P	108 ± 3
TMD 4 (1–267)	LFTGIVPALFLVLNPIIQYTIF (203–224)	*Sc*Sec61p TMD1 (33–55) LIWTGVSLLIFLILGQIPLYGI	P	91 ± 3
		*Sc*Sec61p TMD2 (76–95) TLLELGVSPIITSSMIFQFL	P	108 ± 4
		*Sc*Sec61p TMD3 (120–140) VCAIILILGQALVVVMTGNY	P	84 ± 3

[a] NPPS, non-peroxisomal punctate structures; P, peroxisomes.

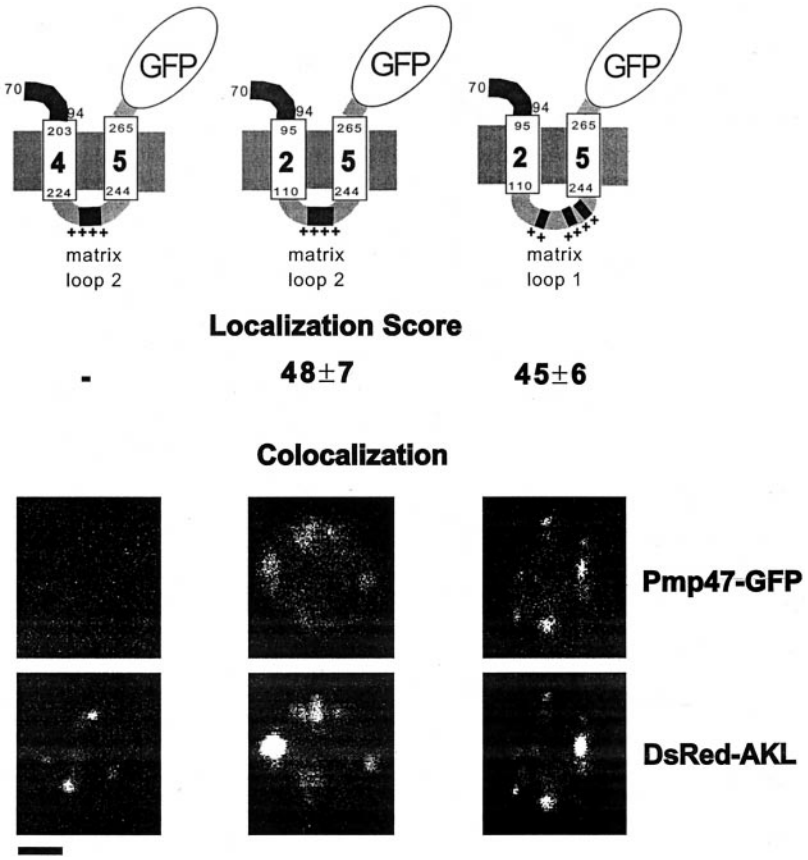

FIG. 4. **Amino acids 70–110, a matrix-facing basic cluster, and an anchoring TMD are sufficient for peroxisomal localization.** *Top*, diagram of the three minimal Pmp47-GFP fusion proteins and their localization scores; *bottom*, representative pairs of fluorescence images from cells doubly transformed with DsRed-AKL and each of the three Pmp47-GFP fusion constructs. *Scale bar*, 2 μm.

DISCUSSION

In this work we have analyzed some important characteristics of the Pmp47 basic cluster and have identified a second area consisting of a transmembrane domain and adjacent cytoplasmic sequence within Pmp47 that is important for targeting. A quantitative localization assay allows us to estimate that a minimal GFP fusion protein consisting of little more than these regions and an anchoring membrane span localizes to peroxisomes with about 50% efficiency compared with wild type protein and with full fidelity (*i.e.* no detectable mistargeting).

Judging Localization in Vivo—Results from our GFP assay correlated well with previous data derived from immunofluorescence and cell fractionation studies (9, 24). Thus, matrix loop 2 was found to be essential for targeting, and substitution of amino acids in block II had by far the most severe effect compared with changes in the other two blocks.

We measured both localization efficiency (approximated by the strength of the fluorescence signal and reflected in the localization score) and fidelity (the degree of correspondence of green fluorescent spots with those from the peroxisomal marker). The use of two distinct but complementary assays allowed us to dissect the basic cluster and analyze other areas of Pmp47 for targeting signals with confidence. The behavior of several Pmp47-GFP fusion proteins that targeted to nonperoxisomal punctate structures in oleate-cultured cells serves to emphasize the importance of a good colocalization assay.

There are limitations inherent in our *in vivo* system, however. It is possible that the rates of synthesis among the fusion proteins are different, although such difference should be minor because the same promoter is used for all fusion proteins. Another potential limitation is that the dynamic range may not fully reflect the extent of targeting; the assay might reach a maximum value at low actual targeting efficiency. However, fluorescence of our positive control (Pmp47-(1–267)-GFP) is very dim relative to that of a GFP peroxisomal matrix marker and we can detect fluorescence in only 70% of cells with this fusion protein, whereas GFP-AKL is detected in over 90% of

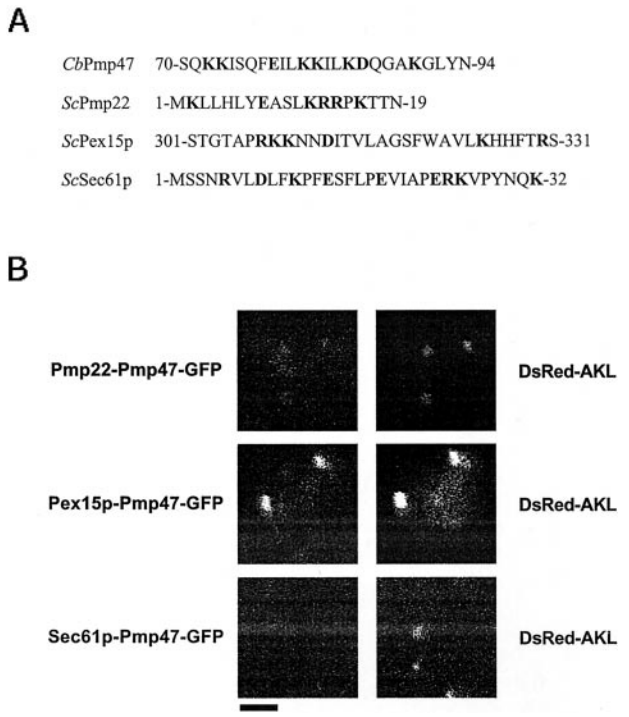

FIG. 5. **Amino acids 70–94 contain peroxisomal targeting information.** *A*, the sequence of Pmp47 residues 70–94 and those from cytoplasmic regions of two peroxisomal membrane proteins containing putative targeting information, *Sc*Pmp22 and *Sc*Pex15p, and one ER membrane protein, *Sc*Sec61p, respectively, are shown. Charged residues are in *boldface* text. All four sequences immediately flank a TMD. *B*, representative pairs of fluorescence images from cells doubly transformed with DsRed-AKL and domain exchange mutants and cultured in glucose-containing medium. *Pmp22-Pmp47*, *Sc*Pmp22-(1–19)-Pmp47-(95–267)-GFP; *Pex15p-Pmp47*, *Sc*Pex15p-(301–331)-Pmp47-(95–267)-GFP; *Sec61p-Pmp47*, *Sc*Sec61p-(1–32)-Pmp47-(95–267)-GFP. *Scale bar*, 2 μm.

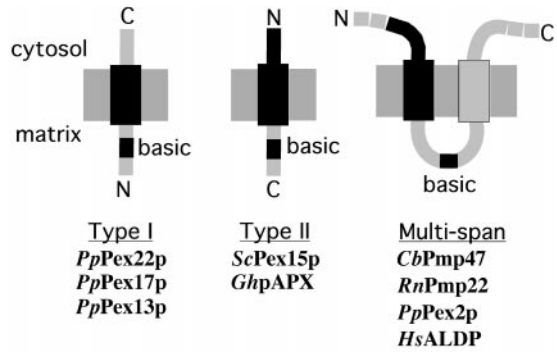

FIG. 6. **Peroxisomal targeting signals may depend on protein orientation.** The peroxisomal targeting signals for three different types of membrane proteins (type I, type II, and multispan transmembrane proteins, respectively) are shown. According to this model, all of them contain an important basic cluster facing the matrix and membrane span containing targeting information. Type II proteins and multispanning proteins may require additional information on the cytoplasmic side of the membrane. Targeting information is indicated in *black*. Listed are proteins with known topology and identified targeting regions. See text for further discussion. References: *Pp*Pex22p, *Pp*Pex17p, *Pp*Pex13p, *Pp*Pex2p (30); *Sc*Pex15p (11); *Gh*pAPX (13); *Cb*-Pmp47, this study; *Rn*Pmp22 (14); *Hs*ALDP (26).

cells (data not shown), suggesting that our assay for peroxisomal membrane targeting does not saturate. However, targeting of fusion proteins with very weak signals may be missed by our assay, although it is probably detected by organelle fractionation and sensitive immunoblotting.

A more substantive caveat of the assay is that aberrant protein folding or low stability, rather than a weak targeting signal, would be reflected as a low localization score. This is a problem with most *in vivo* targeting assays as well as our own. Therefore, we have interpreted low localization scores with caution. The strength of the assay lies in identification of fusion proteins with high sorting scores, indicating that they must contain strong targeting signals.

The Role of the Basic Cluster—The mPTSs of several peroxisomal membrane proteins contain basic clusters (10–13). Our data show that basicity is an important functional attribute of the cluster. However, the loss of two of the four positive charges has only a minor effect. This is in agreement with studies of a basic cluster that directs Pex3p to peroxisomes (10). The fact that both KIKQQ and AAAKR localize well (scores of 70 and 68, respectively) indicates that no particular specific residue within KIKKR is required. But not all substitutions with similar total charge behave similarly: AAKKA localizes much worse that AAAKR (scores of 26 and 68, respectively). Likewise, AAAAR localizes worse than AAAKA (scores of 27 and 8, respectively). This suggests that the position of the residues, in addition to the total charge of the cluster, governs targeting in some way. The ability of glutamines to substitute for all basic residues and sort better than a construct with alanines (both higher localization score and better colocalization with thiolase and DsRed-AKL) further suggests that the side chains of residues in the basic cluster contribute significantly to the binding to a receptor. This is also the case for the binding of the classical NLS to importin α, where the binding is stabilized by the methylene bridges of the basic residues (25).

Our data indicate that the functional equivalent of the basic cluster in other peroxisomal membrane proteins need not contain more than two basic charges, or may function with even fewer in the proper context. Thus, the first matrix loop of Pmp47, containing three pairs of basic residues, can substitute for the second one in our minimal construct without significant loss of targeting. At least one matrix domain with a dibasic sequence or multiple basic residues in a cluster is present in all peroxisomal integral membrane proteins with known topology (11–14, 26–30). These all may serve a similar targeting role.

However, the spacing between these basic clusters and their preceding TMDs are quite different. The fact that deletion of block I in Pmp47-(1–243)-GGG-GFP does not obliterate targeting suggests that this spacing is not critical. These data all support the hypothesis that the matrix-facing basic cluster serves as a targeting motif.

Additional Amino-terminal Signals Are Required for Sorting—A number of experiments motivated us to search for targeting signals upstream of matrix loop 2. First, residual sorting was observed even when the basic cluster KIKKR was replaced with AAAAA, which suggested that additional signals existed elsewhere. Second, fusion proteins missing large regions of the amino-terminal half of Pmp47 have been unstable and unable to target (24). Third, although we had previously shown that a fusion protein consisting of only GFP and matrix loop 2 comigrated on a Nycodenz gradient with peroxisomes, the extent of targeting was poor and could not be independently verified by fluorescence microscopy (9). Finally, Pmp47-GFP fusion proteins containing the second matrix loop flanked only with TMD4, TMD5, or both (Pmp47-(176–243)-GFP, Pmp47-(226–267)-GFP, and Pmp47-(203–267)-GFP, respectively) do not show peroxisomal localization at all (Table II and data not shown).

Our analysis indicates that peroxisomal localization is affected when more than 69 amino acids are deleted from the amino terminus, implicating the end of the first cytoplasmic loop in peroxisomal targeting. Our substitution experiments with cytoplasmic regions of Pmp22 and Pex15p suggest that

these regions may fulfill similar functions. A conserved motif, $YX_3LX_3PX_3$(K/Q/N), was identified in proteins in the Pmp22 family that may serve a targeting function (14). This motif, however, is not found in the Pmp47 cytoplasmic loop, suggesting that the shared targeting signal may be related to secondary *versus* primary sequence structure. Our experiments show that Pmp22 and Pex15p cytoplasmic sequences were only able to complement in cells grown in glucose, not oleate. This may reflect a quantitative or qualitative difference in targeting to basal *versus* induced peroxisomes. There may be enhanced cytosolic or matrix proteolysis in oleate-grown cells, such that poorly folded proteins are more likely to be degraded before or after arrival at peroxisomes. Perhaps the signals on Pmp47 form internal interactions that are not optimally reproduced with the Pex15p or Pmp22 substitutions. Another possibility is that there is an additional targeting step involved requiring a hierarchy of signals to reach induced peroxisomes, perhaps related to the vesicular intermediates identified in *Yarrowia lipolytica* (31), and the mixed Pex15p-Pmp47 or Pmp22-Pmp47 signals are not sufficiently strong to allow passage through this critical step.

TMD2, adjacent to the important region of cytoplasmic loop 1, is necessary for the targeting of Pmp47. The evidence supporting the importance of TMD2 in targeting stems by its necessity for targeting and the inability to be substituted with either TMD4 or spans of analogous mitochondrial carrier proteins. TMD2 has a low hydrophobicity score, in contrast to TMD4 and TMD5 (32). Fusion proteins lacking TMD4 and TMD5 are soluble in the cell (24). Therefore, we added TMD5 to anchor our minimal fusion protein to the membrane. Stabilization of TMD2 by TMD5 within the membrane may occur through an ionic interaction between E97 in TMD2 and K252 in TMD5. We detect no sequence similarity of Pmp47 TMD2 and transmembrane domains of other peroxisomal membrane proteins. The TMD2 signal may be more related to its secondary structure, or it may depend on interactions with the adjacent region of the cytoplasmic loop during targeting. It would be interesting to assess whether these two adjacent regions can be separated while maintaining targeting function.

Membrane Orientation May Affect Requirements for Targeting Signals—There are conflicting reports about the nature of the signal that directs proteins to the peroxisomal membrane. Although matrix-oriented basic clusters have been shown to be important for many peroxisomal membrane proteins, the critical signal on Pmp22 appears to be a cytoplasmic sequence adjacent to a TMD. All membrane proteins require a TMD for membrane interaction, yet a synthetic TMD might function to some extent (13).

Do these data imply several different targeting pathways and receptor systems for peroxisomal membrane proteins or can the observations described above be reconciled into a simpler model? We propose that targeting signals depend in part on the topology of proteins across the peroxisomal membrane (Fig. 6). According to this model type I membrane proteins require a matrix-oriented basic cluster and transmembrane span for targeting and integration. We hypothesize that the transmembrane span has specific targeting information as well as a membrane-anchoring role, because it seems improbable that a basic cluster has sufficient information for targeting by itself. A test of this hypothesis will require a demonstration of peroxisomal colocalization of a single span protein with a synthetic TMD substitution.

In contrast to type I proteins, our model predicts that type II and multispanning proteins require additional information within cytoplasmic-oriented sequences. This may reflect a more complex pathway and different kinetics of membrane insertion compared with those of type I proteins. This is reasonable, because the targeting signal of type I proteins would become available to the receptor machinery sooner during translation. Proteins in the other two classes may require additional chaperones or receptor components that bind earlier.

The model stipulates that the functionality of the matrix basic cluster is conserved in all peroxisomal integral membrane proteins. Although removal of the basic cluster in loop 1 of Pmp22 has no effect on targeting, a minimally sufficient fragment retains the sequence KMR from the second matrix loop (14), which may suffice for a basic cluster, suggesting a redundant function. Redundancy in targeting signals on multispanning peroxisomal membrane proteins is also shown in our substitution of matrix loop 1 for matrix loop 2 in our minimal fusion proteins. Both Pmp47 minimal proteins have similar localization scores. The score for Pmp47-(1–267)-GFP is double that of the minimal proteins, which may reflect the presence of both basic matrix loops, other redundant signals yet undetected, or a more stable protein structure in the membrane.

Beyond our model, there may be another level of complexity of targeting signals, if peroxisomal membrane proteins can enter a series of vesicular precursors at different points in the pathway (31). Determining the signals that direct proteins to specific peroxisomal organelle precursors awaits a more complete understanding of this novel pathway.

The identification of discrete sorting regions on peroxisomal membrane proteins is an important step to understanding the mechanism of targeting and membrane integration. Although Pex19p has been implicated as a receptor for peroxisomal membrane proteins (33), absence of overlap between binding and targeting domains on several peroxisomal membrane proteins suggest that other factors must also be involved (30). We hope the identification of the important targeting elements within Pmp47 will lead to dissection of the precise function of each signal in peroxisomal membrane assembly.

Acknowledgments—We are grateful for the excellent technical assistance of Shary Shelton. We thank Michael Roth for much help with using his confocal microscope and Koji Okamoto for advice with immunofluorescence analysis. We thank Kate Luby-Phelps for helpful suggestions on quantification of sorting. Thanks to Mary Stewart, Johnathan Ballard, Pamela Marshall, and Mark Lehrman for suggestions on this manuscript. Thanks to Richard Baerends for sharing his results before publication.

REFERENCES

1. Subramani, S. (1993) *Annu. Rev. Cell Biol.* **9,** 445–478
2. Hettema, E. H., Distel, B., and Tabak, H. F. (1999) *Biochim. Biophys. Acta* **1451,** 17–34
3. Subramani, S. (1998) *Physiol. Rev.* **78,** 171–188
4. van der Klei, I. J., Hilbrands, R. E., Kiel, J. A., Rasmussen, S. W., Cregg, J. M., and Veenhuis, M. (1998) *EMBO J.* **17,** 3608–3618
5. Zhang, J. W., and Lazarow, P. B. (1996) *J. Cell Biol.* **132,** 325–334
6. Jank, B., Habermann, B., Schweyen, R. J., and Link, T. A. (1993) *Trends Biochem. Sci.* **18,** 427–428
7. Nakagawa, T., Imanaka, T., Morita, M., Ishiguro, K., Yurimoto, H., Yamashita, A., Kato, N., and Sakai, Y. (2000) *J. Biol. Chem.* **275,** 3455–3461
8. Sakai, Y., Saiganji, A., Yurimoto, H., Takabe, K., Saiki, H., and Kato, N. (1996) *J. Cell Biol.* **134,** 37–51
9. Dyer, J., McNew, J., and Goodman, J. M. (1996) *J. Cell Biol.* **133,** 269–280
10. Baerends, R. J. S., Faber, K. N., Kram, A. M., Kiel, J., van der Klei, I. J., and Veenhuis, M. (2000) *J. Biol. Chem.* **275,** 9986–9995
11. Elgersma, Y., Kwast, L., van den Berg, M., Snyder, W. B., Distel, B., Subramani, S., and Tabak, H. F. (1997) *EMBO J.* **16,** 7326–7341
12. Koller, A., Snyder, W. B., Faber, K. N., Wenzel, T. J., Rangell, L., Keller, G. A., and Subramani, S. (1999) *J. Cell Biol.* **146,** 99–112
13. Mullen, R. T., and Trelease, R. N. (2000) *J. Biol. Chem.* **275,** 16337–16344
14. Pause, B., Saffrich, R., Hunziker, A., Ansorge, W., and Just, W. W. (2000) *FEBS Lett.* **471,** 23–28
15. Titorenko, V. I., and Rachubinski, R. A. (1998) *Mol. Cell. Biol.* **18,** 2789–2803
16. McCammon, M. T., Dowds, C. A., Orth, K., Moomaw, C. R., Slaughter, C. A., and Goodman, J. M. (1990) *J. Biol. Chem.* **265,** 20098–20105
17. Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) *J. Bacteriol.* **153,** 163–168
18. van Dijken, J. P., Otto, R., and Harder, W. (1976) *Arch. Microbiol.* **111,** 137–144
19. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A*

Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
20. Sikorski, R. S., and Hieter, P. (1989) *Genetics* **122,** 19–27
21. Kadowaki, H., Kadowaki, T., Wondisford, F. E., and Taylor, S. I. (1989) *Gene* **76,** 161–166
22. Vallette, F., Mege, E., Reiss, A., and Adesnik, M. (1989) *Nucleic Acids Res.* **17,** 723–733
23. Pringle, J. R., Adams, A. E. M., Drubin, D. G., and Haarer, B. K. (1991) in *Guide to Yeast Genetics and Molecular Biology* (Guthrie, C., and Fink, G. R., eds) Vol. 194, pp. 586–597, Academic Press, Inc., San Diego
24. McCammon, M. T., McNew, J. A., Willy, P. J., and Goodman, J. M. (1994) *J. Cell Biol.* **124,** 915–925
25. Fontes, M. R., Teh, T., and Kobe, B. (2000) *J. Mol. Biol.* **297,** 1183–1194
26. Gloeckner, C. J., Mayerhofer, P. U., Landgraf, P., Muntau, A. C., Holzinger, A., Gerber, J. K., Kammerer, S., Adamski, J., and Roscher, A. A. (2000) *Biochem. Biophys. Res. Commun.* **271,** 144–150
27. Gould, S. J., Kalish, J. E., Morrell, J. C., Bjorkman, J., Urquhart, A. J., and Crane, D. I. (1996) *J. Cell Biol.* **135,** 85–95
28. Kamijo, K., Taketani, S., Yokota, S., Osumi, T., and Hashimoto, T. (1990) *J. Biol. Chem.* **265,** 4534–4540
29. Snyder, W. B., Koller, A., Choy, A. J., Johnson, M. A., Cregg, J. M., Rangell, L., Keller, G. A., and Subramani, S. (1999) *Mol. Biol. Cell* **10,** 4005–4019
30. Snyder, W. B., Koller, A., Choy, A. J., and Subramani, S. (2000) *J. Cell Biol.* **149,** 1171–1178
31. Titorenko, V. I., Chan, H., and Rachubinski, R. A. (2000) *J. Cell Biol.* **148,** 29–44
32. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132
33. Sacksteder, K. A., Jones, J. M., South, S. T., Li, X., Liu, Y., and Gould, S. J. (2000) *J. Cell Biol.* **148,** 931–944