A Conserved Tripeptide Sorts Proteins to Peroxisomes

Stephen J. Gould, Gilbert-Andre Keller, Nancy Hosken, Jack Wilkinson, and Suresh Subramani

Department of Biology and Center for Molecular Genetics, University of California, La Jolla, California 92093

Abstract. The firefly luciferase protein contains a peroxisomal targeting signal at its extreme COOH terminus (Gould et al., 1987). Site-directed mutagenesis of the luciferase gene reveals that this peroxisomal targeting signal consists of the COOH-terminal three amino acids of the protein, serine-lysine-leucine. When this tripeptide is appended to the COOH terminus of a cytosolic protein (chloramphenicol acetyltransferase), it is sufficient to direct the fusion protein into peroxisomes. Additional mutagenesis experiments reveal that only a limited number of conservative changes can be made in this tripeptide targeting signal without abolishing its activity. These results indicate that peroxisomal protein import, unlike other types of transmembrane translocation, is dependent upon a conserved amino acid sequence.

In eukaryotic cells, the sorting of proteins to their proper subcellular locations is often mediated by *cis*-acting sequences present in the transported polypeptide (Blobel, 1980). This mode of protein segregation has been demonstrated for protein import into mitochondria (Hase et al., 1984), chloroplasts (Van den Broek et al., 1985), lysosomes (Johnson et al., 1987), the nucleus (Kalderon et al., 1984), and the endoplasmic reticulum (ER)[1]/secretion pathway (for review, see Walter and Lingappa, 1986) as well as for retention of proteins in the ER (Munro and Pelham, 1987). Only recently has this observation been extended to the process of peroxisomal protein import (Gould et al., 1987, 1988; Small et al., 1988). Proteins destined to reside in the peroxisome are transported into the organelle posttranslationally after synthesis on free polysomes. Unlike import of proteins into mitochondria, chloroplasts, or the ER/secretion pathway, peroxisomal import is not associated with removal of a presequence or with any other detectable modification of the imported polypeptide (for a review of peroxisomal protein biogenesis, see Lazarow and Fujiki, 1985).

We have previously demonstrated that the peroxisomal protein luciferase (Keller et al., 1987) contains a peroxisomal targeting signal (PTS) within its COOH-terminal twelve amino acids (Gould et al., 1987). More recently, we identified additional PTSs at or near the COOH terminus of the peroxisomal proteins human catalase, pig D-amino acid oxidase, rat acylcoenzyme A (CoA) oxidase, and rat enoyl-CoA/hydratase-3-hydroxyacyl CoA dehydrogenase (Gould et al., 1988), and also in the *Candida boidinii* PMP-20 protein (Gould et al., manuscript in preparation). Miyazawa et al. (1989) have recently demonstrated that the COOH terminus of rat acyl-CoA oxidase, which had been previously shown to contain a PTS (Gould et al., 1988), is also necessary for its own import into rat liver peroxisomes in vitro.

To further investigate the nature of the PTS, we have created mutations that alter the amino acid sequence of luciferase at or near the COOH terminus of the protein. The mutant proteins were tested for their ability to be imported into peroxisomes in vivo. The results of these experiments indicate that a tripeptide (serine, alanine, or cysteine at the first position; lysine, histidine, or arginine at the second; leucine at the third) is capable of acting as a PTS when present at the COOH terminus of a protein. A search of known peroxisomal protein sequences reveals that this tripeptide is a common feature of many peroxisomal proteins. These data suggest that a consensus three-amino-acid targeting signal is responsible for the import of a large number of peroxisomal proteins. In contrast, other processes that transport proteins across membranes (ER, mitochondrial, or chloroplast transport) are usually mediated through larger NH_2-terminal signals that are cleaved upon import and lack sequence conservation. The unique properties of peroxisomal protein import indicate that it may involve a fundamentally different type of protein translocation mechanism.

Materials and Methods

Reagents

The rabbit antibody against bovine catalase was a gift from A. Schram (Genentech Inc., San Francisco, CA). The chloramphenicol acetyltransferase (CAT) 2 cell line producing the anti-CAT mAb was a gift from C. Gorman (Genentech Inc., San Francisco, CA). All oligonucleotides were purchased from Operon Technologies, Inc. (San Pablo, CA). Sequenase and the plasmid pTZ19U were obtained from United States Biochemical Corp. (Cleveland, OH). Additional reagents were from standard sources.

Plasmids

Manipulations of plasmid DNAs were as described (Maniatis et al., 1982).

1. *Abbreviations used in this paper*: CAT, chloramphenicol acetyltransferase; CoA, coenzyme A; PTS, peroxisomal targeting signal.

DNA sequence analysis was performed by the method of Sanger et al. (1977), as modified by Chen and Seeburg (1985). In place of Klenow fragment of *Escherichia coli* DNA polymerase, a modified T7 polymerase (Sequenase; United States Biochemical Corp.) was used. Construction of the plasmid pRSVL has been described elsewhere (de Wet et al., 1987) as has construction of the plasmid pRSVLΔ C12 (Gould et al., 1987). The other plasmids used in this report (except pSV2CAT and pSV2CAT-SKL) are the same as pRSVL except for the alteration of specific nucleotides between the Cla I and Sac I sites of pRSVL. The 450-bp-long Eco RV to Sac I fragment of pRSVL was isolated and cloned between the Hinc II and Sac I sites of the plasmid pTZ19U (United States Biochemical Corp.). The Cla I site of the luciferase gene is within this fragment ~30 bp from the Eco RV site. The resultant plasmid pTZ-Luc was used as the substrate for site-specific mutagenesis by the method of Zoller and Smith (1983), as modified by Kunkel (1985). After the structure of each mutant was confirmed by DNA sequence analysis, the Cla I to Sac I segment of the mutated luciferase gene fragment was inserted back into pRSVL for subsequent expression in mammalian cells.

The construction of pSV2CAT has been described elsewhere (Gorman et al., 1982). pSV2CAT-SKL was constructed by inserting an oligonucleotide linker into the unique Nae I site of the plasmid pCATC (the structure of which is described in Gould et al., 1988). This oligonucleotide had the sequence 5'-TCAAGCTTTGA-3' and encoded a Hind III restriction enzyme recognition site. When inserted into the Nae I site of pCATC, the predicted amino acid sequence of CAT is altered in such a way that the tripeptide serine-lysine-leucine is appended to the COOH terminus of the protein. In all other respects, the sequence of the putative CAT-SKL protein is the same as the wild-type protein. A signal for mRNA polyadenylylation from SV40 was then placed downstream of the CAT-SKL coding sequence by inserting the approximately 1600-bp-long Kpn I to Eco RI fragment of pRSVL between the Kpn I and Eco RI sites of pCATC-SKL. The product of this construction was pSV2CAT-SKL.

Cell Culture, Transfections, and Immunofluorescence

Conditions for growth of monkey kidney CV-1 cells are the same as previously described (Keller et al., 1987). Transfections were performed by the method of Parker and Stark (1979). Immunofluorescence was performed essentially as described by Keller et al. (1987). In this procedure, fixed cells were first labeled with a guinea pig antibody that recognizes the firefly luciferase and a rabbit antibody raised against bovine catalase. After several washes, the cells were then labeled with a fluorescein-conjugated goat anti-guinea pig IgG antibody and a rhodamine-conjugated goat anti-rabbit IgG antibody. When observed by fluorescence microscopy, the rhodamine label depicted the distribution of catalase and the fluorescein label showed the distribution of luciferase. The goat anti-rabbit antibody does not cross-react with the guinea pig anti-luciferase antibody under the conditions we used. Cells expressing the CAT and CAT-SKL proteins were analyzed in the same manner except that a monoclonal anti-CAT antibody was used in place of the guinea pig antiluciferase antibody and that the secondary labelings were done with a fluorescein-conjugated goat anti-mouse IgG antibody. The culture medium in which the CAT-2 mAb-producing cell line was grown was used as the source of the anti-CAT antibody.

Results

The Minimal PTS in Luciferase

When the gene encoding firefly luciferase was expressed in mammalian, insect, plant, or yeast cells, luciferase was found to be translocated into the peroxisomes of the cells in all cases (Keller et al., 1987, and unpublished data). Because the PTS of this protein was recognized in such diverse organisms, we felt that luciferase would be an excellent model protein with which to determine the minimal PTS. Site-directed mutagenesis was used to alter the predicted amino acid sequence of the luciferase gene at the 3' end of its coding region, the region of the protein previously reported to contain targeting information (Gould et al., 1987). Specifically, deletions were made that removed amino acids either from the amino or carboxy end of the twelve amino acid–long COOH-terminal targeting signal. Plasmids directing the expression of mutant luciferase genes from the Rous Sarcoma virus promoter were introduced into CV-1 monkey kidney cells by calcium phosphate-mediated transfection (Parker and Stark, 1979). The cells were subsequently processed for double indirect immunofluorescence as previously described (Keller et al., 1987). A guinea pig antiluciferase antibody and fluorescein-conjugated goat anti-guinea pig IgG antibody were used to detect the subcellular distribution of the mutant luciferase proteins. The same cells were incubated with a rabbit anticatalase antibody and rhodamine conjugated goat anti-rabbit-IgG antibody to localize the peroxisomes within the transfected cells. Each mutant luciferase protein was identified as either peroxisomal or cytosolic based on whether or not it colocalized with catalase in the cell.

Table I. Mutants Defining the Minimal PTS in Luciferase

Amino acid position in luciferase (5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 / 3 3 3 3 4 4 4 4 4 4 4 4 4 4 5 / 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0)	Plasmid name	Subcellular distribution
R-E-I-L-I-K-A-K-K-G-G-K-S-K-L	pRSVL	P
R-E-I-	pRSVLΔC12	C
R-E-I- G-G-K-S-K-L	pRSVLΔ539-544	P
R-E-I- K-S-K-L	pRSVLΔ539-546	P
R-E-I- S-K-L	pRSVLΔ539-547	P
R-E-I- K-L	pRSVLΔ539-548	C
R-E-I- L	pRSVLΔ539-549	C
R-E-I-L-I-K-A-K-K-G-G-K	pRSVLΔC3	C
R-E-I-L-I-K-A-K-K-G-G-K-S	pRSVLΔC2	C
R-E-I-L-I-K-A-K-K-G-G-K-S-K	pRSVLΔC1	C
R-E-I-L-I-K-A-K-K-G-G-K-S-K-L-S	pRSVL + S	C
R-E-I-L-I-K-A-K-K-G-G-K-S-K-L-I-L	pRSVL + IL	C

The numbers in the table head (*left*) refer to the amino acid positions of the 550-amino acid-long luciferase. Below are the amino acid sequences of each mutant; the name of each mutant to its right; and on the extreme right, the location of the protein within the cell. *P*, peroxisomal; *C*, cytosolic. The subcellular location of each protein was determined as described in the legend to Fig. 1.

Table I provides a summary of the results of these experiments. Also, in Fig. 1 (*A–F*), we have presented the immunofluorescence data obtained with a representative subset of these mutants. Removal of sequences from amino acid 538 towards the COOH terminus of luciferase (which is 550 amino acids long) had no affect on targeting as long as the deletion did not extend to the serine residue 3 amino acids from the end of the protein. The mutant containing a 9-amino acid deletion (Δ539–547) that left only the last 3 amino acids of the 12-amino acid targeting signal had the same punctate distribution as wild-type luciferase (Fig. 1 *A*). Its colocalization with catalase within the same cell (compare *A* and *B* in Fig. 1) confirmed that this mutant protein (that retains serine-lysine-leucine at the COOH terminus) was peroxisomal. In contrast, the mutant Δ539–548 (which had only lysine-leucine at the COOH terminus) exhibited a diffuse, cytosolic distribution (Fig. 1 *C*) and did not colocalize with catalase (compare *C* with *D* in Fig. 1). This lack of import demonstrated a defect in the protein's targeting signal. Likewise, deletion of three (ΔC3), two (ΔC2), or one (ΔC1; Fig. 1, *E* and *F*) amino acid(s) from the COOH terminus abolished the ability of luciferase to be transported to peroxisomes (Table I). Overall, the results obtained with these mutants indicated

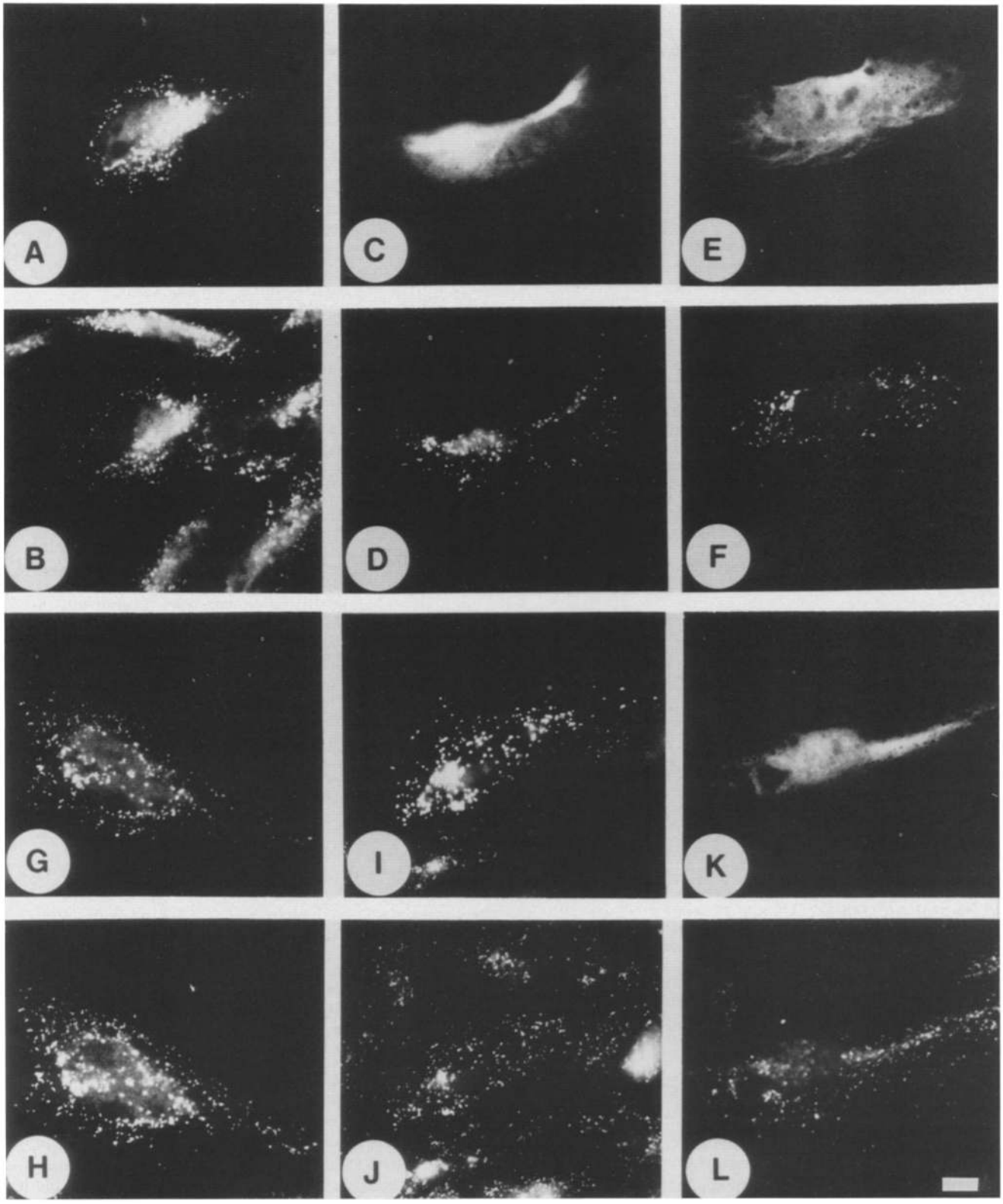

Figure 1. Immunofluorescent localization of mutant luciferases. CV-1 monkey kidney cells were transfected with each of the mutant luciferase genes described in Tables I and II and subsequently processed for double indirect immunofluorescence. We show the data obtained with a representative subset of these mutants. The micrographs are paired such that *A–E* and *G–K* show the distribution of luciferase (fluorescein labeled) inside transfected cells while *B–F* and *H–L* show the distribution of catalase (i.e., peroxisomes) for the cells in the same field (rhodamine labeled). Shown are the subcellular distributions of mutant luciferases encoded by pRSVLΔ539-547 (*A* and *B*); pRSVL-Δ539-548 (*C* and *D*); pRSVLΔC1 (*E* and *F*); pRSVL-AKL (*G* and *H*); pRSVL (*I* and *J*); and pRSVL-SKI (*K* and *L*).

Table II. Acceptable Degeneracy of the PTS

Amino acid position in luciferase 556 556 789 012 345		Plasmid name	Subcellular distribution
R-E-I-L-I-K-A-K-K-G-G-K-S-K-L		pRSVL	P
Serine 548 mutated to	Glutamic acid	pRSVL-EKL	C
	Leucine	pRSVL-LKL	C
	Phenylalanine	pRSVL-FKL	C
	Tyrosine	pRSVL-YKL	C
	Threonine	pRSVL-TKL	C
	Lysine	pRSVL-KKL	C
	Alanine	pRSVL-AKL	P
	Cysteine	pRSVL-CKL	P
Lysine 549 mutated to	Glycine	pRSVL-SGL	C
	Glutamic acid	pRSVL-SEL	C
	Isoleucine	pRSVL-SIL	C
	Serine	pRSVL-SSL	C
	Histidine	pRSVL-SHL	P
	Arginine	pRSVL-SRL	P
Leucine 550 mutated to	Serine	pRSVL-SKS	C
	Valine	pRSVL-SKV	C
	Glutamic acid	pRSVL-SKE	C
	Phenylalanine	pRSVL-SKF	C
	Lysine	pRSVL-SKK	C
	Isoleucine	pRSVL-SKI	C

Mutations affecting each of the last three amino acids of luciferase are shown. For each, the amino acid substitution is noted, and the subcellular distribution of the mutant luciferase is shown to the right. *P*, peroxisomal; *C*, cytosolic. The subcellular distribution of each mutant was determined as described in the legend to Fig. 1.

that the minimal PTS consists of the COOH-terminal three amino acids (serine-lysine-leucine) of luciferase.

The existence of this PTS at the extreme COOH terminus of the protein begs the question regarding the role, if any, the COOH-terminal location of the PTS plays in its function. Mutants were made that appended either one (serine) or two (isoleucine-leucine) amino acid(s) to the COOH terminus of luciferase. Both of these mutant proteins were observed to be cytosolic, indicating that addition of amino acids to the carboxy end of the PTS resulted in its inactivation (Table I).

Effect of Mutations within the Minimal PTS

We previously described a point mutation in the targeting signal of rat enoyl-CoA/hydratase-3-hydroxyacyl-CoA dehydrogenase (which ends in the sequence serine-lysine-leucine-COOH) that abolished the ability of the sequence to act as a PTS (Gould et al., 1988). In this mutation, an asparagine was substituted in place of the penultimate lysine residue. To obtain more complete data on acceptable sequence variation in the minimal PTS, site-directed mutagenesis was used to alter each of the last three codons of the luciferase gene so that a variety of amino acids were substituted for the three COOH-terminal amino acids of luciferase. As before, expression vectors containing these mutant luciferase genes were transfected into CV-1 cells. Double indirect immunofluorescence was performed on the transfected cell populations and the subcellular distribution of each mutant luciferase was determined (Table II). Immunofluorescence data obtained with a subset of these mutants are presented in Fig. 1 (*G–L*). The serine residue could be changed to the structurally similar amino acids alanine (Fig. 1, *G* and *H*) or cysteine, but not to threonine, tyrosine, phenylalanine, leucine, lysine, or glutamic acid. The lysine residue could be replaced by the basic residues arginine (Fig. 1, *I* and *J*) or histidine, but not by glycine, glutamic acid, isoleucine, or serine. Also, the requirement of a leucine at the last position seemed fixed, as neither isoleucine (Fig. 1, *K* and *L*), serine, valine, glutamic acid, lysine, or phenylalanine could be substituted for this amino acid without abolishing import of the protein to the organelle.

The Minimal PTS Is Sufficient to Direct a Cytosolic Protein into the Organelle

Though our results demonstrate that the tripeptide targeting signal identified above is necessary for import of luciferase into peroxisomes, the critical test of a such a signal is whether it is sufficient to direct a heterologous protein into the organelle. To see if the tripeptide serine-lysine-leucine-COOH could function in this manner, we fused a segment of DNA encoding only these three amino acids (followed by a translational stop codon) onto the 3′ end of the CAT coding region. As can be seen in Fig. 2 *A* and *B*, when the three amino acids serine-lysine-leucine were appended to the COOH terminus of CAT, the protein was found to colocalize with catalase in the peroxisomes of the expressing cells. As shown in previous reports (Gould et al., 1987, 1988) wild-type CAT was present in the cytosol of transfected cells (Fig. 2, *C* and *D*). These results demonstrate that the tripeptide serine-lysine-leucine-COOH is sufficient to direct a cytosolic protein to peroxisomes and confirms its ability to function as a PTS.

Discussion

Characteristics of the PTS

The work presented in this paper indicates that at least one type of signal for peroxisomal protein import consists of a 3-amino acid-long peptide with the "consensus sequence" of serine, alanine, or cysteine at the first position; lysine, histidine, or arginine at the second, and leucine at the third position, the COOH terminus of the protein. The results of the amino acid substitution experiments allow us to make some predictions regarding the nature of the PTS. The ability of serine, alanine, or cysteine to function at the first position of the sequence implies that recognition of this residue by a putative PTS receptor would be based primarily on structural properties (i.e., size of side chain). In contrast, because lysine, histidine, and arginine each contain basic side chains, it may be that the key requirement for this second residue is that it possess basic properties. Also, the inability of any substitutions for leucine to retain function implies that it is essential for function. The key properties that must be exhibited at this position cannot be surmised at this time.

We also found that the addition of one (serine) or two

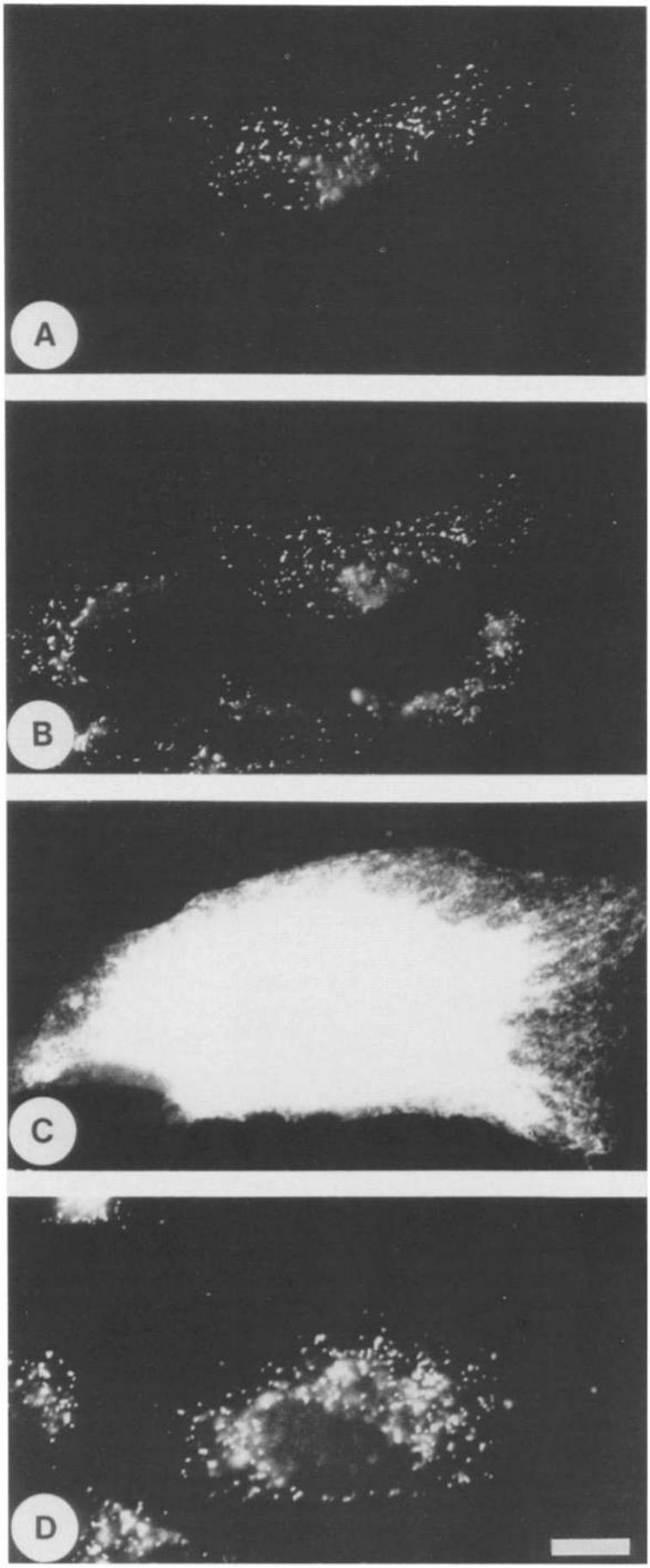

Figure 2. Immunofluorescent localization of the CAT-SKL fusion protein and wild-type CAT. CV-1 cells transfected with pSV2CAT-SKL or pSV2CAT were processed for double indirect immunofluorescence. Micrographs *A* and *C* depict the immunofluorescent staining with the anti-CAT antibody (fluorescein labeled) while micrographs *B* and *D* represent the immunofluorescent staining with the anticatalase antibody (rhodamine labeled). (*A* and *B*) Cells transfected with pSV2CAT-SKL. (*C* and *D*) Cells transfected with pSV2CAT.

Table III. Peroxisomal Targeting Sequences in Peroxisomal Proteins

Peroxisomal protein	Total no. amino acids	Conserved amino acids	Location		Reference
			COOH terminal	Internal*	
Class I					
Rat acyl-CoA oxidase	661	659 661 Ser-Lys-Leu	+	–	Miyazawa et al., 1987
Rat peroxisomal bifunctional enzyme	722	720 722 Ser-Lys-Leu	+	–	Osumi et al., 1985
Mouse or rat sterol-carrier-protein 2	143 partial	141 143 Ala-Lys-Leu	+	–	Scallen, T., personal communication; Morris et al., 1988
Rat uricase	cDNA	Ser-Arg-Leu	+	–	Reddy et al., 1988
Pig D-amino-acid oxidase	347	345 347 Ser-His-Leu	+	–	Ronchi et al., 1982
P. pyralis luciferase	550	548 550 Ser-Lys-Leu	+	–	de Wet et al., 1987
P. plagiophthalamus luciferase	543	541 543 Ser-Lys-Leu	+	–	Wood et al., 1989
Soybean uricase	309	307 309 Ser-Lys-Leu	+	–	Nguyen et al., 1985
Cucumber malate synthase	partial cDNA	Ser-Lys-Leu	+	–	Smith and Leaver, 1986
Spinach glycolate oxidase	369	367 369 Ala-Arg-Leu	+	–	Volokita and Somerville, 1987
C. boidinii PMP-20	167	165 167 Ala-Lys-Leu	+	–	Garrard and Goodman, 1989
Class II					
Human catalase	526	516 518 Ser-His-Leu	–	+	Bell et al., 1986
Rat peroxisomal thiolase	398	140 142 Ser-Arg-Leu	–	+	Hijikata et al., 1987
		252 254 Ala-Lys-Leu	–	+	
D. melanogaster xanthine dehydrogenase (rosy)	1335	56 58 Ala-Lys-Leu	–	+	Keith et al., 1987
C. tropicalis acyl-CoA oxidase PXP-4	709	554 556 Ser-Lys-Leu	–	+	Okazaki et al., 1986
C. tropicalis acyl-CoA oxidase PXP-5	662	522 524 Ser-Lys-Leu	–	+	Okazaki et al., 1986
C. tropicalis catalase	485	331 333 Ser-Arg-Leu	–	+	Okada et al., 1987
H. polymorpha methanol oxidase	664	592 594 Ala-Arg-Leu	–	+	Ledeboer et al., 1985

Class I proteins contain the PTS at the COOH terminus, and Class II proteins contain an analogous sequence in at least one internal location.
* A different COOH-terminal signal cannot be excluded.

(isoleucine-leucine) amino acids to the COOH terminus of luciferase blocks import of the protein into peroxisomes. It may be that the PTS requires the COOH terminal location for function, in which case a role may exist for the free carboxylic acid in recognition of the PTS by its putative receptor. Alternatively, it may be that its ability to function as a PTS is strongly context dependent, and these last two mutations may have placed the PTS in an unrecognizable context. Our results are similar to those observed for soluble proteins of the ER, where a 4-amino acid peptide (lysine-aspartic acid-glutamic acid-leucine, or KDEL) at the COOH terminus of these proteins acts as an ER retention signal (Munro and Pelham, 1987). However, it should be noted that the processes of ER retention and protein transport into the peroxisome differ significantly in that peroxisomal import requires translocation of proteins across a lipid bilayer whereas ER retention probably does not (Munro and Pelham, 1987; Pelham et al., 1988).

The immunofluorescence assay we used to detect import of peroxisomal proteins was qualitative and did not allow us to record subtle differences in import efficiency between various proteins. The microscopy images presented in this report were representative of the distribution of the proteins as determined by analysis of tens or hundreds of transfected cells. Because the transfection procedure may deliver greatly different amounts of plasmid DNA to different cells, the amount of protein produced in different cells may vary accordingly. In cells that express extremely high levels of even a normally peroxisomal protein such as luciferase, significant levels of the protein may be detected in the cytosol. Thus

when we observe a cell in which the protein of interest is detected in the cytosol as well as in peroxisomes, it is difficult to be certain if the dual localization of the protein is due to a low efficiency PTS on the protein or if it is caused by saturation of the import machinery because of hyperexpression of the introduced peroxisomal protein. We could observe no differences among the peroxisomal luciferases in the percentage of expressing cells in which both peroxisomal and cytosolic staining was seen. We are therefore unable to speculate as to the relative efficiencies of the peroxisomal targeting signal sequences that were identified herein. The most pressing limitation of our in vivo transport assay is that it may not be able to detect low efficiency import since a low initial rate of import will result in accumulation of the protein in the cytosol, resulting in a bright cytosolic signal that may mask weak peroxisomal fluorescence.

The Minimal PTS Is a Common Feature of Peroxisomal Proteins

A major prediction that follows from our identification of the minimal PTS is that other peroxisomal proteins should contain this element within their amino acid sequence. In accordance with this prediction, all but two of the peroxisomal proteins for which sequence data is available contain the tripeptide targeting signal somewhere within their sequence. Eleven of these proteins (*Class I*, Table III) contain the sequence serine-lysine-leucine or a conservative variant at their COOH terminus. Furthermore, the proteins in Class I (COOH-terminal PTSs) are from insects, mammals, plants, and yeasts, an observation suggesting that the PTS we have identified may be a general one that has been conserved through evolution. This is consistent with the peroxisomal localization of luciferase when expressed in any of these types of organisms (Keller et al., 1987; unpublished data).

The proteins in Class II (Table III) do not contain this consensus PTS at their C termini. However, they do contain the consensus sequence at one or more internal locations. We do not know whether the tripeptide targeting signal that we have identified functions as a PTS at internal locations within these proteins or whether some other type of PTS is active. Interestingly, the COOH-terminal 27 amino acids of human catalase can act as a PTS and contain, at an internal location, the consensus PTS sequence serine-histidine-leucine (Gould et al., 1988).

The only peroxisomal proteins we are aware of that do not contain the conserved tripeptide PTS sequence are rat catalase (Furuta et al., 1986) and dihydroxyacetone synthase from the yeast *Hansenula polymorpha* (Janowicz et al., 1985). In these proteins, the closest match to the consensus PTS is the sequence alanine-lysine-isoleucine in dihydroxyacetone synthase and serine-histidine-isoleucine in rat catalase. Though the leucine–to–isoleucine mutation inactivated the PTS in luciferase, it is quite possible that these sequences may function as PTSs at internal locations.

Even if the consensus PTS is able to work at internal locations, it is clear that it could not be a PTS in all contexts. A search of the protein identification resource protein sequence databank (June 30, 1988 release) identified a large number of eukaryotic proteins that contain the consensus PTS somewhere within their sequence. Most contain this sequence at internal locations and are not peroxisomal. Of the proteins that contain the PTS at the COOH terminus, where it is known to function, all are either peroxisomal or secretory proteins (excepting those with unknown subcellular distribution). It is not surprising that the search identified secretory proteins that have the COOH-terminal PTS. Because secretory proteins are transported co-translationally into the ER, they are committed to secretion before the synthesis of their COOH terminus, and any peroxisomal targeting information present at their COOH terminus would be effectively muted.

Evidence from the work of Small et al. (1988) suggests that certain polypeptides that lack the PTS we have identified are capable of being imported into *Candida tropicalis* peroxisomes in vitro (Small et al., 1988). The contrast between our results and those obtained in their yeast in vitro import system may be explained by the existence of PTSs in *C. tropicalis* that differ from those that act in mammalian cells. Alternatively, multiple signals may be capable of directing proteins to peroxisomes in both types of organisms. Additional experiments are required to distinguish between these possibilities.

Implications for Peroxisomal Protein Import

We have identified a PTS located at the COOH-terminus of several peroxisomal proteins that consists of a tripeptide with the consensus sequence of serine, alanine, or cysteine at the first position, lysine, arginine, or histidine at the second, and leucine at the third. The observation that a short, COOH-terminal sequence can direct proteins to peroxisomes is in contrast to results obtained in the study of mitochondrial, chloroplast, and ER transport, where rather large (at least 12 amino acids, and commonly even longer) NH_2-terminal pre-sequences are implicated in the import of proteins to the organelle. In addition, these other types of translocation signals lack sequence conservation and are thought to form higher order structures responsible for targeting activity (e.g., amphiphilic helices or β-sheets proposed to act as mitochondrial targeting signals [for review, see Verner and Schatz, 1988]), whereas the PTS we have studied has a consensus sequence.

Peroxisomal protein import is also unique in that it is the only one of these translocation processes not associated with proteolytic modification of the imported protein (Lazarow and Fujiki, 1985). That the consensus tripeptide PTS of peroxisomal proteins is not removed as a consequence of their import is confirmed by protein sequence analysis of D-amino acid oxidase showing that the purified, mature protein terminates with the COOH-terminal sequence serine-histidine-leucine (Ronchi et al., 1982). Whether the differences between peroxisomal import and other types of transmembrane translocation are an indication that the underlying mechanisms of import are fundamentally different is not clear at this time. An answer to this question must await the elucidation of the mechanisms underlying each type of transport.

We would like to thank our colleagues for their invaluable advice and comments during the course of this project and especially S. J. Singer for also allowing us the use of his facilities. We would also like to thank A. Schram for the gift of the rabbit antibodies against catalase and C. Gorman for the mAb-producing cell line CAT-2.

We wish to acknowledge the March of Dimes (grant No. 1081 to S. Subramani), the National Institutes of Health (training grants to N. Hosken and J. Wilkinson), and the Powell Foundation (fellowship to S. J. Gould) for financial support of this work.

Received for publication 9 December 1988 and in revised form 23 January 1989.

References

Bell, G. I., R. C. Najarian, G. T. Mullenbach, and R. A. Hallewell. 1986. cDNA sequence coding for human kidney catalase. *Nucleic Acids Res.* 14:5561–5562.

Blobel, G. 1980. Intracellular protein topogenesis. *Proc. Natl. Acad. Sci. USA.* 77:1496–1500.

Chen, E. J., and P. H. Seeburg. 1985. Supercoil sequencing: a fast simple method for sequencing plasmid DNA. *DNA (NY).* 4:165–170.

de Wet, J. R., K. W. Wood, M. DeLuca, D. R. Helinski, and S. Subramani. 1987. Firefly luciferase gene:structure and expression in mammalian cells. *Mol. Cell. Biol.* 7:725–737.

Furuta, S., H. Hayashi, M. Hijikata, S. Miyazawa, T. Osumi, and T. Hashimoto. 1986. Complete nucleotide sequence of cDNA and deduced amino acid sequence of rat liver catalase. *Proc. Natl. Acad. Sci. USA.* 83:313–317.

Garrard, L. J., and J. M. Goodman. 1989. Two genes encode the major membrane-associated protein of methanol-induced peroxisomes from *Candida boidinii. J. Biol. Chem.* In press.

Gorman, C. M., L. F. Moffat, and B. H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* 2:1044–1051.

Gould, S. J., G.-A. Keller, and S. Subramani. 1987. Identification of a peroxisomal targeting signal at the carboxy terminus of firefly luciferase. *J. Cell Biol.* 105:2923–2931.

Gould, S. J., G.-A. Keller, and S. Subramani. 1988. Identification of peroxisomal targeting signals located at the carboxy terminus of four peroxisomal proteins. *J. Cell Biol.* 107:897–905.

Hase, T., U. Muller, H. Reizman, and G. Schatz. 1984. A 70-kd protein of the yeast mitochondrial outer membrane is targeted and anchored via its extreme amino terminus. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:3157–3164.

Hijikata, M., N. Ishii, H. Kagamiyama, T. Osumi, and T. Hashimoto. 1987. Structural analysis of cDNA for rat peroxisomal 3-ketoacyl-CoA thiolase. *J. Biol. Chem.* 262:8151–8158.

Janowicz, Z. A., M. R. Eckart, C. Drewke, R. O. Roggenkamp, and C. P. Hollenberg. 1985. Cloning and characterization of the DAS gene encoding the major methanol assimilatory enzyme of the methylotrophic yeast *Hansenula polymorpha. Nucleic Acids Res.* 13:3043–3062.

Johnson, L. M., V. A. Bankaitis, and S. D. Emr. 1987. Distinct sequence determinants direct intracellular sorting and modification of a yeast vacuolar protease. *Cell.* 48:875–885.

Kalderon, D., B. L. Roberts, W. D. Richardson, and A. E. Smith. 1984. A short amino acid sequence able to specify nuclear location. *Cell.* 39:499–509.

Keith, T. P., M. A. Riley, M. Kreitman, R. C. Lewontin, D. Curtis, and G. Chambers. 1987. Sequence of the structural gene for xanthine dehydrogenase (rosy locus) in *Drosophila melanogaster. Genetics.* 116:67–73.

Keller, G.-A., S. J. Gould, M. Deluca, and S. Subramani. 1987. Firefly luciferase is targeted to peroxisomes in mammalian cells. *Proc. Natl. Acad. Sci. USA.* 84:3264–3268.

Kunkel, T. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA.* 82:488–492.

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

Ledeboer, A. M., L. Edens, J. Maat, C. Visser, J. W. Bos, C. T. Verrips, Z. Janowicz, M. Eckart, R. Roggencamp, and C. P. Hollenberg. 1985. Molecular cloning and characterization of a gene coding for methanol oxidase in *Hansenula polymorpha. Nucleic Acids Res.* 13:3063–3083.

Maniatis, T., E. F. Frisch, and J. Sambrook. 1982. Molecular Cloning: A laboratory Handbook. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.

Morris, H. R., B. S. Larsen, and J. T. Billheimer. 1988. A mass spectrometric study of the structure of sterol carrier protein SCP_2 from rat liver. Biochem. Biophys. Res. Commun. 154:476–482.

Miyazawa, S., H. Hayashi, M. Hijikata, N. Ishii, S. Furuta, H. Kagamiyama, T. Osumi, and T. Hashimoto. 1987. Complete nucleotide sequence of cDNA and predicted amino acid sequence of rat acyl-CoA oxidase. *J. Biol. Chem.* 262:8131–8137.

Miyazawa, S., T. Osumi, T. Hashimoto, K. Ohno, S. Miura, and Y. Fujiki. 1989. Peroxisome targeting signal of rat acyl-coenzyme A oxidase resides at the carboxy terminus. *Mol. Cell. Biol.* 9:83–91.

Munro, S., and H. R. B. Pelham. 1987. A C-terminal signal prevents secretion of luminal ER proteins. *Cell.* 48:899–907.

Nguyen, T., M. Zelechowska, V. Foster, H. Bergmann, and D. P. S. Verma. 1985. Primary structure of the soybean nodulin-35 gene encoding uricase II localized in the peroxisomes of uninfected cells of nodules. *Proc. Natl. Acad. Sci. USA.* 82:5040–5044.

Okada, H., M. Ueda, T. Sugayama, H. Atomi, S. Mozaffar, T. Hishida, Y. Teranishi, K. Okazaki, T. Kamiryo, and A. Tanaka. 1987. Catalase gene of the yeast *Candida tropicalis. Eur. J. Biochem.* 170:105–110.

Okazaki, K., T. Takechi, N. Kanbara, S. Fukui, I. Kubota, and T. Kariyo. 1986. Two acyl-coenzyme A oxidases in peroxisomes of the yeast *Candida tropicalis*: primary structures deduced from genomic DNA sequence. *Proc. Natl. Acad. Sci. USA.* 83:1232–1236.

Osumi, T., N. Ishii, M. Hijikata, K. Kamijo, H. Ozasa, S. Furuta, S. Miyazawa, K. Kondo, K. Inoue, H. Kagamiyama, and T. Hashimoto. 1985. Molecular cloning and nucleotide sequence of the cDNA for rat peroxisomal enoyl-CoA: hydratase-3-hydroxyacyl-CoA dehydrogenase bifunctional enzyme. *J. Biol. Chem.* 200:8905–8910.

Parker, B. A., and G. R. Stark. 1979. Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infection by virus and virus DNA. *J. Virol.* 31:360–369.

Pelham, H. R. B., K. G. Hardwick, and M. J. Lewis. 1988. Sorting of soluble ER proteins in yeast. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:1757–1762.

Reddy, G. P., M. R. Nemaly, M. K. Reddy, M. N. Reddy, P. M. Yuan, S. Yuen, T. G. Laffler, T. Shiroza, H. K. Kuramitsu, N. Usuda, R. L. Chisholm, M. S. Rao, and J. K. Reddy. 1988. Isolation and sequence determination of cDNA clone for rat peroxisomal urate oxidase: liver-specific expression in the rat. *Proc. Natl. Acad. Sci. USA.* 85:9081–9085.

Ronchi, S., L. Minchiotti, M. Galliano, B. Curti, R. P. Swenson, C. H. Williams, and V. Massey. 1982. The primary sequence of D-amino acid oxidase from pig kidney. *J. Biol. Chem.* 257:8824–8834.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–5467.

Small, G. M., L. J. Szabo, and P. B. Lazarow. 1988. Acyl-CoA oxidase contains two targeting sequences each of which can mediate proteins import into peroxisomes. *EMBO (Eur. Mol. Biol. Organ.) J.* 7:1167–1173.

Smith, S. M., and C. J. Leaver. 1986. Glyoxysomal malate synthase of cucumber: molecular cloning of a cDNA and regulation of enzyme synthesis during germination. *Plant Physiol. (Bethesda).* 81:762–767.

Van den Broek, G., M. P. Timko, A. P. Kausch, A. R. Cashmore, M. Van Montagu, and L. Herrera-Estrella. 1985. Targeting of a foreign protein to chloroplasts by fusion to the transit peptide from the small subunit of ribulose-1, 5-bisphosphate carboxylase. *Nature (Lond.)* 313:358–363.

Verner, K., and G. Schatz. 1988. Protein translocation across membranes. *Science (Wash. DC).* 241:1307–1313.

Volokita, M., and C. R. Somerville. 1987. The primary structure of spinach glycolate oxidase deduced from the DNA sequence of a cDNA clone. *J. Biol. Chem.* 262:15825–15828.

Walter, P., and V. R. Lingappa. 1986. Mechanism of protein translocation across the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* 2:499–516.

Wood, K. W., Y. A. Lam, H. H. Feliger, and W. D. McElroy. 1989. cDNAs encoding click beetle luciferases can elicit bioluminescence of different colors. *Science (Wash. DC).* In press.

Zoller, M. J., and M. Smith. 1983. Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. *Methods Enzymol.* 100:468–500.