The Mouse p97 (CDC48) Gene

GENOMIC STRUCTURE, DEFINITION OF TRANSCRIPTIONAL REGULATORY SEQUENCES, GENE
EXPRESSION, AND CHARACTERIZATION OF A PSEUDOGENE*

(Received for publication, October 23, 1998, and in revised form, January 14, 1999)

Joyce M. M. Müller‡§, Hemmo H. Meyer‡¶, Christiana Ruhrberg‡, Gordon W. Stamp**
Graham Warren‡, and David T. Shimad‡‡

From the §Cell Biology Laboratory, Imperial Cancer Research Fund, Lincoln’s Inn Fields, London WC2A 3PX, England. Tel.: 44-0171-269-417; E-mail: shima@icrf.icnet.uk.
¶ Supported by an Imperial Cancer Research Fund (ICRF) fellowship and the Deutsche Forschungsgemeinschaft.
†† Supported by the ICRF and a Hitchings-Elion fellowship from the Developmental Neurobiology, National Institute of Medical Research, The Ridgeway, Mill Hill NW7 1AA, and the Department of Histopathology, Imperial College School of Medicine, Hammersmith Hospital London, Ducane Road, London W12 0HS, United Kingdom.

THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 274, No. 15, Issue of April 9, pp. 10154–10162, 1999 Printed in U.S.A.

Here we present the first description of the genomic organization, transcriptional regulatory sequences, and adult and embryonic gene expression for the mouse p97(CDC48) AAA ATPase. Clones representing two distinct p97 genes were isolated in a genomic library screen, one of them likely representing a non-functional processed pseudogene. The coding region of the gene encoding the functional mRNA is interrupted by 16 introns and encompasses 20.4 kilobase pairs. Definition of the transcriptional initiation site and sequence analysis showed that the gene contains a TATA-less, GC-rich promoter region with an initiator element spanning the transcription start site. cis-acting elements necessary for basal transcription activity reside within 410 base pairs of the flanking region as determined by transient transfection assays. In immunohistological analyses, p97 was widely expressed in embryos and adults, but protein levels were tightly controlled in a cell type-and cell differentiation-dependent manner. A remarkable heterogeneity in p97 immunostaining was found on a cellular level within a given tissue, and protein amounts in the cytoplasm and nucleus varied widely, suggesting a highly regulated and intermittent function for p97. This study provides the basis for a detailed analysis of the complex regulation of p97 and the reagents required for assessing its functional significance using targeted gene manipulation in the mouse.

p97 belongs to the family of AAA ATPases associated with diverse cellular activities (AAA) occurring in eubacteria, archaeabacteria, and eukaryotes (1). The AAA motif is defined by a conserved sequence of 200 amino acids including the Walker type A and B cassettes, which are important in ATP binding and hydrolysis (2). AAA family members include proteins involved in vesicle and organelle biogenesis (3, 4), components of the 26 S proteasome (5), metalloproteases (6), cell cycle regulators (7), and transcription factors (8).

Mammalian p97 (first termed VCP, for valosin-containing protein) was originally described as a precursor protein containing the biologically active peptide valosin (9). Subsequently, in *Xenopus laevis*, a 14 S, homo-oligomeric ATPase was identified as the homologue of mammalian VCP (10), and in *Saccharomyces cerevisiae*, genetic alterations in the p97 homologue (CDC48) were shown to underlie a mitotic-arrest phenotype (7, 11). More recently, highly conserved p97 homologues have been identified in a diverse range of experimental organisms, including *Drosophila melanogaster* (TER94) (12), *Arabidopsis thaliana* (AtCDC48) (13), and the archaeabacteria *Sulfobus acidocaldarius* (SAV) (14), thus demonstrating that p97 is an ancient protein and implying a fundamental role(s) for this protein within cells.

Purified p97 is a soluble, ringlike hexameric complex with each monomeric subunit containing two copies of the AAA domain. This complex is the active Mg2+-dependent ATPase, which is sensitive to the alkylating agent N-ethylmaleimide (10, 15). The significance of these structural and biochemical features for the cellular function of p97 has not yet been firmly established.

Data from diverse origins have implicated p97 in a remarkable number of cellular processes. In *S. cerevisiae*, genetic analysis has shown that conditional mutants in the yeast CDC48 gene, named cdc48–1, arrest in mitosis as large budded cells with elongated nuclei spanning the mother-daughter junction (11). p97 has been proposed to function in homotypic membrane fusion events, including fusion of the endoplasmic reticulum (16) and the reassembly of Golgi cisternae from vesicles and tubules generated by treatment with specific drugs (17) or mitotic cytosol (4). Using the Golgi reassembly assay, p97’s fusion-promoting activity has recently been shown to be dependent upon a stoichiometric association with a cytosolic protein termed p47 (18). Protein interaction studies have also revealed that p97 binds to clathrin in a stoichiometric complex, suggesting a role for p97 in the endocytic cycle (19). p97 has been shown to physically associate with Ufd3p, a protein that is involved in the regulation of free cellular ubiquitin in yeast (20), and to copurify with the mammalian 26 S proteasome (21). Furthermore, in these same experimental systems, immunodepletion or inactivation of p97 severely compromised ubiquitin-dependent proteolyis of experimental substrates in *vitro* and in *vivo* (20, 21).

Despite the abundance of clues provided by these investigations, a common functional link relating these varied observations remains to be established, and thus the precise cellular
role(s) of p97 is still obscure. To begin to understand the physiological role of p97 in mammals, we have initiated an analysis of p97 structure and function in mouse. Here we present the first description of the genomic structure of the p97 gene and a likely processed pseudogene. We also identify an upstream region that acts as a functional basal promoter and potential regulatory elements, and describe a pattern of p97 expression in mouse embryos and adults that suggests an unexpected degree of regulated expression and subcellular localization in both proliferating and differentiated tissues.

**MATERIALS AND METHODS**

**Gene Isolation and Physical Mapping**—A DASH II genomiclibrary (kindly provided by R. Mortenson, Harvard Medical School, Boston, MA) prepared from 129 SVJ mouse spleen was screened with a random primed 32P-labeled probe corresponding to the mouse p97 coding region spanning amino acids 2–141. Approximately 9 × 10⁸ plaques were screened on charged nylon filters by hybridizing the p97 probe overnight at 42 °C in 1% SDS, 2× SSC, 10% dextran sulfate, 50% deionized formamide, 2.5× Denhardt’s, and 0.1 mg/ml denatured calf thymus DNA. Several positive clones were identified and rescreened in two additional rounds. Phage DNA was isolated and analyzed by restriction digestion and Southern blot. Restriction fragments spanning the relevant probes were subcloned into pBluescript (Stratagene) for further analysis.

**Fluorescence in Situ Hybridization**—Purified phage DNA from three unique clones was used to determine the chromosomal location of the mouse p97 gene by fluorescence in situ hybridization. Briefly, metaphase spreads were prepared using standard cytogenetic techniques from mouse diploid cultures and cell lines. Phage DNA was labeled with biotin dUTP by nick translation (Bionick, Life Technologies, Inc.). The labeled probe was combined with mouse Cot-1 DNA and hybridized to metaphase chromosomes in a solution containing 50% formamide, 10% dextran sulfate, 2× SSC, and 1% Tween 20, pH 7.0. Specific hybridization signals were detected by incubating slides in fluorescein isothiocyanate-conjugated avidin. Slides were counterstained with 4,6-diamino-2-phenylindole dihydrochloride (Sigma) and analyzed with a Zeiss Axioscop microscope.

**DNA Sequencing**—p97 subclones were sequenced with gene-specific primers using the ABI dye termination kit (Perkin Elmer). Primers were designed on the basis of the published mouse p97 cDNA sequence (22). Alignment of cdna and genomic DNA sequence was performed by using the tblastg analysis suite (version 9.0).

**Southern Blot Analysis**—Genomic DNA was prepared from mouse spleen using standard protocols. Restriction-digested DNA (10 μg) was electrophoresed, deparaffinized in 0.25 M HCl, denatured, and transferred in 0.4× NaOH to charged nylon (GeneScreen Plus, NEN Life Science Products). Membranes were rinsed in 2× SSC, dried, and hybridized according to the manufacture’s protocol. Blots were washed with 1× SDS, 5× SSC, 55 °C and visualized by autoradiography.

**Primer Extension Analysis**—The transcription start site of the p97 gene was determined by primer extension assay using two different end-labeled antisense oligonucleotides (primer 1: 5′-CCGGGGCTG-GACTCCGAGAAGGGG-3′; primer 2: 5′-CTCTCGCTTCCTCCCAGGG-GACCC-3′). Total RNA from mouse embryonic fibroblasts was isolated with RNAzol (Biogenesis) and divided into poly(A)⁺ and poly(A)⁻ pools using oligo(dT) affinity beads (Oligotex, Qiagen). The assay was performed using standard protocols. Briefly, radiolabeled oligonucleotide was mixed with approximately 10 μg of each RNA pool and denatured for 90 min at 65 °C in 0.15 M KCl, 1 mM EDTA, and 10 mM Tris-Cl, pH 8.3. The mixture was cooled to room temperature and primer extension was initiated by addition of 10 mM MgCl₂, 5.5 mM dithiothreitol, 150 μg/ml actinomycin D, 0.15 mM dNTPs, and 150 units of Superscript II (Life Technologies, Inc.) at 42 °C for 60 min. Products were separated on 8 M urea, 8% polyacrylamide gels and visualized by autoradiography.

The precise location of the transcription start site was determined by comparison to deoxyxarny sequencing reactions carried out using the same oligonucleotides.

**Construction of Promoter-Luciferase Fusion Vectors**—Three fragments containing 147 bp of the 5′-untranslated region and additionally either 410, 1434, or 3000 bp of 5′-flanking sequence were obtained by polymerase chain reaction and ligated in both orientations with respect to the predicted transcription start site into a promoterless luciferase reporter plasmid (pGL2-basic, Promega). The pGL2-control plasmid (Promega), which utilizes the SV 40 promoter/enhancer to initiate luciferase transcription, was used as a positive control.

**Cell Culture and Transient Transfection**—HeLa cells were maintained in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal calf serum. The day before transfection, cells were split into six-well plates. On the next day, cells were transfected in triplicate with 5 μg of test plasmid using SuperFect (Qiagen). To control for transfection efficiency, an empty vector containing a secreted alkaline phosphatase reporter gene was cotransfected (2 μg) with test plasmids (23). Alkaline phosphatase activity was determined as described previously (24) and used to normalize luciferase assay values. Luciferase activity in cell extracts was assayed 48 h after transfection according to the manufacturer’s protocol (Promega).

**Immunostaining of Mouse Tissues**—Anti-p97 antibodies N2 and N5 were raised in rabbit against two bacterially expressed fusion proteins consisting of glutathione S-transferase (GST) fused to fragments of mouse p97 (N2: amino acids 2–186; N5: amino acids 200–461). The antibodies were affinity-purified using full-length recombinant Histagged p97 and were shown to uniquely recognize p97 in Western blot analyses of rat liver homogenate (data not shown). Immunohistochemical results obtained using N2 or N5 were indistinguishable.

For analyses, samples were fixed in 3% formaldehyde, embedded in paraffin, sectioned, and dewaxed. Endogenous peroxidase activity was quenched, and sections were incubated with purified antibodies (0.5 μg/ml). Binding was detected by subsequent incubation with biotinylated secondary antibody and streptavidin-coupled peroxidase, followed by development with diaminobenzidine. In adult tissues, counterstaining was performed with hematoxylin.

**mRNA in Situ Hybridization**—Samples were fixed and prepared for as for immunostaining. Specific localization of p97 mRNA was accomplished by in situ hybridization using an antisense riboprobe synthesized with SP6 RNA polymerase using 32P-UTP (~800 Ci/mmol; Amer sham, UK). The linearized template consisted of a 500-bp fragment of the mouse p97 3′-untranslated region within the plasmid backbone of pSP73. The 3′-untranslated region of sequence used to produce the riboprobe did not show significant homology to any other known gene sequences in the nucleotide data base (GenBank version 109.0). The methods for pretreatment, hybridization, washing, and dipping of slides in Ilford K5 for autoradiography were basically as described by Senior et al., for formalin-fixed paraffin-embedded tissue (25), with modifications (26). The presence of hybridizable mRNA in all compartments of the tissues studied was established in near serial sections using an antisense β-actin probe. Sections were examined under conventional or reflected light darkfield conditions (Olympus BH2 with epi-illumination) that allowed individual autoradiographic silver grains to be seen as bright objects on a dark background.

**RESULTS**

**Identification of Two Different p97 Loci in M. musculus**—To investigate the genomic structure of the mouse p97 gene, we used a probe spanning the N-terminal region of the protein to isolate three unique clones from a λ phage library. Southern blot analysis on digested phage DNA using probes spanning different regions of the published cDNA (22) was performed to further characterize the isolated clones. Two clones, designated λ1 and λ2, extended 20 and 15.6 kb, respectively, and overlapped by 3.7 kb. λ2 spanned the 5′ end and λ1 the 3′ end of the p97 cDNA sequence (Fig. 1a). A third clone, designated λ3, extended 14 kb and appeared to encompass the whole p97 coding region based on Southern blot analysis; however, initial analyses revealed that restriction patterns within the λ97 coding sequence of λ3 and λ2 differed (Fig. 1a). A more detailed physical map of the different p97 clones was assembled by performing single and double digestions with various restriction enzymes, and Southern blot analysis using radiolabeled probes spanning the published cDNA of p97. Comparison of the physical maps of total genomic and λ clone DNA revealed that mouse p97 is a single copy gene. Southern blot analyses revealed that we consistently found two restriction fragments in the total genomic DNA hybridizing to a p97 specific probe. One fragment corresponded to a similarly sized fragment in λ1/2, while the second fragment corresponded to a fragment in λ3, strongly suggesting the presence of two different p97 loci (results for λ2 versus λ3 are shown in Fig. 1b).

The presence of two mouse p97 loci had recently been pro-
posed based on interspecies backcross mapping and polymerase chain reaction analysis (27). Fluorescence in situ hybridization using the isolated λ clones as probes confirmed this genetic analysis. The overlapping clones λ1 and λ2 reside on mouse chromosome 4 at position B2, while clone λ3 resides on chromosome X at position C-D (data not shown).

**Fig. 1. Genomic structure of the mouse p97 genes.**  
**a**, physical map of the mouse p97 genes as established from λ1/2 and λ3; sites are marked for the enzymes BamHI (B), EcoRI (E), HindIII (H), and XhoI (X). The open reading frame of p97 is indicated in black shading and the untranslated regions in gray shading.  
**b**, Southern blot analysis of phage and genomic DNA. DNA was digested (10 μg) with the indicated enzymes and analyzed in Southern blot using a mouse cDNA probe covering the base pairs 14–159 of the published mouse cDNA.
The genomic structure of both genes was further characterized by nucleotide sequence analysis. Analysis of the gene present in \( l_1/2 \) revealed that it was interrupted by 16 introns, and all exon-intron borders corresponded to consensus splicing signals, except the exon 14/intron 14 boundary (Fig. 2a). Exons were on average 150 bp in size, and both nucleotide-binding cassettes (2) were located in single exons, exon 7 and exon 13. Intron sizes were estimated with polymerase chain reaction analysis using exon-specific primers. The location of exons relative to the restriction map was established by nucleotide sequencing of restriction sites proximal to exons, and restriction digestion analysis of polymerase chain reaction-amplified DNA of the subclones (Fig. 1a).

Alignment of nucleotide sequences of \( l_3 \) to the published mouse cDNA (22) revealed that this gene contained the complete coding region of p97 without disrupting introns. Several mutations were identified, including a substitution of Val207 to Ile and a deletion of Gln 569, Ala570, Ala571, Pro572, and Cys573. A polyadenylation signal and a poly(A) stretch were found in the region corresponding to the 3' end of the published cDNA. Each end of the gene was flanked by a direct repeat of 8 nucleotides. The structure of the gene encoded by \( l_3 \) had all the hallmarks expected of a processed pseudogene (designated herein as \( c_{p97} \)) (28), implying that the gene does not encode a functional product, as reported for many other pseudogenes (29).

In support of this notion, alignment of data base cDNA sequences for p97 gene counterparts in organisms ranging from archaeabacteria to mouse demonstrated that all cDNAs encoded the amino acid sequence AP (found at positions 571–572 in the mouse protein), which was present in the discontinuous p97 gene (\( l_1/2 \)), but which was deleted in \( c_{p97} \) (Fig. 2b). Therefore, we tentatively concluded that \( c_{p97} \) most likely represents a non-functional processed pseudogene, rather than a second, independently evolved, intronless and active p97 gene. Thus, we focused our attention on the further characterization of the p97 gene locus represented by \( l_1/2 \), which appeared to encode the functional p97 mRNA.

To identify the transcription start site, primer extension assays were performed. Based on the 5' end nucleotide sequence of the p97 cDNA and the region of deviation between the p97 gene and its pseudogene, two antisense primers were designed to map the 5' end of the p97 gene. The oligonucleotides were annealed to either purified poly(A)\(^+\) RNA or poly(A)\(^-\) RNA from mouse embryonic fibroblasts using primer 1 (see "Materials and Methods"). A sequencing ladder of the corresponding genomic fragments primed with the same oligonucleotide was electrophoresed in parallel. A single major extension product was identified, indicating the transcription initiation site (the complementary nucleotide is shown by an asterisk); nucleotides encompassing the initiator element are shown. Therefore, we concluded that the primer extension product is mapped to the transcription start site of the p97 gene.
poly(A)\(^-\) RNA as negative control. After reverse transcription and denaturing gel electrophoresis, a single major product was identified for each primer extension reaction. The precise transcription start site, designated +1, was determined by comparison of the migration of the products with a sequencing reaction using the same primers which were run in parallel on the gel (results for primer 1 are shown in Fig. 3). A consensus mRNA-CAP site resided at the identified transcription start, thus supporting the mapping data. Both primers also gave rise to a low abundance extension product, which may correspond to a minor transcription start at position +33 (data not shown).

Characterization of the Promoter of p97—Sequence analysis of the putative transcription initiation region demonstrated the absence of a consensus TATA sequence for RNA polymerase II-initiated transcription. Instead, an initiator element (Inr) absence of a consensus TATA sequence for RNA polymerase of the putative transcription initiation region demonstrated the (results for primer 1 are shown in Fig. 3). A consensus mRNACAP site resided at the identified transcription start, thus supporting the mapping data. Both primers also gave rise to a low abundance extension product, which may correspond to a minor transcription start at position +33 (data not shown).

To determine if the 5'-flanking region of the p97 gene can initiate basal transcription, a fragment containing 3000 bp of the flanking nucleotide sequence and additionally 147 bp of the 5'-untranslated region of the gene was inserted in both orientations into a luciferase reporter construct and assayed for activity in transiently transfected HeLa cells. Fusion of p97 sequence to the reporter gene in the appropriate orientation resulted in approximately 75-fold increase of luciferase activity in transiently transfected HeLa cells. Fusion of p97 sequence to the reporter gene in the appropriate orientation resulted in approximately 75-fold increase of luciferase activity when compared with either promoterless luciferase constructs or constructs in which p97 sequence was fused in the opposite transcriptional orientation (Fig. 5). Deletion of 1566 bp from the 5' end of the promoter fragment resulted in a 1.6-fold increase in reporter activity, whereas deletion of 2590 bp resulted in activity similar to that of the full-length fragment, implying that the minimal sequence necessary for basal transcription of the p97 gene in HeLa cells maps to approximately −1 to −410 with respect to the transcription start site. Comparison of the activity of the p97 promoter fragments to an assayed construct containing the SV 40 promoter/enhancer, which was used as positive control, revealed that the p97 promoter was a relatively powerful transcriptional initiator, with p97 promoter sequences yielding 50–75% of the activity measured for the viral promoter/enhancer. The fragment containing 1434 bp of the 5'-flanking region initiated transcription of the reporter gene more efficiently than the 3000-bp fragment, indicating that potential repressor elements may reside further upstream of the basal promoter. However, more detailed studies will be required to identify the components involved in regulating transcription of the p97 gene.

Distribution of p97 in Mammalian Tissues—To provide clues about the gene regulation and potential function of p97 in vivo, we investigated the distribution of p97 protein in adult and embryonic mouse tissues by immunohistochemical staining using affinity purified antisera (see “Materials and Methods”). For adult tissue, the immunohistochemical analysis clearly showed that p97 was widespread in most investigated tissues (results from the small intestine, liver, testis, and kidney are shown in Fig. 6 and summarized in Table I). However, the degree of staining between different cell types or cell differentiation states within a given tissue showed exceptional heterogeneity. This type of heterogeneity was not observed in an analysis performed in parallel for the Golgi apparatus structural protein giantin, which displayed a relatively uniform distribution in virtually all cells within the tissues examined. An example of the differential expression of p97 could be observed in the intestinal epithelium (Fig. 6a). Whereas within the villi the epithelial cells were largely positive in the cytoplasm and nucleus, the crypts, including the stem cell-rich proliferative zone, were mostly negative. Interestingly, other tissues rich in mitotic cells, such as the epidermis and the proliferative zone of the hair bulb, often showed little or no staining for p97 (data not shown).

As with the small intestine, cells within the liver exhibited heterogeneity in the levels of p97 protein. Yet, in contrast to the uniform cytoplasmic and nuclear staining observed in the intestinal villi, hepatocytes often showed relatively weak cytoplasmic staining and strongly positive nuclei (Fig. 6b). Interestingly, however, nuclear staining was not seen in all hepatocytes (arrow in b); Kupffer cells were consistently negative (arrow in b), whereas the cells of the bile duct were strongly positive in both cytoplasm and nucleus (arrow in b). In testis (Fig. 6c), the interstitial cells were uniformly stained with a relatively strong cytoplasmic and a weak nuclear signal (arrowhead); however, in the seminiferous tubules, differential staining was observed between basal and suprabasal populations, suggesting heterogeneity in p97 expression during different stages of spermatogenesis. Differential levels of p97 were also observed within the kidney (Fig. 6d).
The brown reaction product produced by DAB immunohistochemistry denotes positively stained cells. Sections were counterstained with hematoxylin. a, vili and crypts in the small intestine; a, higher magnification of the border between positive and negative (arrow) epithelium. b, detail of liver tissue near bile duct (arrow); b', hepatocytes. Note negative nuclei of some hepatocytes (arrowhead) and Kupffer cells (arrow). c, cross-section of seminiferous tubule with interstitial cells in testis. c', staining was uniform in the cytoplasm of interstitial cells (filled arrowhead), absent in myoid cells (empty arrowhead), and the majority of leptotene spermatocytes (arrow); strong cytoplasmic and nuclear staining was present in basal and suprabasal layers, which primarily consist of type A spermatogonia (asterisk). Spermatids and spermatozoa nuclei were relatively weakly stained. d, kidney shows little staining within a glomerulus and strong staining of various parts of the unfferous tubules of the renal cortex; d', a cross-section of a proximal tubule in which the majority of cells exhibited a strong cytoplasmic and nuclear staining. Occasionally, however, negative nuclear staining (arrowhead) and negative cells were observed (arrow).

Immunohistochemical analysis of variously staged mouse embryos also revealed a widespread, but non-uniform pattern of p97 protein expression. Wholemount stainings from 9.5 and 10.5 dpc embryos suggested expression of p97 within most regions of the developing embryo, with specific elevations in the emerging limb buds, tail bud, branchial arches, and somites (Fig. 7, a and b). Furthermore, immunostaining of sectioned embryos and comparison to parallel hematoxylin and eosin sections revealed that highly elevated levels of p97 were present within one layer of the somites, most likely the myotome (Fig. 7, d and e). Limb buds exhibited moderate levels of p97 throughout the mesenchyme, but higher levels of staining in a subset of cells within the apical ectodermal ridge and surface epithelium (data not shown). Furthermore, as seen in adult tissue specimens, p97 expression within many embryonic tissues exhibited distinct cell-to-cell heterogeneity. Examples of this heterogeneity are shown for the condensing mesenchyme of the kidney anlagen (10.5 dpc; Fig. 7f) and the Wolffian duct epithelium (10.5 dpc; Fig. 7g), and at a later stage of development within a subset of chondrocytes in the hypertrophic zone of developing limb cartilage (13.5 dpc; Fig. 7h). Heterogeneous staining was also observed within the spinal cord (data not shown). In the embryonic tissues analyzed, increased expression of p97 was associated with an apparent accumulation of the protein in the nucleus.

Finally, mRNA expression in situ was examined to determine if the tissue-specific expression of p97 protein was associated with similar alterations in mRNA levels. Shown in Fig. 8 is an example of results from a comparison of p97 and β-actin mRNA expression in the testes. Unlike β-actin mRNA, which was abundant and widespread across most cellular regions of the seminiferous tubules (the exception being the mature sperm), p97 mRNA was differentially expressed within cells of the tubule (arrow). Similar to our observations for p97 protein (see Fig. 7), elevated mRNA levels were highest in basal and suprabasal cells (arrowheads) located 1–2 cell layers from the basement membrane. A comparison of actin and p97 mRNA expression in several other adult and embryonic mouse tissues (data not shown) consistently yielded a similar contrast in the distribution pattern of mRNAs from the “housekeeping” gene β-actin and p97, and thus suggested a general role for tissue-specific regulation, rather than constitutive production of p97 mRNA. These results also suggest that a contributing factor to the tissue-specific control of p97 protein expression is the regulation of mRNA levels.

**DISCUSSION**

**Tissue-specific and Heterogeneous p97 Expression in the Mouse**—p97 is a highly conserved and ancient protein with homologous counterparts found in organisms ranging from archaebacteria to man (13). Evidence from several different organisms and experimental systems has suggested disparate roles for this protein in cellular physiology. Based on the evolutionary conservation and widespread abundance of p97, it had earlier been suggested that it is a ubiquitously expressed component (10), and several of the currently proposed cellular functions for p97 fall into the category of “housekeeping” tasks, implying a constitutive role for p97 in all cells (10, 15, 19, 20). However, our current findings, as well as recent work from other investigators, support the prospect that p97 may perform a more regulated, rather than constitutive function in cells.

Mouse p97 protein and mRNA is indeed found widely in both developing embryos and the adult, but appears to be regulated in a cell-type- and cell differentiation-dependent manner within a given tissue. For example, p97 was enriched in embryos within the myotome region of the somites, was differentially expressed within regions of the adult small intestine, and varied among cell types and developmental stages in the testis. Likewise, in *Drosophila*, the p97 homologue has been recently reported to undergo temporal and tissue-specific regulation, being undetectable in larvae and found predominantly in the brain and gonads of the imago and adult fly (12).

The most surprising finding to come from the immunohistochemical analyses was that p97 levels varied greatly within neighboring cells of both established and developing tissues, such as hepatocytes, kidney epithelium, and embryonic chondrocytes. This pattern was not found in control stainings for the Golgi apparatus protein giantin. Thus, these findings suggest a specific cell-to-cell regulation of p97 protein levels, which one could envision as a cell autonomous response to alterations in subcellular physiology, or a single cell’s response to local microenvironmental cues. Whether this cell-to-cell heterogeneity, as
seen for the tissue-specific regulation of p97, results from regulation at the mRNA level is still unknown. However, the variable degree of p97 nuclear staining found in our studies, as well as independent work showing cell cycle regulation of p97/Cdc48p nuclear import (31) and inducible tyrosine phosphorylation of p97 (32), together point toward a significance for post-translational regulation in p97 function.

p97 Gene Structure and Identification of the Promoter—The molecular cloning and analysis of the p97 gene and regulatory sequences should facilitate dissection of the molecular basis for p97’s complex expression pattern and the functional relevance. Characterization of the genomic structure of the p97 locus demonstrated that the gene encoding the functional mRNA encompasses 20.4 kb and is interrupted by 16 introns. The two AAA domains of p97 are encoded within 6 (N-terminal AAA) and 4 exons (C-terminal AAA). The gene is flanked by a TATA-less/GC-rich upstream region and is likely to utilize an initiator element, spanning the mRNA start site, to direct basal transcription. Transient transfection assays in HeLa cells provide direct evidence that this region is sufficient for a relatively robust level of basal transcription. Cis-acting elements necessary for basal transcription activity seem to reside within 410 bp of the 5′-flanking region of the gene. Correspondingly, this upstream segment contains consensus binding sites for several transcriptional activators which have been implicated in basal gene transcription and also growth factor/cytokine, developmental, or stress-induced expression, including CBF, Sp1, AP-2, junB, Ets-1, and Y-box proteins.

Analysis of the phage clone encoding a second p97 gene, termed ϕp97, revealed that the region features all the hallmarks of a processed pseudogene, which are thought to evolve by random insertion of a reverse transcript of the mRNA into the genome (28). Since most processed pseudogenes derive from integration of reverse transcribed mRNA generated by RNA polymerase II, the sequence homology between pseudogenes and their functional counterpart ceases at the points corresponding to the beginning and the end of the transcript. However, ϕp97 contained an additional 15 bp of 5′-flanking genomic sequence that was identical to sequences found in the corresponding untranscribed region of the functional p97 gene counterpart. Other cases have been reported in which the homology between processed pseudogenes and their counterparts extends beyond the transcription start site, and it has been suggested that these pseudogenes may result from transcription via RNA polymerase III (28).

In contrast to Mus musculus, p97 is thought to be encoded by a single, intron-containing gene in humans and Mus spretus (27). Together with the lack of evidence of a function for pseudogenes and the apparent absence of a cDNA corresponding to the ϕp97 sequence, it therefore seems unlikely that ϕp97 is functionally significant. However, since the pseudogene has an intact open reading frame, and thus the potential to encode for a nearly intact p97 protein, further studies are required to confirm that ϕp97 is functionally inactive. We are currently in the initial stages of analyzing conditional gene disruption mutants of p97 that have been engineered in mouse embryonic stem cells. Our observations suggest that, similar to yeast, p97 is an essential gene in the mouse, thus demonstrating that the p97 gene and pseudogene are not functionally redundant members of a multi-gene family.

Functional Clues from p97 Expression in Vivo—The tissue-specific and heterogeneous expression of p97 in vivo provides evidence for a regulated or intermittent role for p97. Based on previous investigations, it could be suggested that the protein’s regulated role is in proliferating cells, particularly during mitosis. Yeast cdc48–1 is a mutant arresting in mitosis with undivided nuclei (7), and the CDC48 protein is a nuclear and peripheral endoplasmic reticulum/nuclear envelope protein, which has been speculated to regulate the nuclear fusion step required for the completion of mitosis (16). In mammalian systems, p97 has also been implicated in the fusion step required for reassembling mitotic Golgi fragments into stacks at telophase (4). Also, the results obtained from our analysis of p97 in embryonic mouse tissues and in work on developing Arabidopsis (13) are consistent with a role for p97 in expanding/proliferating tissues, and the heterogeneous p97 levels observed in the embryo could potentially reflect an increase in

---

3 J. M. Müller and D. T. Shima, unpublished observation.
4 J. M. Müller and D. T. Shima, unpublished observation.
p97 protein in actively cycling cells.

However, our data and recent data from other investigators do not support a sole function for p97 in dividing cells. p97 levels appeared to be most abundant in the majority of highly differentiated, non-proliferating cells throughout the adult. In fact, p97 was significantly less abundant within proliferative zones, including the crypts of the small intestine, the hair bulb, and the epidermis. Similarly, the Drosophila homologue of p97 is abundant in non-proliferating cells, predominantly in the nervous system and in the reproductive organs (12). Moreover, recent biochemical and genetic data point to a role for p97 in intracellular signaling (22) and targeted proteolysis (21), which are obviously prevalent in both proliferating and differentiated tissues. Finally, the findings presented here, as well as those discussed above emphasize the necessity of accounting for p97’s complex regulation in vivo in any future models of p97 function and provide the initial framework for a relatively unexplored experimental approach to address the cellular and organismal role(s) of this multifaceted protein.

Acknowledgments—We thank Jill Williams for fluorescence in situ hybridization analysis; Graham Clark and Andrew Davies for DNA sequencing assistance; George Elia and the ICRF Histology Unit for their contributions to the immunohistochemistry; Richard Poulson, Rosemary Jeffery, and Jan Longeroff for skilled assistance with the in situ hybridization studies; Richard Mortenson for providing the
Genomic Organization and Regulated Expression of Mouse p97

REFERENCES

1. Confalonieri, F., and Duguet, M. (1995) Bioessays 17, 639–650
2. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) EMBO J. 1, 945–951
3. Rothman, J. E. (1994) Nature 372, 55–63
4. Rabouille, C., Levine, T. P., Peters, J. M., and Warren, G. (1995) Cell 82, 905–914
5. Dubiel, W., Ferrell, K., and Rechsteiner, M. (1995) Mol. Biol. Rep. 1, 27–34
6. Tomoyasu, T., Yamanaka, K., Murata, K., Suzuki, T., Boulec, P., Kato, A., Niki, H., Hiraga, S., and Ogura, T. (1993) J. Bacteriol. 175, 1352–1357
7. Frohlich, K.-U., Fries, H.-W., Rudiger, M., Erdmann, R., Botstein, D., and Warren, G. (1995) J. Cell Biol. 114, 443–453
8. Shibuya, H., Irie, K., Nimoniyama-Tsuchi, J., Goebi, M., Taniguchi, T., and Matsumoto, K. (1992) Nature 357, 700–702
9. Koller, K. J., and Brownstein, M. J. (1987) Nature 325, 542–545
10. Peters, J. M., Walsh, M. J., and Franke, W. W. (1990) EMBO J. 9, 1757–1767
11. Mair, D., Stewart, S. E., Osmond, B. C., and Botstein, D. (1982) Genetics 100, 547–563
12. Finer, M., Jekely, G., Szepesi, R. J., Farkas, A., Theopold, U., Meyer, H. E., Lindholm, D., Nissel, D. R., Hultmark, D., and Friedrich, P. (1995) Biochim. Biophys. Acta 1255, 91–98
13. Feiler, H. S., Desprez, T., Santoni, V., Kronenberger, J., Caboche, M., and Traas, J. (1995) EMBO J. 14, 5626–5637
14. Confalonieri, F., Marsault, J., and Duguet, M. (1994) J. Mol. Biol. 235, 396–401
15. Zhang, L., Ashendel, C. L., Becker, G. W., and Morre, D. J. (1994) J. Cell Biol. 127, 1871–1883
16. Latterich, M., Frohlich, K. U., and Schekman, R. (1995) Cell 82, 885–893
17. Acharya, U., Jacobs, R., Peters, J. M., Watson, N., Farquhar, M. G., and Malhotra, V. (1995) Cell 82, 895–904
18. Kondo, H., Rabouille, C., Newman, R., Levine, T. P., Pappin, D., Freimont, P., and Warren, G. (1997) Nature 388, 75–78
19. Pleasure, I. T., Black, M. M., and Keen, J. H. (1993) Nature 365, 459–462
20. Ghislain, M., Dohmen, R. J., Levy, F., and Varshavsky, A. (1996) EMBO J. 15, 4884–4899
21. Dai, R. M., Chen, E., Longo, D. L., Garbea, C. M., and Li, C. C. (1998) J. Biol. Chem., 273, 3562–3573
22. Egeron, M., Ashe, O. R., Chen, D., Duker, B. J., Burgess, W. H., and Samuelson, L. E. (1992) EMBO J. 11, 3533–3540
23. Shima, T., Kuroki, M., Deutsch, U., Ng, Y. S., Adams, A. P., and IF, P. A. (1996) J. Biol. Chem. 271, 3877–3883
24. Berger, J., Hauber, J., Hauber, E., Geiger, R., and Cullen, B. R. (1988) Gene (Amst.) 66, 1–10
25. Senior, P. V., Critchley, D. R., Beck, F., Walker, R. A., and Varley, J. M. (1988) Development 104, 431–446
26. Poulsom, R., Longcroft, J. M., Jeffery, R. E., Rogers, L., and Steel, J. H. (1998) Eur. J. Histochem. 42, 121–132
27. Hoyle, J., Tan, K. H., and Fisher, E. M. (1997) Mamm. Genome 8, 778–780
28. Vanin, E. F. (1985) Annu. Rev. Genet. 19, 253–272
29. Richardson, M. P., Braybrook, C., Tham, M., Moore, G. E., and Stanier, P. (1998) Gene (Amst.) 206, 145–150
30. Azizkhan, J., Jensen, D. E., Pierre, A. J., and Wade, M. (1993) Crit. Rev. Eukaryotic Gene Exp. 3, 229–254
31. Madeo, F., Schlauer, J., Zischka, H., Mecke, D., and Frohlich, K. U. (1998) Mol. Biol. Cell 9, 131–141
32. Schulte, R. J., Campbell, M. A., Fischer, W. H., and Sefston, B. M. (1994) J. Immunol. 153, 5465–5472