Article / Book Information

( )	
Title(English)	Analysis of Genes for Intracellular Endo-Oligopeptidases, EP 24.15 (Thimet Oligopeptidase) and EP 24.16 (Neurolysin or Oligopeptidase M)
( )	
Author(English)	Akira Kato
( )	: , : , : 3756 , :1998 3 26 , : , :
Citation(English)	Degree:Doctor of Science, Conferring organization: Tokyo Institute of Technology, Report number: 3756 , Conferred date:1998/3/26, Degree Type:Course doctor, Examiner:
( )	
Type(English)	Doctoral Thesis

Analysis of Genes for Intracellular Endo-Oligopeptidases, EP 24.15 (Thimet Oligopeptidase) and EP 24.16 (Neurolysin or Oligopeptidase M)

Structures of the genes and targeting of EP 24.16 to different subcellular compartments by alternative usage of multiple promoters

Akira Kato

Graduate School of Bioscience and Biotechnology

Tokyo Institute of Technology

1998

Advisor: Professor Shigehisa Hirose

CONTENTS

ABBREVIATIONS

EP	endopeptidase
ER	endoplasmic reticulum
TOP	thimet oligopeptidase
MOP	oligopeptidase M
MHC	major histocompatibility complex
ABC	ATP-binding cassette
PSA	puromysin-sensitive aminopeptidase
PE	prolyl endopeptidase
peptidyl-NH-Mec	peptidyl-7-amino-4-methylcoumarin
SDS	sodium dodecyl sulfate
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
RACE	rapid amplification of cDNA ends
SINE(s)	short interspersed repetitive element(s)
bp	base pair(s)
nt	nucleotide(s)

GENERAL INTRODUCTIONS

Turnover of proteins in eukaryotic cells

Proteins are one of the major components of cells, and they are continuously turning over (1-3). The quality of proteins is maintained by degradation of abnormal proteins (4-7). The amounts of proteins are controlled by both synthesis and degradation. In mammals, turnover of proteins produces oligopeptides presented by MHC molecules, which help in the recognizition of self and non-self by the immune system (8-12). Depending on the protein, some are stable for several weeks, whereas others are rapidly degraded within a few minutes in cells. Rapid degradation of regulatory proteins is one of the most important ways to control cell function (3, 7). Hydrolysis of proteins are irreversible processes of protein maturation and degradation that are strictly controlled.

Cleavage of the peptide bonds of proteins or peptides are the reactions that are catalyzed by enzymes. Protease is a general term for enzymes which hydrolysze the peptide bond, and every protease that has been identified is a protein. The terms "proteinase" and "peptidase" are also commonly used for enzymes which hydrolyze proteins and peptides, respectively. Eukaryotic cells have a large variety of proteinases and peptidases, which constitute one of the largest families of proteins (13-17). Proteases show a wide variety of substrate specificities, localizations, and functions in organisms. In eukaryotic cells, there are organized systems of proteases to degrade intracellular proteins. These are classified into two categories: 1) lysosomal system and 2) non-lysosomal system.

Lysosomes are intracellular membrane compartments commonly found in animal cells (18, 19). Plants and yeast have an equivalent organelle called "vacuole". Lysosomes contain many kinds of acid hydrolases, such as proteases, nucleases, glycosidases, lipases, and so on. These acid hydrolases work only inside lysosomes, which maintain a low pH (about 5) in their interior. The substrates of hydrolases are taken in by fusion of membranes, then digested inside lysosomes. Many acidic proteinases and peptidases are enclosed in this organelle, e.g., such as cathepsins A, D,

E, B, H, K (O), L, S, W and X, legumain, and pro-X carboxypeptidase which degrade

substrate proteins to amino acids and oligopeptides (20-22). Some of these proteases are

ubiquitous and others are tissue specific (23-26). The lysosomal degradation of proteins

in macrophages of mammals produces class II MHC peptides, which are involved in

humoral imunity (11, 12). Lysosomes can digest proteins in cytosol, mitochondria,

intracellular membrane compartments, endosomes and phagosomes. A process whereby

endoplasmic reticulum (ER) encloses parts of cells (cytoplasm and organelles) and fuses

with lysosomes is called autophagy (27). Autophagy is one of the most important

processes in turnover of proteins and other components of the cell (Fig. 1). Most proteins

that have long half lives are digested by autophagy. The digestion of intracellular proteins

by autophagy is thought to be non-selective, and is regulated by the process of fusion of

intracellular membranes. Recent work has shown phosphorylation of regulatory protein

for membrane fusions control autophagy (28-31).

The major pathway for non-lysosomal degradation of intracellular proteins is the

ubiquitin-proteasome pathway (Fig. 1) (7, 32). Proteasome is an 26S complex of

multiple proteases (20S subunit termed 20S proteasome) and ATP hydrolases (19S

subunit termed PA700). Proteasomes are ubiquitously found in the cytosol and nucleus

of eukaryotic cells and can degrade proteins in cytosol, nucleus, and ER. The hydrolysis

of proteins by proteasomes is strictly regulated. The signal for the degradation of protein

by proteasome is poly-ubiquitination. Ubiquitin is a small (8.5 kDa), highly conserved

protein, widely found in eukaryotic cells. Ubiquitins are conjugated to free amino groups

of proteins through its C-terminal glycine residue. Ubiquitination of protein is strictly

regulated by ubiquitin-conjugating systems that involve cascades of reactions by ubiquitin-

activating (termed E1) enzyme, ubiquitin-conjugating (E2) enzymes, and ubiquitin ligases

(E2 or E3). These systems ubiquitinate short-lived or abnormal proteins, which are then

degraded by proteasomes (Fig. 1). The ubiquitin-proteasome pathway is the most

important pathway for maintaining quality of proteins and degrading short-lived proteins

which regulate cell functions, such as transcriptional regulators (33-35), cyclins (36, 37),

receptors (38-42), metabolic enzymes (43-45), and so on. This pathway also produces

3

oligopeptides presented by class I MHC molecules, which are involved in the cellular immunity system of mammals (8-10).
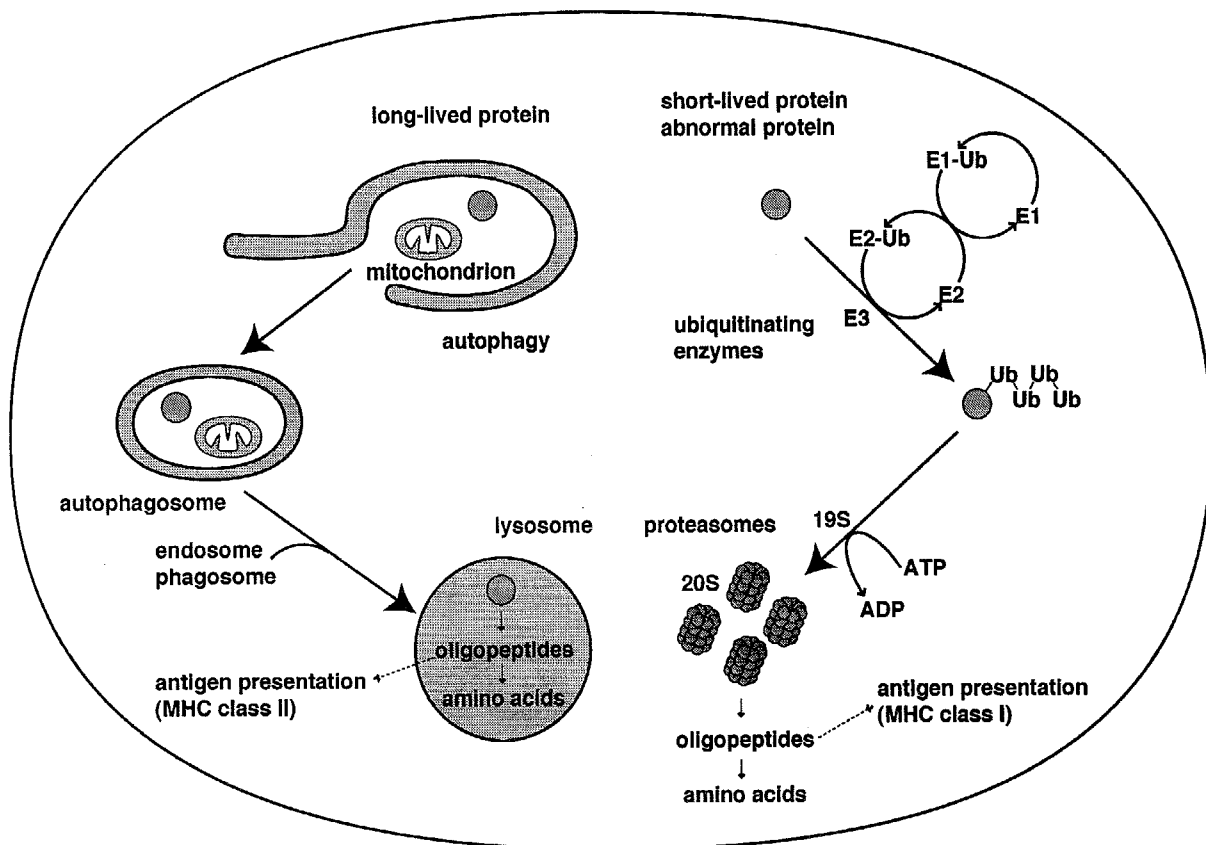

Fig. 1. **Intracellular protein turnover.** *Left,* the pathway metabolizing long-lived proteins by autophagy. *Right,* degradation of short-lived or abnormal proteins by the ubiquitine-proteasome system.

Turnover of oligopeptides in cytosol

Oligopeptides are the first products produced during degradation of proteins, and they are further hydrolyzed into amino acids by oligopeptidases. Oligopeptidases are a group of enzymes that act only on small peptides. The structure of oligopeptidases allows them to associate only with small substrates. Lysosomes contain a variety of endo- and exo-peptidases (cathepsins and other peptidases) that degrade peptides into amino acids. However, cytosolic systems for the degradation of peptides are not well understood.

The major source of oligopeptidases in the cytosol is the proteasome (32). Peptides are degraded into amino acids, or presented by MHC class I molecules after being transported by the ABC transporter from cytosol to ER (Fig. 2) (8, 10). Some reports describe identification of oligopeptidases in cytosolic fractions of mammalian cells. These include endopeptidase 24.15 (46), endopeptidase 24.16 (47-49), prolyl endopeptidase (50), puromycin-sensitive aminopeptidase (PSA) (51), aminopeptidase H (52), and acyl-peptide hydrolase (53) (Table 1). These peptidases are considered to be strong candidates for enzymes that degrade peptides derived from the ubiquitin/proteasome system because of their 1) solubilities and cytosolic localizations, 2) abilities to cleave short peptides, 3) widely distributed presence in many kinds of cells and tissues, and 4) high activities at neutral pH. The details for each enzyme are as follows:

Puromycin-sensitive aminopeptidase (PSA)— The enzymes that catalyze the final steps of degradation of peptides in the cytosol are considered to be aminopeptidases (51). Aminopeptidases hydrolyze N-terminal amino acids of oligopeptides. Puromycin-sensitive aminopeptidase (PSA) is a soluble monomeric metallopeptidase (100 kDa) homologous to membrane-type aminopeptidases such as aminopeptidase N and A, but is 100-fold more sensitive to puromycin than is aminopeptidase N. This enzyme shows broad specificity for amino-acyl-NHMec substrates with alanine, leucine, arginine, tyrosine, methionine, or lysine in P1 positions at neutral pH (54). Substrates with proline, glycine, or acyl-leucine at P1 are not cleaved or are slowly cleaved by PSA in vitro. At first, PSA was thought to degrade neuropeptides such as enkephalins. However, the function of PSA in neuropeptide metabolism has been questioned, since it

was found to be a cytosolic protein. In cytosol, PSA associates with the spindle apparatus during mitosis by its two internal microtubule-binding motifs, which show similarities to microtubule-associated proteins (MAP-2 and MAP-4), tau protein, and alpha-type proteasome subunits. Inhibition of PSA arrests the cell division cycle in G2/M, suggesting its essential role in mitosis (51), which includes degradation of peptides derived from proteasomes.

Aminopeptidase H— Aminopeptidase H, also called "bleomycin hydrolase", is a cysteine peptidase homologous to papain and cathepsins B, H, L, and S (52, 55). This 52 kDa protein forms a trisymmetrical homohexamer of 310 kDa, and is localized in the cytosolic/nuclear fraction. In vitro, aminopeptidase H hydrolyzes substrates containing citrulline, methionine, leucine, alanine, or glutamate, but is weakly or not active against valine, proline, and asparate in the P1 positions at neutral or slightly alkaline pH (52). This enzyme also degrades bleomycin, which is an antineoplastic glycopeptide antibiotic used for therapy against human tumors. Increase of aminopeptidase H in several tumors causes resistance of the tumors to bleomycin (56).

Acyl-peptide hydrolase (acylaminoacyl-peptidase)— Many intracellular proteins of eukaryotes are found to be NH_2-terminally acetylated (57). A serine-type oligopeptidase that catalyzes the removal of an $N(\alpha)$-acetylated residue from peptides is present in cytosol (53). This enzyme has a molecular mass of 300 kDa, consists of four subunits of 80 kDa, and is considered to play a central role in the degradation of $N(\alpha)$-acetylated peptides in cytosol.

Prolyl endopeptidase— Prolyl endopeptidase (PE) cleaves peptide bonds on the C-terminal side of prolyl residues in peptides that are up to approximately 30 amino acids long (50). PE and acyl-peptide hydrolase belong to a subfamily of serine proteases together with some other membrane-type peptidases. PE is considered to play a major role in the degradation of peptides containing proline residue(s) (58).

Endopeptidase 24.15 and endopeptidase 24.16— Two related endopeptidases are considerd to play roles in the degradation of peptides in cytosol. EP 24.15 and EP 24.16 are soluble Zn^{2+} metallopeptidases (78—80 kDa) that are active at neutral pH (46,

59). EP 24.15 is localized in cytosol and nucleus (46, 59, 60), and EP 24.16 is present in cytosol and the intermembrane space of mitochondria (49, 61). EP 24.15 and EP 24.16 hydrolyze oligopeptides 3—17 amino acids long at one or two position(s) (Table 2, (46, 62)). Although many peptides were tested to find rules for predicting the bonds cleaved by these peptidases, no rules were found (59, 62). However at least the size of the substrate seems to be important, because these peptidases do not act on oligopeptides more than 20 amino acids in length.

These peptidases were first found as enzymes that metabolize neuropeptides or peptide hormones (59). Most available data on these enzymes are derived from studies on the metabolism of bioactive peptides, although major amounts of these enzymes are present in intracellular compartments such as cytosol, nucleus, and mitochondria. Recent studies have focused on the importance of these enzymes in the metabolism of cytosolic peptides (46, 59, 61, 63). However, the roles of these peptidases in cytosol, nucleus, and mitochondria have not been established.

In this study, I report 1) isolation of cDNAs corresponding to six isoforms of EP 24.16 mRNA that have different 5' ends, 2) structures of porcine genes for EP 24.15 and EP 24.16, 3) a mechanism of generation of six mRNA isoforms from a single gene for EP 24.16 by alternative usage of multiple promoters and splicing, and 4) difference in subcellular localizations of the six EP 24.16 isoforms, in order to investigate the molecular bases of the intracellular functions of EP 24.15 and EP 24.16 in mammalian cells (64-67).

ER

CYTOSOL

proteasome

ABC transporter

oligopeptides

oligopeptides

oligopeptidases in ER

oligopeptidases in cytosol

oligopeptides

endopeptidase 24.15
endopeptidase 24.16
prolyl endopeptidase

puromycin-sensitive aminopeptidase
aminopeptidase H
acyl-peptide hydrolase

MHC class I

amino acids
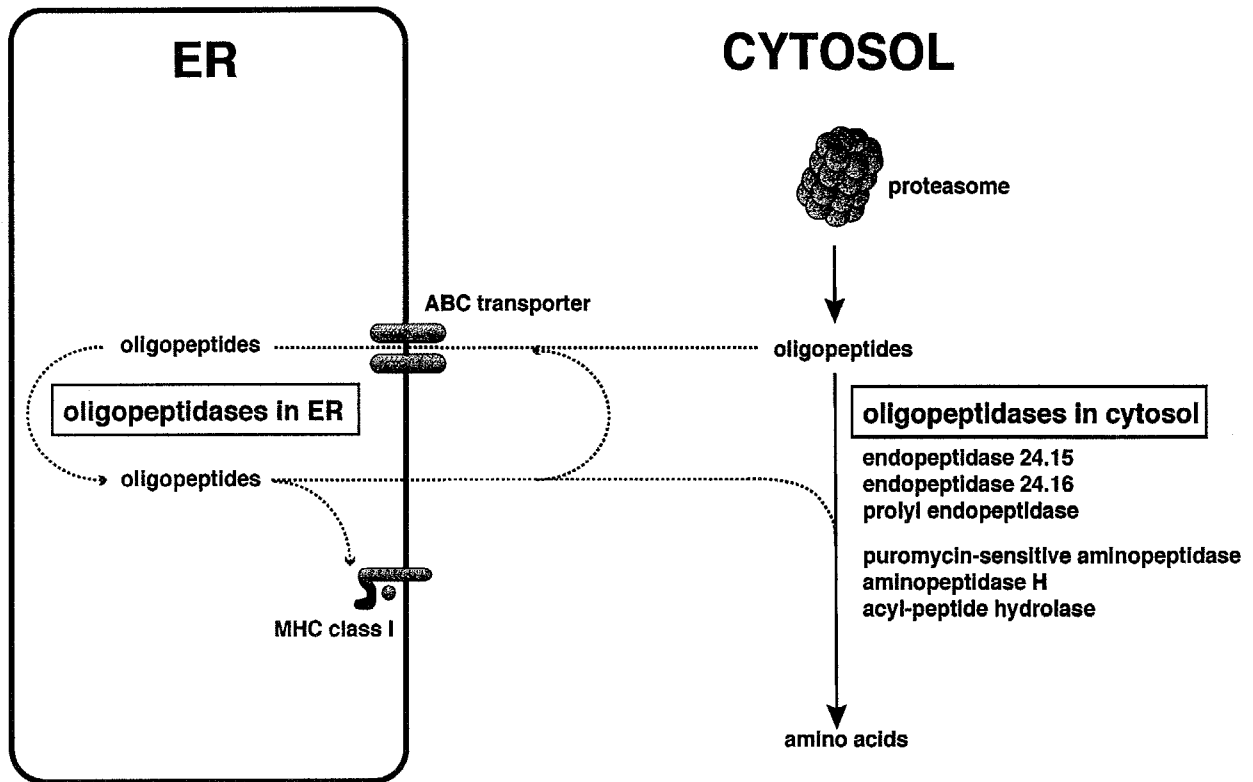
Fig. 2. **Metabolism of cytosolic peptides by oligopeptidases.** Cytosolic peptides are degraded into amino acids or processed into antigen peptides. ABC transporter transport peptides from cytosol to ER. The presence of a system which transport peptides from ER to cytosol has been proposed by Dr. Roelse (124) suggesting network of metabolism of peptides by oligopeptidases in cytosol and ER.

Table 1

Cytosolic peptidases for which structural and compositional data are available

Name	Source	Subunit mass (kDa)	No. of subunits	Classification	Intracellular localization	Refs
Proteasome					Cytosol, nucleous	
20 S proteasome		20—30	28	Threonine		32
PA700 (19S complex, ATPase)		28—112	≈20	—		32
PA28 (11S regulator)		27—29	2	—		32
Endo-peptidases						
Endopeptidase 24.15 (thimet oligopeptidase)	Testis (rat) Brain (rat)	78	1	Metallo (Zn^{2+})	Cytosol, nucleus (secreted)	46, 59
Endopeptidase 24.16 (neurolysin, oligopeptidase M)	Brain (rat) Liver (rat, pig, rabbit)	78	1	Metallo (Zn^{2+})	Cytosol, mitochondria (secreted)	46, 48
Prolyl endopeptidase	Kidney (pig)	75	1	Serine	Cytosol	50
Exo-peptidase						
Puromysin-sensitive aminopeptidase	Liver (rat)	100	1	Metallo (Zn^{2+})	Cytosol, nucleus	51, 54
Aminopeptidase H (Bleomycin hydrolase)	Liver (rabbit), lung (rabbit), skeletal muscle (chicken)	52	6	Cysteine	Cytosol, nucleus	52
Acyl peptide hydrolase	Liver (rat)	80	4	Serine	Cytosol	57

Table 2

Sites of hydrolysis of peptides by EP 24.15 (TOP) and EP 24.16 (MOP) (↓)

A. Cleavage positions that are common to both enzymes. *B*. Differences between the enzymes. *C*. Effect of the size of substrates on the cleavage by EP 24.15 (TOP). *D*. Effect of amidation of the C-terminus of GnRH. *E*. Effect of addition or substitution of glutamine residues at C-terminus of substrates. Abbreviations: Bz, benzoil; Dnp, 2,4-dinitrophenyl; Mcc, 7-methoxycomumarin-3-carboxylyl; MCA, 7-methoxycoumarin-4-yl; Dyn, dynorphin; BAM, bovine adrenal medula enkephalin peptide; βEnd, β-endorphin; GnRH, gonadotropin-releasing hormone; qf, quenched fluorescence; Abz, *o*-aminobenzoil; EDDnp, N-(2,4-dinitrophenyl)ethylenediamine; and Bk, bradykinin.

	EP 24.15 (TOP)	EP 24.16 (MOP)
A		
Pz-peptide	Pz-P-L- ↓ -G-P-ᴅ-R	Pz-P-L- ↓ -G-P-ᴅ-R
Orlowski substrate	Bz-G- ↓ -A-A-F-pAb	Bz-G- ↓ -A-A-F-pAb
QF01	Dnp-P-L- ↓ -G-P-W-ᴅ-K	Dnp-P-L- ↓ -G-P-W-ᴅ-K
QF02	Mcc-P-L- ↓ -G-P-K(Dnp)	Mcc-P-L- ↓ -G-P-K(Dnp)
Bradykinin	R-P-P-G-F- ↓ -S-P-F-R	R-P-P-G-F- ↓ -S-P-F-R
Dynorphin A (1–8)	Y-G-G-F-L- ↓ -R-R-I	Y-G-G-F-L- ↓ -R-R-I
B		
Neurotensin	Pca-L-Y-E-N-K-P-R- ↓ -R-P-Y-I-L	Pca-L-Y-E-N-K-P-R-R-P- ↓ -Y-I-L
Dynorphin A (1–17)	Y-G-G-F-L-R-R-I-R-P-K-L-K-W-D-N-Q (Resistant)	Y-G-G-F-L-R-R-I-R-P-K- ↓ -L- ↓ -K-W-D-N-Q
QF34	Mca-G-G-F-L- ↓ -R-R-A-K(Dnp)NH₂	Mca-G-G-F-L- ↓ -R- ↓ -R-A-K(Dnp)NH₂
QF37	Mca-G-G-F-I-R- ↓ -R-A-K(Dnp)NH₂	Mca-G-G-F-I-R-R- ↓ -A-K(Dnp)NH₂

EP 24.15 (TOP)	
C	
Dyn A (1–8)	Y-G-G-F-L- ↓ -R-R-I
Dyn A (1–11)	Y-G-G-F-L- ↓ -R-R-I-R-P-K
Dyn A (1–12)	Y-G-G-F-L-R-R-I-R-P-K-L (Resistant)
Dyn A (1–17)	Y-G-G-F-L-R-R-I-R-P-K-L-K-W-D-N-Q (Resistant)
BAM-12P	Y-G-G-F-M- ↓ -R-R-V-G-R-P-E
BAM-18P	Y-G-G-F-M-R-R-V-G-R-P-E-W-W-M-D-Y-Q (Resistant)
βEnd (1–9)	Y-G-G-F-M- ↓ -T-S-E-K
βEnd (1–16)	Y-G-G-F-M-T-S-E-L-K-S-Q-T-P-L-V-T-L (Resistant)
D	
GnRH Free Acid	pE-H-W-S-Y- ↓ -G-L-R-P-G
GnRH (1–9)	pE-H-W-S-Y- ↓ -G-L-R-P
GnRH	pE-H-W-S-Y-G-L-R-P-G-NH₂ (Resistant)
E	
qf-Dyn (2–8)	Abz-G-G-F-L- ↓ -R-R-V-EDDnp
qf-Q⁸Dyn (2–8)	Abz-G-G-F-L-R- ↓ -R-Q-EDDnp
qf-Bk	Abz-R-P-P-G-F- ↓ -S-P-F-R-EDDnp
qf-Bk-Q	Abz-R-P-P-G-F-S-P- ↓ -F-R-Q-EDDnp
qf-Bk (4–9)	Abz-G-F- ↓ -S-P-F-R-EDDnp
qf-Bk (4–9)-Q	Abz-G-F-S-P- ↓ -F-R-Q-EDDnp

A, B. Serizawa, A. *et al*. (1995) *J. Biol. Chem.* 270, 2092–2098; *C, E*. Camargo, A. C. M. *et al*. (1997) *Biochem. J.* 324, 517–522; *D*. Lew R. L. *et al*. (1995) *Bichem. Biophys. Res. Commun.* 209, 788–795.

10

SUMMARY

Endopeptidase 24.16 or mitochondrial oligopeptidase, abbreviated here as EP 24.16 (MOP), is a thiol- and metal-dependent oligopeptidase that is found in multiple intracellular compartments in mammalian cells. From an analysis of the corresponding gene, I found that the distribution of the enzyme to appropriate subcellular locations is achieved by the use of alternative sites for the initiation of transcription. The pig EP 24.16 (MOP) gene spans over 100 kb and is organized into 16 exons. The core protein sequence is encoded by exons 5—16 which match perfectly with exons 2—13 of the gene for endopeptidase 24.15, another member of the thimet oligopeptidase family. These two sets of 11 exons share the same splice sites, suggesting a common ancestor. Multiple species of mRNA for EP 24.16 (MOP) were detected by the 5' rapid amplification of cDNA ends (5'RACE) and they were shown to have been generated from a single gene by alternative choices of sites for the initiation of transcription and splicing. Two types of transcript were prepared, corresponding to transcription from distal and proximal sites. Their expression *in vitro* in COS-1 cells indicated that they encoded two isoforms (long and short) which differed only at their amino termini: the long form contained a cleavable mitochondrial targeting sequence and was directed to mitochondria; the short form, lacking such a signal sequence, remained in the cytosol. The complex structure of the EP 24.16 (MOP) gene thus allows, by alternative promoter usage, a fine transcriptional regulation of coordinate expression, in the different subcellular compartments, of the two isoforms arising from a single gene.

INTRODUCTION

Metalloendopeptidases form a large family of peptidases that have a His-Glu-X-X-His (HEXXH) zinc-binding motif and preferentially cleave short substrates. For example, endopeptidase 24.15 (EP 24.15), a member of this family, acts on peptides of 6—18 amino acid residues and exhibits no or only very weak proteolytic activity against proteins (Table 2, (13, 46, 59, 68)). Among the members of this family, thimet oligopeptidase (TOP or EP 24.15) and oligopeptidase M (MOP or EP 24.16) are unique in their sensitivities to thiol reagents and they constitute a subfamily, the thimet (thiol- and metal-dependent) oligopeptidase subfamily. Recent molecular cloning revealed the presence of a cysteine residue unique to members of this subfamily near position 483 (Fig. 3). This residue is absent from the other members that exhibit no thiol dependence (46, 69). In addition to the members of this family of mammalian origin, certain oligopeptidases of microbial origin that belong to this family have also been identified, including oligopeptidase A (OpdA) and dipeptidyl carboxypeptidase (Dcp) of *Escherichia coli* and *Salmonella typhimurium* (70), peptidase F of *Lactococcus lactis* (71), mitochondrial intermediate peptidase of rat and yeast (72-74), and saccharolysin (YCL57w or proteinase yscD) of yeast (75). This report deals with the two best-characterized mammalian enzymes, namely, EP 24.15 (TOP) and EP 24.16 (MOP), which are members of the thimet oligopeptidase family. This family has also been called the M3 family of metalloendopeptidases in the classification of Rawlings and Barrett (Table 3, (13, 17)).

EP 24.15 (TOP) was first identified as a collagenase-like peptidase or Pz-peptidase in experiments with the Pz-peptide that was originally designed by Wünsch and Heidrich (76) as a substrate for collagenase. Although the Pz-peptide was a good substrate for clostridial collagenase, it turned out not to be a substrate for avian and mammalian collagenases (77). The Pz-peptide-hydrolyzing activities found in avian and mammalian tissues have, therefore, been designated collagenase-like peptidases or simply Pz-peptidases. Independent studies on the metabolism of brain peptides led to the discovery

of two enzymes: one was described by Camargo *et al.* (78) in 1972 and was named

neutral endopeptidase and, later, endo-oligopeptidase A; and the other, first discribed by

Orlowski *et al.* (79) in 1983, was initially named soluble metalloendopeptidase and

subsequently endopeptidase 24.15. All these enzymes turned out to be the same and are

now known as thimet oligopeptidase (80). In this report I use the abbreviated designation

EP 24.15 (TOP). cDNA sequences for the mammalian enzyme are now available for the

rat (69, 81, 82), pig (64), and human (83) (Fig. 3).

EP 24.16 (MOP) was also discovered independently in several different

laboratories. 1) Heidrich *et al.* (84) demonstrated a Pz-peptide-hydrolyzing activity in a

mitochondrial fraction of rat liver, which was later shown to be distinct from EP 24.15

(TOP) by both biochemical characterization (85) and partial amino acid sequencing of the

purified enzyme; it was named oligopeptidase M (49). 2) Our laboratory (86, 87) and

Kiron and Soffer (88) identified a soluble angiotensin-binding protein in pig and rabbit

liver during the course of studies aimed at identifying hepatic receptors for angiotensin II.

After our publication of the cDNA sequence of the binding protein from pig (86), McKie

et al. (81) pointed out the strong similarity between our sequence and that of rat EP 24.15

(TOP) which had been determined by Pierotti *et al.* (69, 82). I then obtained a second

cDNA clone which was very similar to but clearly different from that of the cDNA for the

binding protein, and I showed that the second clone represented the pig homolog of rat EP

24.15 (TOP) (64). The angiotensin-binding protein, although originally identified as a

binding protein, did indeed have thiol- and metal-dependent oligopeptidase activity (64).

At that time, therefore, the binding protein appeared to represent a new member of the

thimet oligopeptidase family since the amino acid sequence of oligopeptidase M or EP

24.16 (MOP) from no mammalian spacies had yet been determined. 3) Kawabata *et al.*

(89-92) isolated an endopeptidase and the corresponding cDNA clone as a candidate for an

enzyme responsible for the post-transcriptional processing of γ-carboxyglutamic acid-

containing blood coagulation factors. They failed to notice the strong similarity to our

binding protein, which was later pointed out by McKie *et al.* (81). 4) Checler *et al.*

demonstrated the presence of a novel proteolytic activity capable of inactivating

neurotensin (93, 94). They purified the peptidase from rat brain synaptic membranes and characterized it (95). The enzyme, termed neurolysin or endopeptidase 24.16, was shown to be distinct from EP 24.15 (TOP) and neprilysin (also known as enkephalinase or endopeptidase 3.4.24.11) and to have a relatively broad substrate-specificity and tissue distribution. Recent determination of its amino acid sequence by cDNA cloning clearly indicated that neurolysin is identical to the three enzymes mentioned above (47). Thus, four saparate lines of research have converged in the discovery of a single new member of the thimet oligopeptidase family. In this report I use the abbrevation EP 24.16 (MOP) for this protein, whose identity has been only recently established.

EP 24.15 (TOP) and EP 24.16 (MOP) are very similar in terms of size and enzymatic properties: both are intracellular proteins of 78—80 kDa, consisting of about 680—700 amino acids, and their sequences are 65% homologous (Fig. 3) (64). They are, however, clearly distinguishable in several respects. For example, they have different specificities for inhibitors, different immunoreactivity, and different cleavage-site specificities. EP 24.15 (TOP) hydrolyzes neurotensin exclusively at the Arg-Arg bond whereas EP 24.16 (MOP) cleaves it at the Pro-Tyr bond (46, 49) (Table 2). Another difference is found in the subcellular localizations of these enzymes. EP 24.15 (TOP) is found in the cytosol while EP 24.16 (MOP) is found in both, the cytosolic and mitochondrial compartments (Table 4). How can the product of a single gene be localized to more than one intracellular compartment? To answer this question and to characterize evolutionary relationships among the members of the thimet oligopeptidase family, I investigated the structural organization of the pig genes for EP 24.15 (TOP) and EP 24.16 (MOP) and of their 5'-proximal flanking regions. I discovered six species of mRNA for EP 24.16 (MOP) that are generated from one single gene as a result of the utilization of alternative sites for the initiation of transcription. The six species of mRNA can be classified into two categories: those containing an additional sequence that encods a mitochondrial targeting sequence and those that lack such a sequence. The use of different promoters for the eventual targeting of proteins to appropriate subcellular compartments appears to be a useful mechanism for adjustment of local concentrations of proteins that

function at different intracellular sites in response to the physiological requirements of the cell.

Table 3

M3 family of metallopeptidases

Name	EC	Size (kDa)	Accession numbers	Locus	Refs.
Mitochondrial intermediate peptidase (MIP)	3.4.24.59				
Human		80	U80034	13q12	b
Rat		80	M96633		72
Saccharomyces cerevisiae		88	U10243	11	73
Schizophyllum commune		86	L43072		c
Schizosaccharomyces pombe		86	Z70690		a
Endopeptidase 24.15 (thimet oligopeptidase [TOP], endo-oligopeptidase A)	3.4.24.15 (3.4.22.19)				
Human		78	Z50115	19p13.3	83
Rat		78	M61142		69
Pig		78	D21871 (AB000426—38)		64, 66
Endopeptidase 24.16 (neurolysin, oligopeptidase M [MOP], microsomal metalloendopeptidase [MEP], soluble angiotensin-binding protein [sABP])	3.4.24.16				
Human		—	—	5q12–13	a
Rat		80	X87157		47
Rabbit		80	D13310		91
Pig		80	D11336 (AB000411—25)	16	86 66
Saccharolysin (yscD, YCL57w)	3.4.24.37				
Saccharomyces cerevisiae		82	X76504	3	75
MepB					
Aspergillus fumigatus	—	82	U85769		115
Oligopeptidase A (OPA)	3.4.24.70				
Escherichia coli		77	M93984		125
Salmonella typhimurium		77	M84574		70
Haemophilus influenzae		78	U32706		d
Synechocystis sp.		80	D90916		e
Peptidyl-dipeptidase (DCP)	3.4.15.5				
Escherichia coli		77	X57947		126
Salmonella typhimurium		77	M84575		70
pepF	—				
Lactococcus lactis		69	Z32522, X99710		71
Mycoplasma genitalium		71	U39695		f
Mycoplasma pneumoniae		71	U34795		g

[a] Direct submissions; [b] Chew, A. *et al.* (1997) *Genomics* **40**, 493-96; [c] Isaya G. *et al.* (1995) *Genomics* **28**, 450-61; [d] Fleischmann, R. D. *et al.* (1995) *Science* **269**, 496-512; [e] Peterson, S. N. *et al.* (1993) *J. Bacteriol.* **175**, 7918-30; [f] Kaneko, T. *et al.* (1996) *DNA Res* **3**, 109-36; [g] Herrmann, R. *et al.* (1996) *Nucleic Acids Res.* **24**, 4420-49

```
TOP (human)    1                           MKPPAACAGDMADAASPCSVV--NDLRWDLSAQQIEERIRELIEQTKRVYDQVGTQEFEDV-SYESTL    65
TOP (rat)      1                           MKPPAACAGDVVDTVSPCSTV--NHLRWDLSAQQTRALTTQIIEQTKCVYDRVCAQDFEDV-SYESTI    65
TOP (pig)      1                           MKPPAACAGDALDVAAPCSAV--NYLRWDLSAQQTGELTTELIEQTKRVYDRVCTQELQDV-SYENTI    65
MOP (pig)      1   MIVRCLSAARRLHRVGGSGILLRMTLGREAMSPLQAMSSYTVDGRNVLRWDLSPEQIKRRIEELIAQTKQVYDDIGMLDIEEV-TYENCL    89
MOP (rat)      1   MITLCLSTLRGLHRAGGSRLQLTMTLGKELASPLQAMSSYTAAGRNVLRWDLSPEQIKTRIEQLIAQTKQVYDTVCTIALKEV-TYENCI    89
MOP (rabbit)   1   MIARCFSAVRGLHRVGGSRILFKMTLGREVMSPLQAVSSYTAAGRNVLRWDLSPEQIKTRIEELIAQTKQVYDSVSVGMLDIKDV-TYENCL    89
Saccharolysin  1           MRLLLCKNWFASPVISPLLY-TRSLYSMANTTSFPIAP-QAPPNWSFTPSDISGKINEIINNSNNEYDSMSKVESPSVSNFVEPF    83

TOP (human)    66  KALADVEVTYIVQRNILDFPQHVSPSKDIRTASTEADKKLSEFDVEMSMREDVYQRIVWL-QE-KVQKDSLRPEAARYLERLIKLGRRNG   153
TOP (rat)      66  KALADVEVTYIVQRNILDFPQHVSPNKDIRAASTEADKKLSEFDVEMSMRQDVYQRVVWI-QE-KIPKDSLKPFAARYLFRLIKLGRRNG   153
TOP (pig)      66  KALADVEVSYIVQRNILDFPQHVSPCKDIRTASTEADKKLSETDVFMSMRQDVYQRIVWI-QE-KVQKDSLRPFAARYLERLIKLGRRNG   153
MOP (pig)      90  QALADVEVKYIVERTMLDFPQHVSSDKEVRAASTEADKRLSRFDIEMSMREDIFLRIVRL-KE-TCDLGKIKPEARRYLEKSVKMGKRNG   177
MOP (rat)      90  QVLADIEVTYIVERTMLDFPQHVSSDREVRAASTEADKKLSRFDIEMSMREDVFQRIVHI-QE-TCDLEKIKPEARRYLEKSIKMGKRNG   177
MOP (rabbit)   90  QALADVEVKYIVERTMLDFPQHVSTDREVRAASTEADKRLSRFDIEMSMREDIFQRIVHL-QE-TCDLEKIKPEARRYLEKSVKMGRRNG   177
Saccharolysin  84  MKF-ENLLG-PII-NQLTPLQHVSSDKEIRDASVNSSMKLDELNIDISIRHDIFLQFARVWQDVQSKADSVERETFKYVEKSYKDYIHSG   170

TOP (human)    154 LHIPRETQENIKRIKKKLSILCIDFNKNLNEDTIFIPFTLQELGGLP-----EDF-LNSLEKMEDGK-IKVTLKYPIYFPLLKKCHVPETR   237
TOP (rat)      154 LHIPQDTQEKIKNIKKRLSILCIDFNKNLNEDTIFIPFTREELGGLP-----EDF-LNSLEKTEDGK-IKVTLKYPHYFPLLKKCHVPETR   237
TOP (pig)      154 LHIPKETQEKIKSIKKKLSILCIDFNKNLNEDTIFIPVTREELGGLP-----EDF-LNSLFKTEDEK-IKVTLKYPHYFPLLKKCHVPETR   237
MOP (pig)      178 LHIPEQVQNEIKAMKKRMSELCIDFNKNLNEDDIFLVFSKAELGALP-----DDF-IDSLEKTDDNK-YKITLKYPHYFEVMKKCCIPETR   261
MOP (rat)      178 LHISEHIRNEIKSMKKRMSELCIDFNKNLNEDDTSLVFSKAELGALP-----DDF-IDSLEKTDEDK-YKVTLKYPHYFPVMKKCCVPETR   261
MOP (rabbit)   178 LHLPEEVQNEIKSMKKRMSELCIDFNKNLNEDDIFLVFSKAELGALP----DDF-IDSLEKMDDDK-YKIILKYPHYFPVMKKCCIPETR   261
Saccharolysin  171 LFLDEGNRLKIKEIKKKISVNSINFSKNIGEQKEYITFTKEQLEGVLPDSILTQFETIKSDKDSNEILYKVIFKYEDIPPVMKLASSAQTR   260

TOP (human)    238 RKVBEAFNCRCKEENCAIIKELVTLRAQKSRLLGFHTHADYVLBMNMAKTSQTVATFLDELAQKLKPLGEQERAVIIELKRAECERRCLP   327
TOP (rat)      238 RLLFEAFNCRCKEENCAIIKEIVSLRAQKSNLLGFRTHADYVIPMNMAKTSQTVATFLDEIARKLKPIGFQERAVIIELKEAFSAKRCLP   327
TOP (pig)      238 RKVBEAFNCRCKEENCAIIREIVRLRAQKSSLLGFSTHADYVLDMNMAKTSQVVATFLDEIAQKLKPLGFQERAVIIELKKAECTKRCLD   327
MOP (pig)      262 RKMBMAFNTRCKEENTIILQELLPLRAKVAKLLGYSTHADFVLBMNTAKSTHHVTAFLDDLSQKLKPLGBAEREFILNLKKKECEEKGFE   351
MOP (rat)      262 RKMBMAFHTRCKCENTAIILQQILPLRAQVAKLLGYNTHADFVLBLNTAKSTSRVAAFLDDLSQKLKPLGBAEREFIISLKKKECEERCFE   351
MOP (rabbit)   262 RRMBMAFNTRCKEENTVILQQILPLRAQVAKLLGYSTHADFVLBMNTAKSTSRVTAFLDDLSQKLKPLGBAEREFILSLKKKECEEKGFE   351
Saccharolysin  261 KQAFIADQNK-VPENEAIILDTLKLRDELASLLGYDIYANYNIYDKMAEDSTTVMNFLNDLKDKLTPLGRKELQVLQDMKAEDVKKLNQG   349

TOP (human)    328 FDGRIRAWDMRYYMNQVEETRYCVDQNLLIKEYFPVQVVIHGLIGIYQELLGLAHHHEEGA---SAWHEDV-RIL-YTARDAASG-EVVCKF   411
TOP (rat)      328 FDGRIHAWDMRYYMNQVEEDSYRVDQNLLIKEYFPMQVVIRGLIAIYQELLGITFTLEEGA---AAWHEDV-RIL-YSVRDAASG-EEICKF   411
TOP (pig)      328 FDGRINAWDMRYYMNQVFFTRYRVDQNLLIKEYFPMQVVTRGLIGIYQELLGLTDHLEEGA---AVWHEDV-MIL-YSVRDAASG-KVICKT   411
MOP (pig)      352 YDGKINAWDLHYYMTQTEBLKYSVDQEIIKEYFPIEVVIEGLLNIYQELLGLSBEQVTDA---HVWNKSV-TLL-YTVKDKATG-EVLGQF   435
MOP (rat)      352 YDGKINAWDLEYYMTQTEBLKYSVDQESLKEYFPIEVVTEGLISIYQELLGISBEQVPDA---HVWNKSV-SIL-YTVKDKATG-EVLGQF   435
MOP (rabbit)   352 YDGKINAWDLHYYMTQTEBLKYSIDQEFIKEYFPIEVVIEGLLNIYQELLGLSBEQVADA---HVWNPSV-TLL-YTVKDKATG-EVLGQF   435
Saccharolysin  350 ADPNYYIWDHRYYDNKYLLENFNVLEKISEYFPLEATITGMIBIYETIFNIKEIETKDSQNKSVWHDDVKQIAVWNMDDPKSPNFVGWI   439
                                                                        *    *        ↓
TOP (human)    412 YLDLYPREGKYGHAACFGLQPGCIRQDGSRQIAIAAMVANFTKFTADAPSLLQHDEVETYFHEFGHVMHQLCSCAEFAMFSG-THVERDF   500
TOP (rat)      412 YLDLYPREGKYGHAACFGLQPGCIRQDGSRQLAIAAMVANFTKFTPDVPSLLQHDEVETYFHEFGHVMHQLCSCAEFAMFSG-THVFRDF   500
TOP (pig)      412 YLDLYPREGKYGHAACFGLQPGCIRQDGSRQIAIAAMVANFTKFTPDAPSLLQHDEVETYFHEFGHVMHQLCSCAEFAMFSG-THVRRDF   500
MOP (pig)      436 YLDLYPREGKYGHAACFGLQPGCILPDGSRMMSVAALVVNFSQRAGRPSLLRHDEVRTYFHEFGHVMHQICACTDFARFSG-TNVETDF   524
MOP (rat)      436 YLDLYPRFGKYNHAACFGIQPGCILPDGSRMMSVAALVVNFSQFVAGRPSLLRHDEVRTYFHEFGHVMHQICACTDFARFSG-TNVETDF   524
MOP (rabbit)   436 YLDLYPREGKYNHAACFGLQPGCILPDGSRMLSVAALVVNFSQFVAGRPSLLRIIDEVRTYFHEFGHVMHQICACTDFARFSG-TNVETDF   524
Saccharolysin  440 YFDLFPRDGKYGHAANFGLSSSFMIDDTTRSYPVTAIVCNESKSTKDKPSLLKINEIVIFFHBLGHGIHDLVGCNKESRENCPGSVPWDF   529

TOP (human)    501 VEAPSQMLENWVWEQEFLLRMSRHYRTGSAVPRELLEKIIESRQANIGLFNLRQIVLAKVDQALIIIQTDADPABEY-ARIC-QEILGVPA   588
TOP (rat)      501 VEAPSQMLENWVWEKEFIMRMSQHYRTGGEAPEDLLEKTIIKSRQANAGIFNLRQTVIAKVDQVIHTQTDVDPABEY-ARIC-QEILGVPA   588
TOP (pig)      501 VEAPSQMLENWVWEAEFLLRMSQHYRTGSAIPQBLLEKLIKSRQANTGLFNLROIVIAKVDQALHIQTAADPABEY-ARIC-QEILGVPA   588
MOP (pig)      525 VEVPSQMLENWVWDTDSIRRLSKHYKDCGSPITDDLLEKLVASRLVNIGLLTLRQIVLSKVDQSLHINTSLDAASEY-AKYC-TEILGVAA   612
MOP (rat)      525 VEVPSQMLENWVWDVDSIRKLSKHYKDGHPITDELLEKLVASRLVNIGLLTLRQIVISKVDQSLHINATLDAASFY-AKYC-TEILGVAA   612
MOP (rabbit)   525 VEVPSQMLENWVWDIDSLRRLSKHYKDGNPIADDLLEKLVASRLVNIGLLTLRQIVISKVDQSLHINSSLDAASEY-ARYC-TDILGVAA   612
Saccharolysin  530 VEAPSQMLFFWINKNEDINLSSHYKTGEKIPESLINSIIKTKHVNGAIFTLRQLHFGLFDMKVHTCKDLQNLSICDTWNQLRQDISLIS   619

TOP (human)    589 TPGT-NMP-ATFGH-LAGGYDAQYYGYLWSEVYSMDMFHTRFKQEGVLNSKVGMDYRSCLLRPGGSELASAMLRRELGRDEKQDAFIISK   675
TOP (rat)      589 TPGT-NMP-ATFGH-LAGGYDAQYYGYLWSEVYSMDMFHTRFKQEGVLSPKVGMDYRTSIIRPGGSELASTMIKQFIGRDEKQDAFILSK   675
TOP (pig)      589 TPGT-NMP-ATFGH-LAGGYDAQYYGYLWSEVYSADMFHTRFKQEGILSGKVGMDYRSCILRPGGSELASVMLKLFLGRDEKQDAFIISK   675
MOP (pig)      613 TPGT-NMP-ATFGH-LAGGYDGQYYGYLWSEVFSMDMFYSCFKKEGIMNPEVGMKYRNLILKPGGSLDGMDMLQNFIKREPNQKAFIMSR   699
MOP (rat)      613 TPGT-NMP-ATFGH-LAGGYDGQYYGYLWSEVFSMDMFHSCFKKEGIMNPEVGIKYRNLILKPGGSLDGMDMIQNFIQREPNQKAFIMSR   699
MOP (rabbit)   613 TPGT-NMP-ATFGH-LAGGYDGQYYGYLWSEVFSMDMFYSCFKKEGIMNPEVGIKYRNLILRPGGSIDGMDMLQNFLQRBPNQKAFIMSR   699
Saccharolysin  620 NGGILSKGYDSFGHIMSDSYSAGYYGYLWAEVFAIDMYHTKBAKDPLNAKN-GIQYRDIVIARGGLYLINDNIKEPIGREPSKDAFIKEL   708

TOP (human)    676 GIQVGGCEPEPQVC   689
TOP (rat)      676 GIQVEGCEP-P-AC   687
TOP (pig)      676 GIQVEGCEP-PAS   687
MOP (pig)      700 GIHAP   704
MOP (rat)      700 GING    704
MOP (rabbit)   700 GIQAP   704
Saccharolysin  709 GIQN   712
```

Fig. 3. **Alignment of amino-acid sequences of EP 24.15 (TOP), EP 24.16 (MOP) and saccharolysin.** Amino acids are shown in single-letter code. Identical amino acids in EP 24.15 (TOP) and EP 24.16 (MOP) are shown in *black boxes*. Histidine residues involved in coordinated zinc binding are indicated by asterisks (*). Cys 483 of EP 24.15 (TOP) is indicated by an arrow.

16

EXPERIMENTAL PROCEDURES

Identification of 5'-Terminal Sequences of cDNAs for EP 24.16 (MOP) — The 5'-ends of cDNAs for EP 24.16 (MOP) were cloned with 5'RACE (rapid amplification of cDNA ends) system (Clontech, Palo Alto, CA). Two μg of poly(A)$^+$RNA, isolated from pig liver (86), were reverse-transcribed with a specific primer for the cDNA for pig EP 24.16 (MOP), 5RA-1 (5'-GTCTAGCATGGTTCGTTCC-3'), and AMV reverse transcriptase. The first-strand cDNA was ligated at the 3' end with an anchor (5'-CACGAATTCACTATCGATTCTGGAACCTTCAGAGG-3') by T4 RNA ligase. A nested specific primer for the cDNA for EP 24.16 (MOP), 5RA-2 (5'-CCGTCTACACCTTCACTTC-3'), was used with an anchor primer (5'-CTGGTTCGGCCCACCTCTGAAGGTTCCAGAATCGATAG-3') for amplification of the 5' ends of the cDNAs by PCR. The products of PCR were fractionated on a 3% agarose gel, and fragments of 300—650 bp were isolated and cloned into pBluescript II (Stratagene, La Jolla, CA). Positive clones were identified by colony hybridization, with the ^{32}P-labeled *Eco*RI-*Eco*RV 592-bp fragment of PAB-L1 (86) as probe, and sequenced.

Sequencing of DNA— DNA was sequenced by the dideoxy chain-termination method of Sanger *et al.* (96) with double-stranded plasmids as templates. Termination reactions were performed with SequiTherm DNA polymerase (Epicentre Technologies, Madison, WI) and IRD41-labeled M13 universal or reverse primer (LI-COR, Lincoln, NE). The products were analyzed with a DNA sequencer (model 4000; LI-COR). Sequences were organized and analyzed with GENETYX-MAC program (Software Development, Tokyo, Japan).

Isolation of Genomic Clones for Pig EP 24.16 (MOP) and EP 24.15 (TOP) — A pig liver genomic library constructed in λEMBL3 SP6/T7 (Clontech) was screened with the 2.7-kb *Eco*RI-*Eco*RI fragment of a cDNA for EP 24.16 (MOP) clone (PAB-L1); (86)

17

or with the 2.5-kp *Eco*RI-*Eco*RI fragment of a cDNA clone for EP 24.15 (TOP) (PABH-L7); (64), both of which had been labeled with [α-^{32}P] dCTP (Amersham, Little Chalfont, UK) with a random priming kit (Takara, Kyoto, Japan). Phage clones (2 × 10^6) were plated at a density of 30,000 pfu per 135 × 95 mm plate on *Escherichia coli* NM538, from which duplicate replications were made on cellulose-nitrate filters (Schleicher & Schuell, Dassel, Germany) and allowed to hybridiz with the ^{32}P-labeled probe in a solution of 6× SSPE (1× SSPE is 0.15 M NaCl, 15 mM NaH$_2$PO$_4$, pH 7.0, 1 mM EDTA), 50% formamide, 0.1% SDS, and 5× Denhardt's solution at 42°C for 16 hr. The filters were rinsed twice at room temperature in 2× SSC (1× SSC is 0.15 M NaCl, 15 mM sodium citrate, pH 7.0) that contained 0.1% SDS and washed twice at 60°C in 1× SSC that contained 0.1% SDS for 1 hr. Positive plaques were identified by autoradiography and purified by the additional rounds of screening.

Restriction Mapping of λ *Phages* — Positions of the *Eco*RI, *Sac*I, and *Xba*I restriction sites in genomic clones were determined by complete or partial digestion with restriction enzymes and subsequent Southern blot analysis. UV irradiation and formation of pyrimidine dimers were used for preparation of incompletely digested genomic clones. λEMBL3 SP6/T7 contains two unique *Sfi*I or *Sal*I sites and bacteriophage promoters (SP6 and T7) that flank the insert. Arms were separated with *Sfi*I or *Sal*I from the inserts, which still contained promoter sequences at the both ends. DNA samples were UV-irradiated for 0 or 20 min with UV Stratalinker 2400 (Stratagene) in 10 mM Tris, pH 7.5, 10 mM MgCl$_2$, and 1 mM dithiothreitol. UV-irradiated samples (500 ng) were digested incompletely with *Eco*RI, *Sac*I, or *Xba*I (10 U) for 1 hr at 37°C, fractionated on a 0.7% agarose gel, and transferred to nylon membranes (Magnagraph; MSI, Westboro, MA). A set of filters was prepared and allowed to hybridize with end-labeled oligoprobes for T7 or SP6 promoter sequence, for 14 hr at 37°C in 6× SSC, 5× Denhardt's solution, 0.5% SDS and 100 μg/ml herring sperm DNA. The filters were washed twice in 1× SSC, 0.1% SDS at 42°C for 30 min, exposed to imaging plates, and analyzed with a Bioimage Analyzer (model BAS 2000; Fuji Film, Tokyo, Japan).

S1 Nuclease Protection Assay — Three primers, namely, #108L

(CAAGCCTTGCGGCGGCCTAGCAAAGGAGGCAACAG) for exon 1, #107L

(GGTGTCCCTCGGGGTAGACCATGTGGGCTGTAGAA) for exon 2, and #106b

(GTCTCTCCATGAGAATGCTCCT) for exon 3, were designed for the synthesis of

single-stranded antisense DNA probes that would protect pig 5'-ends of mRNAs for EP

24.16 (MOP). Ten pmols of each primer were labeled with $[\alpha\text{-}^{32}P]ATP$ (Amersham) by

polynucleotide kinase (Takara) and used for the synthesis of probes. End-labeled primers

were annealed with 5 μg of plasmid DNA that contained genomic fragments of the pig

gene for EP 24.16 (MOP) (*Apa*I-*Xho*I 837-bp fragment of λPAB-G33 for exons 1 and 2;

*Bgl*II-*Eco*RI 923-bp fragment of λPAB-G32 for exon 3), and antisense probes were

synthesized with T7 DNA polymerase (Pharmacia, Uppsala, Sweden). The 3' ends of the

probes were digested with restriction enzymes (*Sma*I for exon 1, *Bss*HII for exon 2, and

*Bgl*II for exon 3), fractionated by electrophoresis on a 5% polyacrylamide gel that

contained 7 M urea and exposed to X-ray film. Probes were detected as bands of the

expected mobility and extracted in 0.5 M ammonium acetate, 10 mM magnesium acetate, 1

mM EDTA, 0.1% SDS and 10 μg/ml yeast tRNA at 37°C for 12 hr. Extracted probes

were precipitated in ethanol, and probes (1×10^5 cpm each) were annealed with 5 μg of

poly(A)$^+$RNA from pig liver or with 10 μg of yeast tRNA, as a control, for 12 hr at

30°C in 80% formamide, 40 mM PIPES, pH 6.4, 1 mM EDTA and 400 mM NaCl. Non-

annealed nucleic acids were digested with S1 nuclease (Boehringer Mannheim,

Mennheim, Germany) at a final concentration of 1,000 U/ml in 0.28 M NaCl, 0.05 M

sodium acetate, pH 4.5, 4.5 mM $ZnSO_4$ and 20 μg/ml denatured herring sperm DNA.

The protected fragments were purified by extraction with phenol/chloroform and

precipitation with ethanol, and electrophoresed in 5% polyacrylamide gels containing 7 M

urea. Gels were dried and exposed to imaging plates for 48 hr. Images were analyzed

with the Bioimage Analyzer.

Production of Antiserum against Purified Pig EP 24.16 (MOP) — Rabbits

were injected subcutaneously with 75 μg of purified pig EP 24.16 (MOP) [formerly

referred to as soluble angiotensin-binding protein, sABP (86)] in complete Freund's

adjuvant. Booster injections with 75 μg of purified protein in incomplete Freund's

adjuvant were given 2, 4 and 6 weeks after the initial injection. Rabbits were bled 10 days

after the fourth injection.

Construction and Expression of cDNAs for Isoforms of Pig EP 24.16 (MOP)

— Six plasmids, pcDNA3-MOP1 (exon 1-[5—16]), -MOP1' (exon 1-4-[5—16]),

-MOP2 (exon 2-[5—16]), -MOP2' (exon 2-4-[5—16]), -MOP3 (exon 3-[5—16]), and

-MOP3' (exon 3-4-[5—16]), were constructed for expression analysis. For pcDNA-

MOP1, a 2732-bp *Eco*RI-*Eco*RI fragment of PAB-L1 (86), which contained the entire

open reading frame of type 1 cDNA for EP 24.16 (MOP), was subcloned to pcDNA3

(Invitrogen, San Diego, CA). For the other plasmids, PAB-R5, -R302, -R305, -R8, and

-R1 (Fig. 4), which encoded only the 5' ends of type 1', 2, 2', 3, and 3' cDNAs,

respectively, were digested with *Alw*NI at their 3' termini, ligated with the 2291-bp

*Alw*NI-*Eco*RI fragment of PAB-L1 and subcloned into pcDNA3. COS-1 cells were

maintained in DMEM (Gibco/BRL, Gaithersburg, MD) that contained 10 mM HEPES,

pH 7.2, 10% fetal bovine serum, 50 units/ml penicillin and 50 μg/ml streptomycin, in a

controlled atmosphere of 5% CO_2 in air at 37°C. Approximately 6 × 10^6 cells were

electroporated with 20 μg of each plasmid at 220 V at a capacitance setting of 960 μF in a

Gene Pulser apparatus (Bio-Rad, Hercules, CA) and harvested 48 hr after electroporation.

Subcellular Fractionation of Cells and Western Blotting — All steps were

performed at 4°C. Cells were washed by centrifugation in Dulbecco's phosphate-buffered

saline (2.7 mM KCl, 138 mM NaCl, 1.2 mM KH_2PO_4, and 8.1 mM Na_2HPO_4, pH 7.4)

at 700 rpm for 2 min. Approximately 2 × 10^7 cells were suspended in 1 ml of 2.5 M

sucrose and homogenized for 2 min. Nuclear fractions were removed by centrifugation at

3,000 rpm (700 × *g*) for 10 min, and supernatants were centrifuged at 9,200 rpm (7,000 ×

g) for 10 min to recover mitochondrial fractions as pellets. Mitochondrial fractions were washed twice by centrifugation at 25,000 rpm (24,000 × *g*) for 10 min. The post-mitochondrial supernatants were centrifuged at 50,000 rpm (105,000 × *g*) for 100 min, and the pellets (microsomes) and supernatants (cytosol) were recovered. The concentration of protein in each fraction was determined with the BCA protein assay reagent (Pierce, Rockford, IL). Five μg of each protein sample were fractionated by SDS-PAGE (10% polyacrylamide) in standard glycine running buffer (192 mM glycine, 25 mM Tris, and 0.1% SDS) or high-resolution running buffer (492 mM glycine, 75 mM Tris, and 0.1% SDS). The separated proteins were transferred to a PVDF membrane (ATTO, Tokyo, Japan) and probed with 2,000-fold diluted rabbit antiserum against pig EP 24.16 (MOP). Bound antibodis were detected with alkaline phosphatase-conjugated second antibodies, with 4-nitro blue tetrazolium chloride and 5-bromo-4-chloro-3-indolyl phosphate (NBT/BCIP) as chromogen.

RESULTS

Differences among 5'-Terminal Sequences of mRNAs for EP 24.16 (MOP) —

To delineate the complete structure of the gene for EP 24.16 (MOP), I determined the 5'

end of the corresponding mRNA by rapid amplification of 5' ends of cDNA (5'-RACE)

using preparations of poly(A)$^+$RNA from pig liver. More than five prominent bands of

fragments of 340—630 nucleotides were obtained. DNA sequencing of these fragments

revealed the presence of several mRNAs whose sequences were completely different from

others starting 28 nt upstream from the ATG initiation codon (Fig. 4; M3). These results

suggested that usage of alternative promoters and exons might be involved in the

generation of the observed mRNA diversity. To determine the precise molecular

mechanism responsible for generation of such heterologous mRNAs, I isolated and

characterized the pig gene for EP 24.16 (MOP), which had previously been shown to be

present as only a single copy (86).

Structural Organization and 5'-Untranslated Exons of EP 24.16 (MOP) Gene

— To isolate the pig gene for EP 24.16 (MOP), I screened approximately 2 × 10^6

independent plaques of a pig genomic library in λEMBL3 (Clontech) using the PAB-L1

cDNA clone (86) as the probe. I isolated and mapped more than 50 clones, and then I

subcloned and sequenced the phage fragments for the identification of exons. The exon-

intron organization of the pig gene for EP 24.16 (MOP) was deduced from an analysis of

11 independent clones, each of which contained part of the gene (Fig. 5). The gene

extends over 100 kb and contains 16 exons and 15 introns of various sizes (Fig. 5*B* and

7B). All the introns have typical splice donor and acceptor boundaries (Fig. 7*D*) (97).

Comparison of the nucleotide sequences of the genomic clones λPAB-G32 and

λPAB-G33 (Fig. 5*D*) with those of the products of 5'-RACE (Fig. 4) allowed us to

identify the alternatively spliced leader exons (Fig. 5*A*, 5*B*, and 5*A*). Six distinct species

of mRNA for EP 24.16 (MOP) appeared to be generated by differential use of three sites

for initiation of transcription located upstream of exons 1, 2, and 3, respectively, and by

the alternative splicing of exon 4; exons 1, 2, and 3 are mutually exclusive (Fig. 9A).

Exon 1 encodes a putative mitochondrial targeting sequence, (M)IVRCLSAARRLHR

(Fig. 9D), which is rich in basic amino acids and can be expected to form an amphipathic

helix (98). The common exons 5 through 16 are used to assemble the functional domain

of the enzyme. The zinc-binding motif HEFGH is encoded by exon 12 (Fig. 7B). The

extreme 3' exon, exon 16, encodes the last 44 amino acid residues, the termination codon,

and the 3' untranslated sequences that include three polyadenylation signals, a short

interspersed repetitive element (SINE or PRE-1), and an AT repeat, all of which were

identified previously by cDNA cloning (86).

There appears to be a "pseudo-exon" that encodes a protein that resembles a

ribosomal protein [11.5 kDa, L44, (99)] in reverse orientation (3' to 5') within the

untranslated region of the 3'-most exon (Fig. 7B). The sequence encoding the homolog

of ribosomal L44 is flanked by the direct repeat TGTTTTAGAGAATTT and has a

poly(A) tract, suggesting that the pseudogene might have arisen as a result of

retroposition.

Structural Comparison of the Genes for EP 24.15 (TOP) and EP 24.16 (MOP)

— I wondered whether the complexity of organization of the gene for EP 24.16 (MOP)

might be reflected in the genes for other members of the thimet oligopeptidase family and,

to this end, I also characterized the gene for EP 24.15 (TOP). The gene for EP 24.15

(TOP) was isolated from the same pig genomic DNA library as that used for isolation of

genes for EP 24.16 (MOP), and it was found to have a much simpler structure in its 5'

region (Fig. 6 and 7B). The gene exists as a single copy, as revealed by Southern blot

analysis (data not shown); it spans approximately 45 kb (Fig. 7A); and it is organized into

13 exons. The overall organization of the two genes is very similar with the exception of

the length of introns and the 5' leader and untranslated exons (Fig. 7B). For example,

exons 2—12 of the gene for EP 24.15 (TOP) correspond precisely to exons 5—15 of the

gene for EP 24.16 (MOP) and there is strong conservation of the respective exon-intron

boundaries (Fig. 7D), suggesting evolution from a common ancestor. The zinc-binding

motif HEFGH is encoded within exon 9. The 3'-terminal exon 13 is composed of a short coding sequence, the termination codon, and the entire 3'-untranslated sequence.

The promoter region of the gene for EP 24.15 (TOP) lacks the TATA box but contains several putative binding sites for ubiquitous factors including one CCAAT box, three Sp1 sites, one NF-1 site, one AP-1 site, and two AP-2 sites (data not shown).

Identification of Three Major Sites of Transcription Initiation — Characterization of the 5' ends of mRNAs for EP 24.16 (MOP) by 5'-RACE (Fig. 4) revealed the presence of multiple sites for initiation of transcription, at least one each in the upstream of exons 1, 2, and 3. To determine the transcription start sites, I performed S1 nuclease mapping using poly(A)$^+$RNA from pig liver and using three probes, which were complementary to exons 1, 2, and 3, respectively. The locations of these probes are indicated in Fig. 8*A*. As I had expected, I found three sites (Fig. 8*A*): one located 172 nt upstream of the first Met (ATG) codon of exon 1; another located at 106 nt upstream of the Met codon of exon 2; and the third located 23 nt downstream of the TATA box close to exon 3.

Putative Sites for Binding of Transcription Factor Near Sites of Initiation of Transcription — Inspection of the sequence of 5'-flanking regions of exons 1, 2, and 3, which I designated promoter regions P1, P2, and P3, respectively (Fig. 7*B*), revealed potential *cis*-acting DNA elements (Fig. 8*B* and 8*C*). Promoter regions 1 and 2 are very GC-rich and lack the TATA and CAAT boxes that are typical of eukaryotic class II promoters; promoter 3 contains a conserved TATA box, which begins 29 nt upstream of the previously identified 5' end of exon 3. The sequences upstream of exons 1 and 2 contain several putative binding sites for transcription factors AP-2 and Sp1. AP-2 mediates enhanced transcription as a result of stimulation by the protein kinase C, cAMP-dependent protein kinase A, and retinoic acid (100-102); Sp1 is a protein that binds to the GC box specifically and is often involved in the regulation of so-called housekeeping genes (103). The upstream region of exon 3 includes consensus binding sites for the

24

transcription factors Myb (product of the myeloblastosis oncogene), AP-1, and GATA-1.

The presence of multiple binding sites for hematopoiesis-specific factors is intriguing:

Myb has been demonstrated to be important in the control of the proliferation and

differentiation of hematopoietic cells (104), while GATA-1 was originally found as an

erythroid-specific factor (105). The Myb-binding site immediately downstream of the

TATA box for exon 3 is of parrticular interest since such a juxtaposed arrangement of a

TATA box and a Myb target sequence was recently demonstrated to serve as a Myb-

suppressible promoter (106).

Differential Subcellular Localization of EP 24.16 (MOP) Directed by

Alternatively Generated Speces of mRNA — The results described above suggest that

the organization of the 5' region of the genes for EP 24.16 (MOP) is unusually complex

and that six mRNA species with different 5' termini are generated as a consequence of the

use of separate promoters (Fig. 8) and the splicing of the 5' leader exons 1 through 3 (in a

mutually exclusive manner) and of exon 4. The cDNA sequences corresponding to the six

species of mRNA are shown schematically in Fig. 9*A*, and they were used for the

expression experiments described below. It should be noted that exon 1 has an in-frame

ATG codon (designated M1), when connected directly to exon 5, and the open reading

frame in exon 1 encodes a putative signal peptide for import into mitochondria; exon 2

also has an in-frame ATG codon (M2) in an appropriate context for the initiation of

translation (Fig. 4) (107) and the open reading frame predicts an enzyme with 64 more

amino acids at its amino terminus then the product generated by the open reading frame

that starts with an ATG codon (M3) in the common exon 5 (Fig. 4 and 9*A*). The fact that

exon 1 could encode an amino-terminal leader sequence for targeting to mitochondria

strongly suggests that, upon selection or elimination of the sequence of exon 1 via

differential utilization of the multiple promoters, the subcellular localization of the products

of the gene for EP 24.16 (MOP) is strictly and efficiently controlled. To confirm this

possibility, I carried out the following experiments.

The six cDNA constructs depicted on the right side of Fig. 9A (labeled types 1 through 3 and 1' through 3') were inserted separately into the mammalian expression vector pcDNA3 and used to transfect COS-1 cells. Then subcellular organelles were isolated from the transfectants and the levels of EP 24.16 (MOP) in these organelles were examined by Western blotting (Fig. 9B and 9C). The type 1 (1-[5—16]) construct directed the synthesis of EP 24.16 (MOP) that was targeted to mitochondria (Fig. 9B, lane 3 and Fig. 9C, MOP[M1b]); the mitochondrial enzyme was slightly smaller than the unprocessed precursor that remained, as a consequence of overexpression of the protein, in the cytosol (Fig. 9B, lane 2, and Fig. 9C, MOP[M1a]). This difference in size indicates that the amino-terminal mitochondrial targeting sequence is cleaved after translocation of the protein into mitochondria. Type 1' (1-4-[5—16]), in which the connection between exons 1 and 5 is interrupted by insertion of exon 4 which includes a stop codon (Fig. 9D), yieled only the cytosolic form of EP 24.16 (MOP) generated from the ATG initiation codon (M3) in exon 5 (Fig. 9B, lanes 5—8). Type 2 (2-[5—16]) allowed the synthesis of an amino-terminally extended cytosolic form (Fig. 9B, lane 10, upper band). Again, as seen with type 2' (2-4-[5—16]), insertion of exon 4 generated a stop codon and only the short cytosolic form was expressed (Fig. 9B, lanes 9—16). With type 1 and type 2, products of translation from the ATG codon in exon 5 (M3) were also detected (Fig. 9B, lane 10, lower band, and Fig. 9C, MOP[M3]), suggesting that these mRNAs generate two isoforms of the protein by alternative usage of codons for the initiation of translation (M1 and M3 for type 1 and M2 and M3 for type 2). The constructs having exon 3 as the 5' leader exon (types 3 and 3') produced only the cytosolic form of the enzyme (Fig. 9B, lanes 17—24), as expected from the fact that exon 3 contains no in-frame ATG codon.

```
          exon no.  clone no.    10      20      30      40      50      60      70      80      90     100     110     120
type 1    MOP(1-5)    (L1):  GGACTGCCGCACGTGGGAAGGCGCGGGCAGAGCGGGCCGCAGCCGCCGCTGTTGCCTCCTTTGCTAGGCCGCCGCAAGGCTTGCCAGCGCCCCGCATACCCCACCCCCAAACCAGAGGAC
type 1'   MOP(1-4-5)  (R5):                                                                                              CCGCATACCCCACCCCCAAACCAGAGGAC
                                                        G
                                                  GCCACCATGG
type 2    MOP(2-5)  (R302):  GACGCTCCTGGAACCCTCCCTTTCTACAGCCCACATGGTCTACCCGAGGGACACCTGGCGCGGGAACTTGGGGCCACCTTTTCTTCCTCTGCTCCCTTGGGAGGGCATCCATTCCCCTT
type 2'   MOP(2-4-5)  (R305):                       ACAGCCCACATGGTCTACCCGACGGACACCTGGCGCGGGAACTTGGGGCCACCTTTTCTTCCTCTGCTCCCTTGGGAGGGCATCCATTCCCCTT
                                                        M2
type 3    MOP(3-5)    (R8):
type 3'   MOP(3-4-5)  (R1):                                                                                                             GGAAATGT

                                      G  M1
                                   GCCACCATGG
type 1    MOP(1-5)    (L1):  AGCGCGCCCGGAGCCGCCGTCGCCTCTCGGCGCTCCCATGATCGTCCGGTGCCTTTCGGCTGCGAGACGCCTCCACAG----------------------------------------
type 1'   MOP(1-4-5)  (R5):  AGCGCGCCCGGAGCCGCCGTCGCCTCTCGGCGCTCCCATGATCGTCCGGTGCCTTTCGGCTGCGAGACGCCTCCACAGAGAGAGACTCCTCTGTCTAAGTTCCATATTTTAGCCCTGGTGCC
type 2    MOP(2-5)  (R302):  TGTTTGGGACTGCCTCAGCTGCAAGCAGGGAGACTGGTCACAAGCCCGGCCCAAAACGAATGCAGAAAGACGGTCGGG----------------------------------------
type 2'   MOP(2-4-5)  (R305):  TGTTTGGGACTGCCTCAGCTGCAAGCAGGGAGACTGGTCGCAAGCCCGGCCCAAAACGAATTCAGAAAGACGGTCGGGAGAGACTCCTCTGTCTAAGTTCCATATTTTAGCCCTGGTGCC
type 3    MOP(3-5)    (R8):  ATTAGAAAAGATTCTGCCAGCTTGAGTGTAACCATCATCCCGTTTGTTTCACACAGGAGCATTCTCATGGAGAGACGG----------------------------------------
type 3'   MOP(3-4-5)  (R1):  ATTAGAAAAGATTCTGCCAGCTTGAGTGTAACCATCATCCCGTTTGTTTCACACAGGAGCATTCTCATGGAGAGACGGAGAGACTCCTCTGTCTAAGTTCCATATTTTAGCCCTGGTGCC
                                                                                     ▲

                                                                                                  G
                                                                                               GCCACCATGG
type 1    MOP(1-5)    (L1):  -----------------------------------------------------------------------GGTTGGTGGTTCCGGGATTTTATTGAGAATGACTTTAGGGAGAGAAG
type 1'   MOP(1-4-5)  (R5):  CCTGTCCATCCTTCCAGTTCCAAGACAGAAGCCTGGATGTCATCCTGAGTCTCTCACCTCCTCTACACATCTGGGTTGGTGGTTCCGGGATTTTATTGAGAATGACTTTAGGGAGAGAAG
type 2    MOP(2-5)  (R302):  -----------------------------------------------------------------------GGTTGGTGGTTCCGGGATTTTATTGAGAATGACTTTAGGGAGAGAAG
type 2'   MOP(2-4-5)  (R305):  CCTGTCCATCCTTCCAGTTCCAAGACAGAAGCCTGGATGTCATCCTGAGTCTCTCACCTCCTCTACACATCTGGGTTGGTGGTTCCGGGATTTTATTGAGAATGACTTTAGGGAGAGAAG
type 3    MOP(3-5)    (R8):  -----------------------------------------------------------------------GGTTGGTGGTTCCGGGATTTTATTGAGAATGACTTTAGGGAGAGAAG
type 3'   MOP(3-4-5)  (R1):  CCTGTCCATCCTTCCAGTTCCAAGACAGAAGCCTGGATGTCATCCTGAGTCTCTCACCTCCTCTACACATCTGGGTTGGTGGTTCCGGGATTTTATTGAGAATGACTTTAGGGAGAGAAG
                                                                            ▲                                          M3

type 1    MOP(1-5)    (L1):  CAATGTCTCCTCTTCAGGCAATGTCTTCCTATACTGTGGATGGCAGAAATGTTTTAAGATGGGATCTTTCACCAGAGCAAATTAAAAGAAGAACTGAGGAGCTGATTGCGCAGACCAAAC
type 1'   MOP(1-4-5)  (R5):  CAATGTCTCCTCTTCAGGCAATGTCTTCCTATACTGTGGATGGCAGAAATGTTTTAAGATGGGATCTTTCACCAGAGCAAATTAAAAGAAGAACTGAGGAGCTGATTGCGCAGACCAAAC
type 2    MOP(2-5)  (R302):  CAATGTCTCCTCTTCAGGCAATGTCTTCCTATACTGTGGATGGCAGAAATGTTTTAAGATGGGATCTTTCACCAGAGCAAATTAAAAGAAGAACTGAGGAGCTGATTGCGCAGACCAAAC
type 2'   MOP(2-4-5)  (R305):  CAATGTCTCCTCTTCAGGCAATGTCTTCCTATACTGTGGATGGCAGAAATGTTTTAAGATGGGATCTTTCACCAGAGCAAATTAAAAGAAGAACTGAGGAGCTGATTGCGCAGACCAAAC
type 3    MOP(3-5)    (R8):  CAATGTCTCCTCTTCAGGCAATGTCTTCCTATACTGTGGATGGCAGAAATGTTTTAAGATGGGATCTTTCACCAGAGCAAATTAAAAGAAGAACTGAGGAGCTGATTGCGCAGACCAAAC
type 3'   MOP(3-4-5)  (R1):  CAATGTCTCCTCTTCAGGCAATGTCTTCCTATACTGTGGATGGCAGAAATGTTTTAAGATGGGATCTTTCACCAGAGCAAATCAAAAGAAGAACTGAGGAGCTGATTGCGCAGACCAAAC
```

Fig. 4. **Nucleotide sequences of the six types of 5' end of cDNAs for pg EP 24.16 (MOP).** The six different sequences revealed by 5'-RACE are aligned. Candidates for codons for initiation of translation are boxed and labeled M1, M2, and M3, respectively. These initiation sites are aligned with the consensus sequence for sites of initiation of translation (GCCA/GCCATGG) (107) which is shown in italics. Three nucleotide replacements, due possibly to allelic polymorphism, were found in exon 2 (underlined). Exon boundaries, determined by comparison of genomic and cDNA sequences (Fig. 7D), are indicated by vertical arrows.
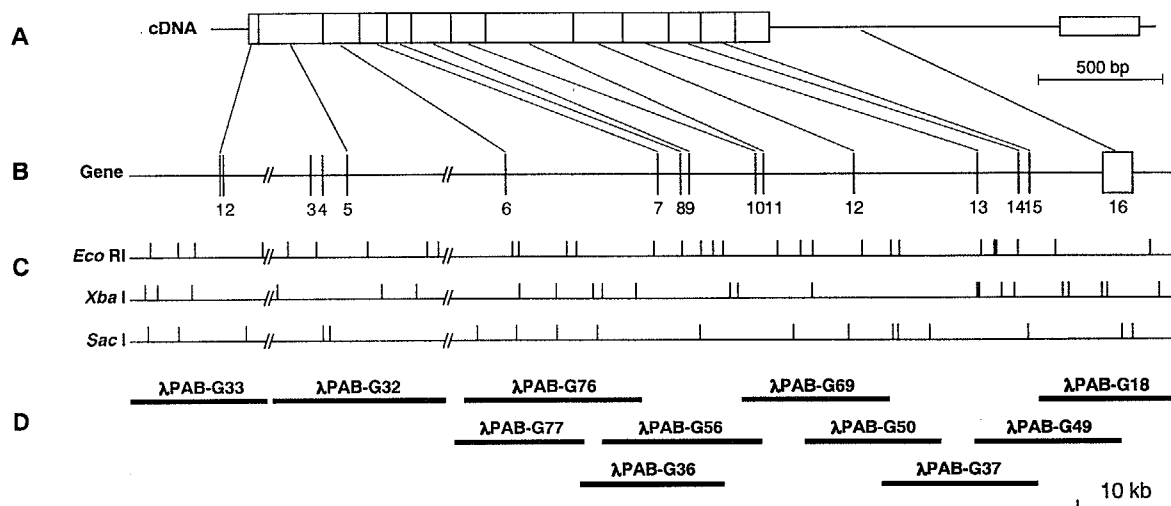
Fig. 5. **Structure of the cDNA for pig EP 24.16 (MOP) in relation to the structure of the gene (A), exon-intron organization of the gene (B), restriction maps for EcoRI, SacI, and XbaI (C), and the relative locations of genomic clones used for the analysis (D).** The structure of cDNA for pig EP 24.16 (MOP) is described elsewhere. (86).

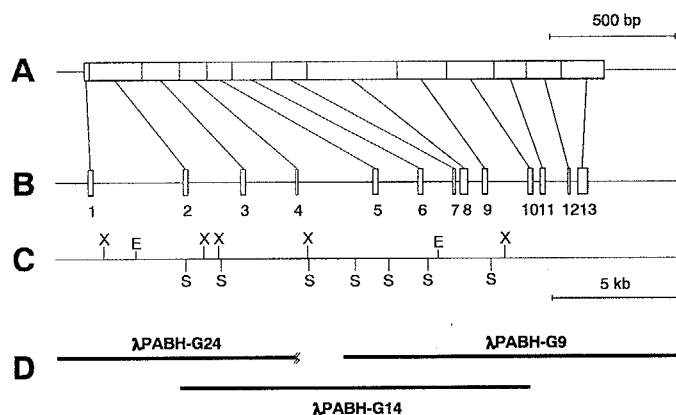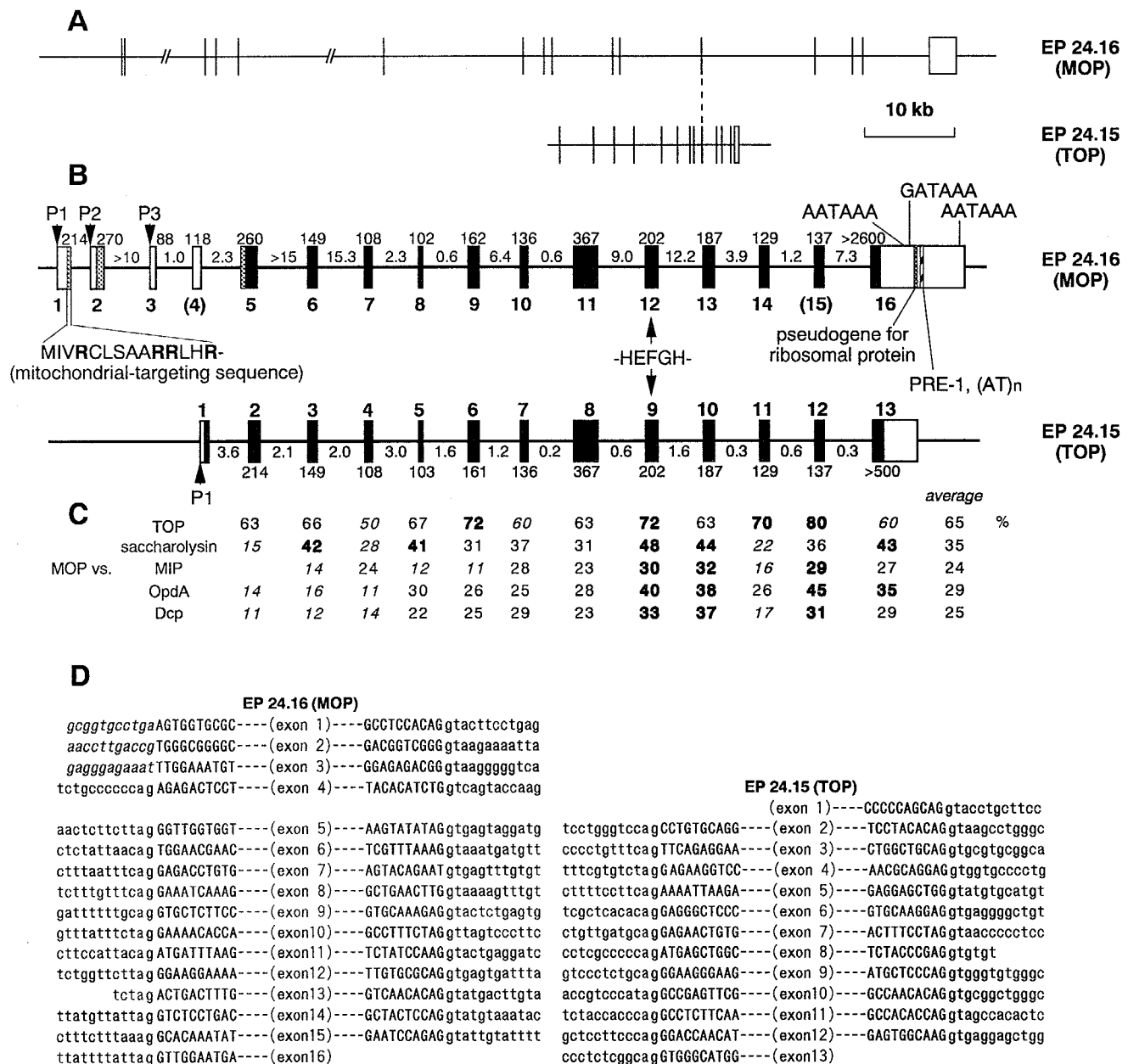Fig. 6. **Structure of the pig gene for EP 24.15 (TOP).** A, structural relationship to the cDNA. The structure of cDNA for pig EP 24.15 (TOP) is described elsewhere. (64). B, exon-intron organization. C, restriction maps. The following abbreviations are used for the restriction enzymes: E, EcoRI; S, SacI; X, XbaI. D, relative locations of genomic clones.

28

A

EP 24.16 (MOP)

10 kb

EP 24.15 (TOP)

B

P1 P2 P3

GATAAA
AATAAA AATAAA

| 214 | 270 | 88 | 118 | 260 | 149 | 108 | 102 | 162 | 136 | 367 | 202 | 187 | 129 | 137 | >2600 |

| | >10 | 1.0 | | 2.3 | >15 | 15.3 | 2.3 | 0.6 | 6.4 | 0.6 | 9.0 | 12.2 | 3.9 | 1.2 | 7.3 |

1 2 3 (4) 5 6 7 8 9 10 11 12 13 14 (15) 16

MIVRCLSAARRLHR-
(mitochondrial-targeting sequence)

-HEFGH-

pseudogene for
ribosomal protein

PRE-1, (AT)n

EP 24.16 (MOP)

1 2 3 4 5 6 7 8 9 10 11 12 13

| | 3.6 | 2.1 | 2.0 | 3.0 | 1.6 | 1.2 | 0.2 | 0.6 | 1.6 | 0.3 | 0.6 | 0.3 |
| | 214 | 149 | 108 | 103 | 161 | 136 | 367 | 202 | 187 | 129 | 137 | >500 |

P1

EP 24.15 (TOP)

C

														average	
	TOP	63	66	50	67	72	60	63	72	63	70	80	60	65	%
	saccharolysin	15	42	28	41	31	37	31	48	44	22	36	43	35	
MOP vs.	MIP		14	24	12	11	28	23	30	32	16	29	27	24	
	OpdA	14	16	11	30	26	25	28	40	38	26	45	35	29	
	Dcp	11	12	14	22	25	29	23	33	37	17	31	29	25	

D

EP 24.16 (MOP)

gcggtgcctgaAGTGGTGCGC----(exon 1)----GCCTCCACAGgtacttcctgag
aaccttgaccgTGGGCGGGGC----(exon 2)----GACGGTCGGGgtaagaaaatta
gagggagaaatTTGGAAATGT----(exon 3)----GGAGAGACGGgtaaggggggtca
tctgcccccagAGAGACTCCT----(exon 4)----TACACATCTGgtcagtaccaag

aactcttcttagGGTTGGTGGT----(exon 5)----AAGTATATAGgtgagtaggatg
ctctattaacagTGGAACGAAC----(exon 6)----TCGTTTAAAGgtaaatgatgtt
ctttaatttcagGAGACCTGTG----(exon 7)----AGTACAGAATgtgagtttgtgt
tctttgtttcagGAAATCAAAG----(exon 8)----GCTGAACTTGgtaaaagtttgt
gattttttgcagGTGCTCTTCC----(exon 9)----GTGCAAAGAGgtactctgagtg
gtttatttctagGAAAACACCA----(exon10)----GCCTTTCTAGgttagtcccttc
cttccattacagATGATTTAAG----(exon11)----TCTATCCAAGgtactgaggatc
tctggttcttagGGAAGGAAAA----(exon12)----TTGTGCGCAGgtgagtgattta
tctagACTGACTTTG----(exon13)----GTCAACACAGgtatgacttgta
ttatgttattagGTCTCCTGAC----(exon14)----GCTACTCCAGgtatgtaaatac
ctttctttaaagGCACAAATAT----(exon15)----GAATCCAGAGgtattgtatttt
ttattttattagGTTGGAATGA----(exon16)

EP 24.15 (TOP)

(exon 1)----CCCCCAGCAGgtacctgcttcc
tcctgggtccagCCTGTGCAGG----(exon 2)----TCCTACACAGgtaagcctgggc
cccctgtttcagTTCAGAGGAA----(exon 3)----CTGGCTGCAGgtgcgtgcggca
tttcgtgtctagGAGAAGGTCC----(exon 4)----AACGCAGGAGgtggtgccctg
cttttccttcaaAAAATTAAGA----(exon 5)----GAGGAGCTGGgtatgtgcatgt
tcgctcacacagGAGGGCTCCC----(exon 6)----GTGCAAGGAGgtgaggggctgt
ctgttgatgcagGAGAACTGTG----(exon 7)----ACTTTCCTAGgtaaccccctcc
cctcgcccccagATGAGCTGGC----(exon 8)----TCTACCCGAGgtgtgt
gtccctctgcagGGAAGGGAAG----(exon 9)----ATGCTCCCAGgtgggtgtgggc
accgtcccatagGCCGAGTTCG----(exon10)----GCCAACACAGgtgcggctgggc
tctaccacccagGCCTCTTCAA----(exon11)----GCCACACCAGgtagccacactc
gctccttcccagGGACCAACAT----(exon12)----GAGTGGCAAGgtgaggagctgg
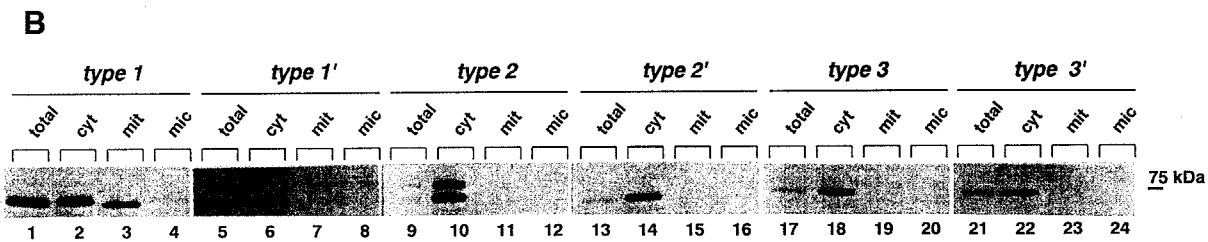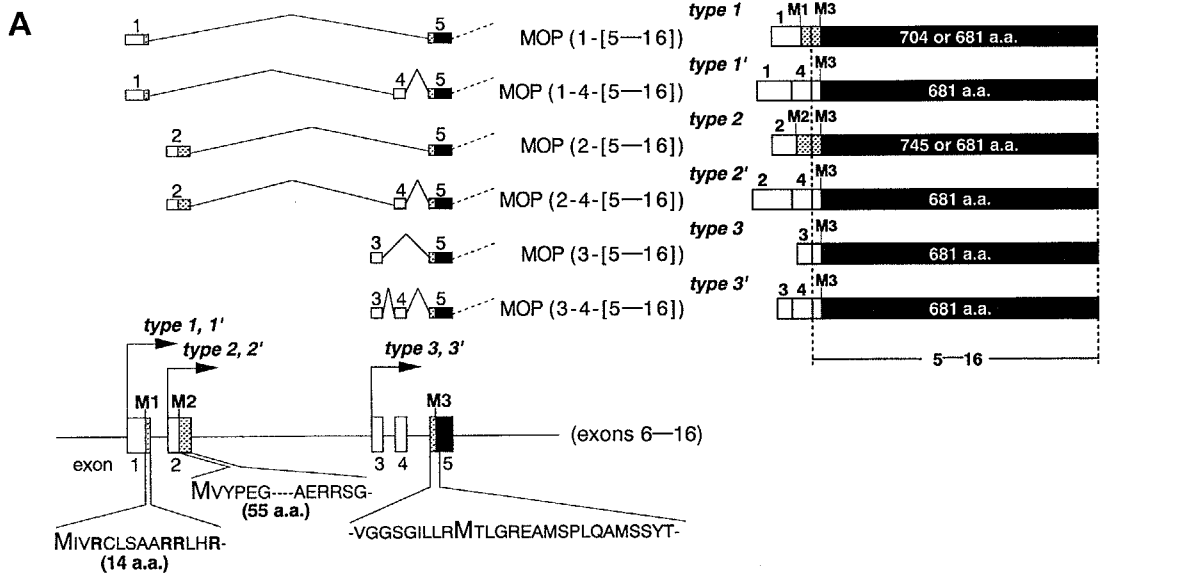ccctctcggcagGTGGGCATGG----(exon13)

Fig. 7. **Comparison of the structures of genes for EP 24.16 (MOP) and EP 24.15 (TOP).** *A*, comparison of the lengths of pig genes for EP 24.16 (MOP) and EP 24.15 (TOP). The exons corresponding to the zinc-binding motifs [exon 12 of EP 24.16 (MOP) and exon 9 of EP 24.15 (TOP)] are aligned (dotted vertical line). *B*, comparison of the exon-intron organization of the gene for EP 24.16 (MOP) with that of EP 24.15 (TOP). Exons are indicated by boxes and are numbered in bold type. Coding regions for the peptidases are shown by filled boxes and amino-terminal extension sequences generated from the alternative sites of initiation of translation of EP 24.16 (MOP) are hatched. Numbers indicate the sizes of exons and introns in bp and kb, respectively. *C*, the extent of sequence homology between the product of the gene for EP 24.16 (MOP) and related peptidases. Coding exons of the gene for EP 24.16 (MOP) (exon 5—16) were converted into deduced amino acid sequences and compared with those of other peptidases, such as pig EP 24.15 (TOP), saccharolysin from *Saccharomyces cervisiae* (75), rat mitochondrial intermediate peptidase (MIP) (72), oligopeptidase A (OpdA) from *Escherichia coli* (125), and dipeptidyl-carboxypeptidase (Dcp) from *Escherichia coli* (126). The extent of homology is indicated, as a percentage, below the corresponding exons. Bold numbers indicate relatively strongly conserved regions, with homology scores more than five points above the average. Italic numbers indicate poorly conserved regions, with homology scores more than five points lower than the average. *D*, nucleotide sequences of exon-intron junctions. The exons are shown in capital letters and the introns in lowercase letters. The 5'-flanking regions of three sites of initiation of transcription are shown in italics.

A

GATC 1 2   GATC 3 4   GATC 5 6

P1   P2   P3

Apa I
Sma I
Sma I
Bss HII
Xho I

1   2

probe 1   probe 2 *

Bgl II   Eco RI

3

probe 3 *

B

```
        10        20        30        40        50        60        70        80        90
                                        AP-2
gggccccggggggtctgtaagccggtccgcctgcccctggtcctcgggcttgagtctggccacgaatcgcagcactgcccacggtggcac
            AP-2                        AP-2
aaacttccccggctcccaggccgccgccgccgcccctcagcccacccggctgtgctctccgctcccagacgccgggtcttgacagaaccg
                    Sp1        RCE                                              Sp1
     AP-2
ccccccgggcggtgctagaggcacggggcggggttcaggcctgggcggggcctgggcggggctgaAGTGGGCGCCGGCGTGGGACTGCCG
Sp1           Sp1            Sp1        Sp1                                        exon 1
CACGTGGGAAGGCGCGGGCAGAGCGGGCCGCAGCCGCCGCTGTTGCCTCCTTTGCTAGGCCGCCGCAAGGCTTGCCAGCGCCCCGCATAC
AP-2
CCCACCCCCAAACCAGAGGACAGCGCGCCCGGAGCCGCCGTCGCTCTCGGCGCTCCCATGATCGTCCGGTGCCTTTCGGCTGCGAGACGC
RCE                                                          M  I  V  R  C  L  S  A  A  R  R
CTCCACAGgtacttcctgagcgcggttttaaccttgaccgTGGGCGGGGCGGGGTCTTCCGCCCGGTTCGAGCCAGGCGTCCCGGCCCAGC
L  H  R                                   Sp1 Sp1           Sp1                   exon 2
CTTTACCGACAGCACCCCTGTGACGCTCCTGGAACCCTCCCTTTCTACAGCCCACATGGTCTACCCCGAGGGACACCTGGCGCGGGAACT
                                                           M  V  Y  P  E  G  H  L  A  R  E  L
TGGGGCCACCTTTTCTTCCTCTGCTCCCTTGGGAGGGCATCCATTCCCCTTTGTTTGGGACTGCCTCAGCTGCAAGCAGGGAGACTGGTC
 G  A  T  F  S  S  A  P  L  G  G  H  P  F  P  F  V  W  D  C  L  S  C  K  Q  G  D  W  S
ACAAGCCCGGCCCAAAACGAATGCAGAAAGACGGTCGGGgtaagaaaattatttcactttcacccccatcagaactgcagacctatcgtt
 Q  A  R  P  K  T  N  A  E  R  R  S  G

gaaagtcgttacagctgttgactcgag
```

C

```
        10        20        30        40        50        60        70        80        90
agatctgagctgcatctgcaacctacaccacagctcacagcaacaccgagtccttaacccactgaacgagggcaggggtctaacctgcat

cctcatggacactatgttggattcttaacctttttgagccacaatgggaactcctgtatattaattttcaaatatttcgatgaaaaaacaa
                     AP-1
tgtcttttgttaagtgtgaatcacttgattggcattcctcactgtccttgcatgaactttttccaaaaatgcattattcttcatttccag

ttttagtttcttcaaagtggagcaggaacactccaggcacacataagcaatctgaattaagaagccttctgaatctccataattcttttc
                                                     Myb
tcagtcttctgtaccagcattttttttcctttggcttttagtaactgtcttttttaaaagtattcccaagacagtttacctaaacaacag
                                                    GATA
atcattctttatgtacatacctttaatcagtaacatttaaattacacgaaattgagtgtgataggcataaacaggttttggaacaatagg
           -29          Myb
catgactcctggtgaatatataactcaactggttgagggagaaatTTGGAAATGTATTAGAAAAGATTCTGCCAGCTTGAGTGTAACCAT
           TATA box                                                                  exon 3
CATCCCGTTTGTTTCACACAGGAGCATTCTCATGGAGAGACGGgtaaggggtcattgttctggtgtctcttgtgttaaaaatgcacccc

tctgcgtctcagcccccttctgtggtgttagctctccttgttttcacagccagctctgtggagaattgctgaggtgtatttcctccttccc
                     AP-2
ctgttctttaccttctccaaagcccgcaggcatagtaatctgacttgcaccctcagcatcctgctaacactgcctttgccaggatcctttt

atggctttcctcttgctgaattc
```

31

Fig. 8. **Identification of site of initiation of transcription of the pig gene for EP 24.16 (MOP) by S1 nuclease mapping (A) and sequences of promoter regions (B and C).** A, five μg of poly(A)$^+$RNA from pig liver (lanes 2, 4, and 6) or yeast tRNA (lanes 1, 3, and 5) were used. Sequencing with the same oligonucleotides was used for calibration of mobilities. Strategies for the preparation of single-stranded antisense DNA probes are shown on the right. B, the sequence of the 837-bp *Apa*I-*Xho*I fragment of pig genomic DNA that contained exons 1 and 2. Sites of initiation of transcription are indicated by arrows. Exon 1 and exon 2 located in a very small region with a GC content of 67%. Nine binding sequences for Sp1, five binding sequences for AP-2, and two Rb control elements (RCE) are indicated. Capital letters represent exons and the deduced amino acid sequences are shown. C, the sequence of the 923-bp *Bgl*II-*Eco*RI fragment of pig genomic DNA that contained exon 3. A site of initiation of transcription is indicated by an arrow. The sequence includes a TATA box at position -23 relative to the site of initiation of transcription of exon 3, three binding sites for Myb, one for AP-1, and one for AP-2. The binding site for Myb site near the TATA-box is very similar to the sequence that is found in the *c-erbB-2* promoter that has been shown to suppress this gene (106).

A

type 1 — MOP (1-[5—16]) — 1 M1 M3 — 704 or 681 a.a.
type 1' — MOP (1-4-[5—16]) — 1 4 M3 — 681 a.a.
type 2 — MOP (2-[5—16]) — 2 M2 M3 — 745 or 681 a.a.
type 2' — MOP (2-4-[5—16]) — 2 4 M3 — 681 a.a.
type 3 — MOP (3-[5—16]) — 3 M3 — 681 a.a.
type 3' — MOP (3-4-[5—16]) — 3 4 M3 — 681 a.a.

5—16

type 1, 1'
type 2, 2'
type 3, 3'

M1 M2   M3
exon 1   2   3 4   5   (exons 6—16)

MVYPEG----AERRSG-
(55 a.a.)

MIVRCLSAARRLHR-
(14 a.a.)

-VGGSGILLRMTLGREAMSPLQAMSSYT-

B

| type 1 | type 1' | type 2 | type 2' | type 3 | type 3' |
| total cyt mit mic | total cyt mit mic | total cyt mit mic | total cyt mit mic | total cyt mit mic | total cyt mit mic |

75 kDa

1 2 3 4   5 6 7 8   9 10 11 12   13 14 15 16   17 18 19 20   21 22 23 24

C

type 1

cyt   total   mit

MOP (M1a) —
MOP (M3) —
— MOP (M1b)

D

type 1 MOP (1-[5–16])
                                                exon 1 | exon 5   M3
CCTCTCGGCGCTCCCATGATCGTCCGGTGCCTTTCGGCTGCGAGACGCCTCCACAGGGTTGGTGGTTCCGGGATTTTATTGAGAATGACTTTAGGGAGAGAAGCAATGTCTCCTCTTCAG
          M I V R C L S A A R R L H R V G G S G I L L R M T L G R E A M S P L Q

GCAATGTCTTCCTATACTGTGGATGGCAGAAATGTTTTAAGATGGGATCTTTCACCAGAGCAAATTAAAAGAAGAACTGAGGAGCTGATTGCGCAGACCAAACAGGTGTATGATGATATT
A M S S Y T V D G R N V L R W D L S P E Q I K R R T E E L I A Q T K Q V Y D D I

type 1' MOP (1-4-[5–16])
                                                exon 1 | exon 4
CGTCGCCTCTCGGCGCTCCCATGATCGTCCGGTGCCTTTCGGCTGCGAGACGCCTCCACAGAGAGACTCCTCTGTCTAAGTTCCATATTTTAGCCCTGGTGCCCCTGTCCATCCTTCCAG
          M I V R C L S A A R R L H R | E T P L S K F H I L A L V P L S I L P V

                                                exon 4 | exon 5   M3
TTCCAAGACAGAAGCCTGGATGTCATCCTGAGTCTCTCACCTCCTCTACACATCTGGGTTGGTGGTTCCGGGATTTTATTGAGAATGACTTTAGGGAGAGAAGCAATGTCTCCTCTTCAG
P R Q K P G C H P E S L T S S T H L G W W F R D F I E N D F R E R S N V S S S G
                                                                M T L G R E A M S P L Q

GCAATGTCTTCCTATACTGTGGATGGCAGAAATGTTTTAAGATGGGATCTTTCACCAGAGCAAATTAAAAGAAGAACTGAGGAGCTGATTGCGCAGACCAAACAGGTGTATGATGATATT
N V F L Y C G W Q K C F K M G S F T R A N *
A M S S Y T V D G R N V L R W D L S P E Q I K R R T E E L I A Q T K Q V Y D D I

type 2 MOP (2-[5–16])
          M2   C
TCTACAGCCCACATGGTCTACCCCGAGGGACACCTGGCGCGGGAACTTGGGGCCACCTTTTCTTCCTCTGCTCCCTTGGGAGGGCATCCATTCCCCTTTGTTTGGGACTGCCTCAGCTGC
          M V Y P E/DG H L A R E L G A T F S S S A P L G G H P F P F V W D C L S C

                                                exon 2 | exon 5   M3
AAGCAGGGAGACTGGTCACAAGCCCGGCCCAAAACGAATGCAGAAAGACGGTCGGGGGTTGGTGGTTCCGGGATTTTATTGAGAATGACTTTAGGGAGAGAAGCAATGTCTCCTCTTCAG
K Q G D W S Q A R P K T N A/SE R R S G | V G G S G I L L R M T L G R E A M S P L Q

type 3 MOP (3-[5–16])
                                                exon 3 | exon 5   M3
TGAGTGTAACCATCATCCCGTTTGTTTCACACAGGAGCATTCTCATGGAGAGACGGGGTTGGTGGTTCCGGGATTTTATTGAGAATGACTTTAGGGAGAGAAGCAATGTCTCCTCTTCAG
                                                M T L G R E A M S P L Q

Fig. 9. **Generation of cytosolic and mitochondrial forms of EP 24.16 (MOP) from a single gene by alternative usage of three promoters (P1— P3) and three codons for initiation of translation (M1—M3).** *A*, schematic representation of the six isoforms of the mRNA for EP 24.16 (MOP) and the organization of the 5' region of the gene for EP 24.16 (MOP), showing how the various isoforms are generated. Exons are indicated by boxes and numbered (box patterns: black, coding regions for the peptidase; hatched, reading frames encoding the amino-terminal extensions; and white, non-coding regions). Three sites for initiation of transcription (Fig. 8) are indicated by arrows. The deduced amino acid sequences corresponding to the reading frames of exons 1, 2, and 5 are shown in single letter code, that contains translational initiation sites indicated by bold and large letters. Basic amino acid residues which are necessary for mitochondrial targeting sequences are indicated by bold letters. Alternative initiation of the transcription of exons 1—3 (P1—P3) and alternative splicing of exon 4 generate six isoforms of the mRNA. *B*, subcellular localization of the products of translation of isoforms of the cDNA for EP 24.16 (MOP) expressed in COS-1 cells. Six cDNA species, identified by 5'-RACE (Fig. 4 and 9*A*), were expressed in COS-1 cells and the products were detected by Western blotting. Subcellular fractions were obtained by differential centrifugation, as follows: cyt, cytosol (100,000 × *g* supernatant); mit, mitochondria (7,000 × *g* pellet); mic, microsomes (100,000 × *g* pellet). Constructions of cDNA used for expression is indicated by boxes on the right of panel (*A*). *C*, cesolution of mitochondrial and cytosolic forms of EP 24.16 (MOP) by SDS-PAGE in the high-resolution buffer system described in the text. M1a, precursor of the mitochondrial form generated by use of the first site (M1) for initiation of translation; M1b, the processed mature form imported into mitochondria; M3, the cytosolic form generated by translation from the M3 site of initiation of translation (for details see Fig. 10). *D*, nucleotide and amino acid sequences of the type 1, 1', 2, and 3 isoforms. Three codons for initiation of translation, M1, M2, and M3, are present in exon 1, exon 2, and exon 5, respectively. The mitochondrial targeting sequence of EP 24.16 (MOP), containing six arginine residues (49), is underlined. Serizawa *et al.* (49, 61) determined the amino-terminal amino acid sequences of the isoforms of EP 24.16 (MOP) purified from mitochondria and the cytosol and showed that the mitochondrial form has a Ser residue at its amino terminus and that the major cytosolic form begins with Thr; the Ser and Thr residues are indicated by white lettering on a black background.

DISCUSSION

In this study, I demonstrated the heterogeneity at the 5' end of the mRNA for EP 24.16 (MOP). Moreover, I showed that the heterogeneity is generated by alternative usage of promoters and splicing of multiple 5' leader and untranslated exons and that it is responsible for the differential subcellular localization of the products of translation.

Targeting of Proteins to Different Subcellular Locations by Alternative Usage of Promoters: EP 24.16 (MOP) Represents the First Example of Such a Mechanism for Intracellular Peptidases/Proteinases — Proteins, after their synthesis, must be delivered to their sites of action. Delivery is usually accomplished with the help of terminal or internal targeting sequences. Sequences for the targeting proteins to the following sites have been identified: mitochondria, endoplasmic reticulum, lysosomes, nuclei, and peroxisomes.

The presence of a putative mitochondrial targeting sequence at the amino terminus of the precursor to EP 24.16 (MOP) was first deduced by Serizawa *et al.* (49) from the potential ability of this sequence to form an amphipathic a-helix with a hydrophobic and a positively charged face of the type expected for a mitochondrial leader sequence (98). This scenario explains the presence of EP 24.16 (MOP) in mitochondria. The enzyme is, however, known also to be present in the cytosol and, prior to the present study, the mechanism responsible for this distribution of EP 24.16 (MOP) has remained unclear. Discovery of 5'-end variants of the mRNA for EP 24.16 (MOP) by the 5'-RACE technique led us to investigate the genetic basis for such diversity. Through an analysis of the structure of the gene, which led to the identification of the three 5'-leader exons that are selected, in a mutually exclusive manner, by use of alternative promoters and splicing, I provided the following resolution of this problem (Fig. 10). If promoter 1 is used, the mitochondrial isoform of EP 24.16 (MOP) is generated by splicing of exon 1, which has a sequence that encodes a signal for transport to mitochondria, to exon 5, which is the beginning of the common translated region that encodes the mature portion of the protein

(type 1 in Fig. 9*A* and 9*D*). The precursor form (704 amino acid residues) with the mitochondrial targeting sequence is processed to the mature mitochondrial form of 667 residues (Fig. 10). The type 1 transcript can also yield the cytosolic form of 681 amino acids when the M3 site of initiation of translation is used instead of the M1 site. If promoter 3 is used and exon 3, which lacks an in-frame ATG codon, is joined to exon 5 (type 5 in Fig. 9*A* and 9*D*), the cytosolic isoform is produced from the ATG initiation codon in exon 5. The use of promoter 2, which directs the synthesis of a cytosolic variant, is discussed below.

Similar senarios have been reported for several other enzymes that are known to occur and function in more than one subcellular compartment (for a recent review, see (108, 109)). Typical examples are the histidine and valine tRNA synthetases of *Saccharomyces cerevisiae* that are involved in protein synthesis in the cytosol and the mitochondria (110, 111). In these cases, two types of transcript (long and short) are produced by alternative usage of promoters, and the long transcript yields the mitochondrial isoform exclusively, while the short transcript yields the cytosolic enzyme. In this way, adjustment of the levels of the proteins to the needs of each compartment is possible. Although the biological significance of this mechanism in the present case is not immediately apparent since the true substrates of the enzyme have not yet been identified, the general regulation of expression of the gene for an oligopeptidase by transcription regulatory factors and the unique regulation, as reported herein, of targeting of the product by use of alternative promoters seems to provide a powerful method by which cells can modulate the concentration of specific peptides in certain intracellular compartments to reflect the metabolic state of the cell.

Complex Organization of the Pig Gene for EP 24.16 (MOP) — As compared to the gene for EP 24.15 (TOP), another member of the thimet oligopeptidase family, the gene for EP 24.16 (MOP) have quite a complicated structure in its 5' and 3' regions. The two genes do, however, exhibit extensive similarity in the regions that encode the mature proteins, which consist of 11 exons, namely, exons 2 through 12 in the case of the gene

36

for EP 24.15 (TOP) and exons 5 through 15 in the case of the gene for EP 24.16 (MOP). The similarity suggests that these two genes and, probably, the genes for other members of this family were generated from an ancestral gene as distinct sequences as a consequence of gene duplication. The presence of a SINE in the 3'-untranslated region of the gene for EP 24.16 (MOP) suggests that insertion of a SINE after the gene duplication might have destabilized the gene for EP 24.16 (MOP) and stimulated extensive diversification of 5' and 3' regions by recruiting the entire gene for ribosomal protein L44 (in the reverse orientation) into the 3'-most exon (exon 16) and the 5'-leader exons into the 5'-flanking region by, perhaps, retroposition and gene conversion.

Long and Short Forms of EP 24.16 (MOP) — The use of promoter 2 of the gene for EP 24.16 (MOP) yields the type 2 transcript, which is predicted to have an amino-terminally elongated product (Fig. 10). Consistent with this prediction, I detected a long form of the protein in the cytosol of COS-1 cells transfected with the type 2 construct after SDS-PAGE and Western blotting (Fig. 9*B*, lane 10). Although the relative abundance of the corresponding mRNA, as estimated from the data after 5'-RACE, in the pig liver is low (<10%), the physiological significance of this form clearly merits further study. The roles of the extended amino-terminal region of 64 amino acid residues (Fig. 9*D*) could include stabilization of the enzyme, modulation of the substrate specificity, and/or mediation of interactions with other cytosolic proteins.

In our analysis, I also noticed the presence of a splice variant that lacked the sequence of exon 15 (A. Kato, N. Sugiura, and S. Hirose, unpublished observation). This variant should encode a protein with a short and slightly different carboxy-terminal tail. The functional significance of this variant and the tissue- and development-specific regulation of the splicing will be the subject of futher research. The presence of at least two forms of EP 24.16 (MOP) has also been demonstrated in purified preparations of the enzyme from rabbit and pig liver (112, 113).

The type 1, type 2, and type 3 species of mRNA all have splice variants, designated type 1', type 2', and type 3', respectively, with an extra exon sequence (5'-

exon 1 2 3 4 5

P1 P2 P3 (exons 6-16)

M1 M2 M3

Nucleus

transcription
and splicing

transcription from exon 1
without exon 4

transcription from exon 2
without exon 4

transcription from exon 3 or
transcription with exon 4

1 M1 M3

2 M2 M3

3 M3

1 4 M3

2 4 M3

3 4 M3

Cytosol

translation

MOP (M3)

| MTL... | (681 a.a.) |

MOP (M1a)

| MIV..SSY... | (704 a.a.) |

MOP (M3)

| MTL... | (681 a.a.) |

MOP (M2)

| MVY... | (745 a.a.) |

MOP (M3)

| MTL... | (681 a.a.) |

translocation and
processing

MOP (M1b)

| | SSY... | (667 a.a.) |

Mitochondria

Fig. 10. **Schematic representation of the mechanism for generation of multiple forms of EP 24.16 (MOP).** Three promoters (P1, P2, and P3) and three codons for the initiation of translation (M1, M2, and M3) are used for the expression of EP 24.16 (MOP). Activation of both P1 and M1 is necessary for expression of EP 24.16 (MOP) in mitochondria. Insertion of exon 4 prevents the translocation of EP 24.16 (MOP) to mitochondria.

untranslated exon 4, Fig. 10), but none of them results in a long open reading frame from the first initiation codon. It is unknown therefore, whether these variants have any biological significance. However, the possibility exists that the variant species of mRNA might contribute to the regulation of rates of translation. Alternatively, they might produce short peptides with as yet unidentified functions. In the case of the type 1' transcript, the insertion of the sequence of exon 4 provides a mechanism by which the synthesis of the mitochondrial isoform is suppressed even under conditions under which promoter 1 is active but there is no mitochondrial requirement for the oligopeptidase.

CONCLUSIONS AND PERSPECTIVE

Conclusions— This is the first study that shows the structures of genes for EP 24.15 (TOP) and EP 24.16 (MOP). Analysis of the gene for EP 24.16 (MOP) revealed the very complex organization of the gene, and the presence of a variety of transcripts generated by differential use of multiple sites of transcription initiation and by alternative splicing of exons 2, 3, 4, and 15. In contrast to these complexities, a simple and definitive answer was obtained to the question of how the product of a single gene for EP 24.16 (MOP) is delivered to two different cellular compartments, namely, the cytosol and the mitochondria.

Perspective: Functions of EP 24.16 (MOP) and EP 24.15 (TOP)— It is useful to consider the function of genes by comparing corresponding genes in mammals and other organisms, especially yeast, a model organism for eukaryotes. *Saccharomyces cerevisiae*, the whole genome sequence of which was determined recently by the genome project (114), has one gene closely related to EP 24.16 (MOP) and EP 24.15 (TOP), named saccharolysin (75). Another homologous gene was found in *Aspergillus fumigatus*, and named mepB (115). Saccharolysin was found in cytosolic and mitochondrial fractions, and mepB was found in cytosol (75, 115). Homologies among EP 24.15 (TOP), EP 24.16 (MOP), saccharolysin, and mepB are very high (65% within mammals, 35—38% between mammals and fungi), and these genes form a single branch within the phylogenetic tree of the M3 family of metallopeptidases (Fig. 11). Wide distribution of these genes among species (from fungi to mammals) and highly conserved primary structures suggest their essential roles in eukaryotic cells. However, mutant cell lines of *Saccharomyces cerevisiae* and *Aspergillus fumigatus* that lack the saccharolysin and mepB genes, respectively, did not show any grossly altered phenotypes (115). The yeast mutant exhibited a decrease in the intracellular degradation of peptides, suggesting saccharolysin's function in degradation of peptides in cytosol (75). Similarly, in

mammals, EP 24.15 (TOP) and EP 24.16 (MOP) are also considered to play roles in metabolizing cytosolic peptides. It is also possible that these enzymes take part in the system for antigen presentation.

It seems important that mammalian cells and *Saccharomyces cerevisiae* contain these peptidases in mitochondria (EP 24.16 [MOP] and saccharolysin) (Table 4), suggesting their essential roles in mitochondria. Mitochondria have their own proteolytic systems, consisting of many proteases and peptidases, in the mitochondrial interior. Those proteases are 1) processing proteases such as MPP (116, 117), IMP (118), and MIP (72, 73), or 2) proteases such as m-AAA, i-AAA, Lon, and ClpP (119, 120), which degrade mitochondrial proteins (Fig. 12). Peptides arising during processing or degradation of mitochondrial proteins may be substrates for EP 24.16 (MOP) or saccharolysin.

It is also interesting to compare the regulatory mechanisms by which EP 24.16 (MOP) and saccharolysin become localized both in cytosol and mitochondria (75). In the case of EP 24.16 (MOP), whose gene has a complex organization at the 5' end, localization of the enzyme is regulated by alternative use of multiple promoters, alternative splicing, and alternative translation initiation. In contrast, saccharolysin is encoded by a single exon, and the localization seems to be regulated by usage of multiple transcription initiation sites or multiple translation initiation codons (Fig. 13).

Finally, in contrast to the essential roles of the intracellular endo-oligopeptidases described above, the mammalian enzymes EP 24.15 (TOP) and EP 24.16 (MOP) are expected to have some other unique functions. The phylogenetic tree shows that the duplication of genes for EP 24.15 (TOP) and EP 24.16 (MOP) occurred after the division of mammals and fungi, and mammals are unique in having two closely related genes. The complex structure of the gene for EP 24.16 (MOP) at the 5' end also suggests the presence of second or third functions of this enzyme, which are regulated through second or third promoters. EP 24.15 (TOP) and EP 24.16 (MOP) are widely distributed among various tissues (64), but the mRNA level of EP 24.15 (TOP) in testis is much higher than in other tissues (69, 121), suggesting some function of EP 24.15

(TOP) in the male reproductive system. The roles of secreted EP 24.15 (TOP) and EP

24.16 (MOP) in metabolizing bioactive peptides are also considerable (122, 123).

However, the mechanism for the secretion of these peptidases is still not established.

```
                                    MIP (Human)
                                    MIP (Rat)
                                    MIP (Saccharomyces cerevisiae)
                                    MIP (Schizosaccharomyces pombe)
                                    MIP (Schizophyllum commune)
                                    EP 24.16 (MOP) (Rat)
                                    EP 24.16 (MOP) (Pig)
                                    EP 24.16 (MOP) (Rabbit)
                                    EP 24.15 (TOP) (Rat)
                                    EP 24.15 (TOP) (Pig)
                                    EP 24.15 (TOP) (Human)
                                    Saccharolysin (Saccharomyces cerevisiae)
                                    MepB (Aspergillus fumigatus)
                                    DCP (Escherichia coli)
                                    DCP (Salmonella typhimurium)
                                    OPA (Synechocystis sp.)
                                    OPA (Haemophilus influenzae)
                                    OPA (Escherichia coli)
                                    OPA (Salmonella typhimurium)
                                    pepF (Lactococcus lactis)
                                    pepF (Mycoplasma pneumoniae)
                                    pepF (Mycoplasma genitalium)
```

Fig. 11. **Phylogenetic tree based on the protein sequences of metallopeptidases which belong to the M3 family.** Sequences were aligned using the program GENETYX-MAC (Software Development, Tokyo, Japan) to construct a phylogenetic tree by the unweighted pair group method using arithmetic averages (UPGMA). The tree forms four branches: 1) MIP, 2) EP 24.15/EP 24.16/saccharolysin/mepB, 3) DCP/OPA, and 4) pepF. The branches are indicated by brown, red, purple, and green colores, respectively. The following abbreviations are used: MIP, mitochondrial intermediate peptidase; DCP, peptidyl-dipeptidase; and OPA, oligopeptidase A.

42
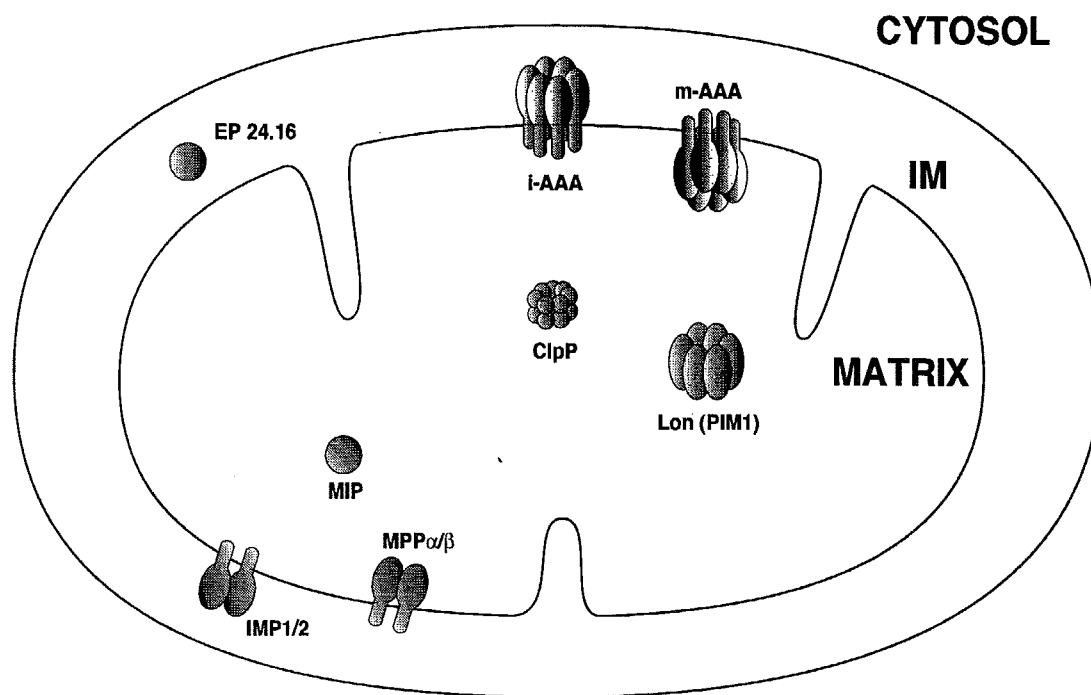
Fig. 12. **Mitochondrial proteolytic systems.** Mitochondrial proteases that have been isolated from eukaryotic cells are illustrated. MPP, IMP, and MIP are necessary for processing the N-termini of mitochondrial proteins which are encoded by nuclear genes. Mitochondrial proteins are degraded by ATP-dependent proteases such as m-AAA, i-AAA, ClpP, and Lon (PIM1). Chaperone-like activities of these proteases are consider to be important not only for proteolysis but also for the assembly of mitochondrial protein complexes. The following abbreviations are used: MPP, mitochondrial processing peptidase; IMP, mitochondrial inner membrane peptidase; MIP, mitochondrial intermediate peptidase; and AAA, ATPase associated with a variety of cellular activities.

43
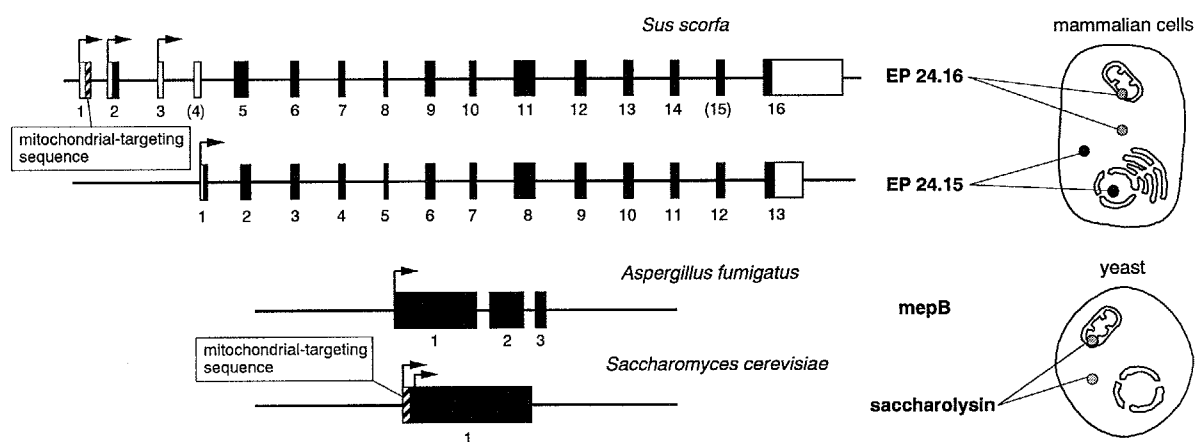
Fig. 13. **Comparison of gene structures and intracellular localizations of EP 24.16 (MOP), EP 24.15 (TOP), mepB, and saccharolysin.** Exons are indicated by *boxes* and are *numbered*. Coding regions for the peptidases are shown by *filled boxes* and mitochondrial-targeting sequences are *hatched*. Arrows indicate initiation sites for transcription. Intracellular localizations of EP 24.16 (MOP), EP 24.15 (TOP), and saccharolysin are illustrated on the *right*.

44

Table 4

Intracellular localization of EP 24.15 (TOP), EP 24.16 (MOP), and saccharolysin
Localizations in each compartment are: ⊚, well examined, and targeting signals are present; ○, examined, but targeting signals are not found; △, examined, but they were questioned by other studies; ✕, not detected; or *ND*, not examined.

Name	Cytosol	Nucleus	Mitochondria	Microsome or membrane	Secretion
Endopeptidase 24.15 (TOP)	⊚	⊚	✕	△	○
Endopeptidase 24.16 (MOP)	⊚	ND	⊚	△	○
Saccharolysin	⊚	ND	⊚	ND	ND

45

ACKNOWLEDGMENTS

REFERENCES

1.   Schimke, R. T. (1973) Control of enzyme levels in mammalian tissues. *Adv Enzymol Relat Areas Mol Biol* **37**, 135-87

2.   Goldberg, A. L., and St, J. A. (1976) Intracellular protein degradation in mammalian and bacterial cells: Part 2. *Annu Rev Biochem* **45**, 747-803

3.   Rogers, S., Wells, R., and Rechsteiner, M. (1986) Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* **234**, 364-8

4.   Gottesman, S., Wickner, S., and Maurizi, M. R. (1997) Protein quality control: triage by chaperones and proteases. *Genes Dev* **11**, 815-23

5.   Kopito, R. R. (1997) ER quality control: the cytoplasmic connection. *Cell* **88**, 427-30

6.   Seufert, W., and Jentsch, S. (1990) Ubiquitin-conjugating enzymes UBC4 and UBC5 mediate selective degradation of short-lived and abnormal proteins. *Embo J* **9**, 543-50

7.   Hershko, A., and Ciechanover, A. (1992) The ubiquitin system for protein degradation. *Annu Rev Biochem* **61**, 761-807

8.   Heemels, M. T., and Ploegh, H. (1995) Generation, translocation, and presentation of MHC class I-restricted peptides. *Annu Rev Biochem* **64**, 463-91

9.   Koopmann, J. O., Hammerling, G. J., and Momburg, F. (1997) Generation, intracellular transport and loading of peptides associated with MHC class I molecules. *Curr Opin Immunol* **9**, 80-8

10. York, I. A., and Rock, K. L. (1996) Antigen processing and presentation by the class I major histocompatibility complex. *Annu Rev Immunol* **14**, 369-96

11.   Watts, C. (1997) Capture and processing of exogenous antigens for presentation on MHC molecules. *Annu Rev Immunol* **15**, 821-50

12.   Cresswell, P. (1994) Assembly, transport, and function of MHC class II molecules. *Annu Rev Immunol* **12**, 259-93

13. Rawlings, N. D., and Barrett, A. J. (1995) Evolutionary families of metallopeptidases. *Methods Enzymol* **248**, 183-228

14. Rawlings, N. D., and Barrett, A. J. (1994) Families of serine peptidases. *Methods Enzymol* **244**, 19-61

15. Rawlings, N. D., and Barrett, A. J. (1995) Families of aspartic peptidases, and those of unknown catalytic mechanism. *Methods Enzymol* **248**, 105-20

16. Rawlings, N. D., and Barrett, A. J. (1994) Families of cysteine peptidases. *Methods Enzymol* **244**, 461-86

17. Rawlings, N. D., and Barrett, A. J. (1993) Evolutionary families of peptidases. *Biochem J* **290**, 205-18

18. Bainton, D. F. (1981) The discovery of lysosomes. *J Cell Biol* **91**, 66s-76s

19. Kornfeld, S., and Mellman, I. (1989) The biogenesis of lysosomes. *Annu Rev Cell Biol* **5**, 483-525

20. Bohley, P., and Seglen, P. O. (1992) Proteases and proteolysis in the lysosome. *Experientia* **48**, 151-7

21. Tan, F., Morris, P. W., Skidgel, R. A., and Erdos, E. G. (1993) Sequencing and cloning of human prolylcarboxypeptidase (angiotensinase C). Similarity to both serine carboxypeptidase and prolylendopeptidase families [published erratum appears in J Biol Chem 1993 Dec 5;268(34):26032]. *J Biol Chem* **268**, 16631-8

22. Chen, J. M., Dando, P. M., Rawlings, N. D., Brown, M. A., Young, N. E., Stevens, R. A., Hewitt, E., Watts, C., and Barrett, A. J. (1997) Cloning, isolation, and characterization of mammalian legumain, an asparaginyl endopeptidase. *J Biol Chem* **272**, 8090-8

23. Li, Y. P., Alexander, M., Wucherpfennig, A. L., Yelick, P., Chen, W., and Stashenko, P. (1995) Cloning and complete coding sequence of a novel human cathepsin expressed in giant cells of osteoclastomas. *J Bone Miner Res* **10**, 1197-202

24. Inaoka, T., Bilbe, G., Ishibashi, O., Tezuka, K., Kumegawa, M., and Kokubo, T. (1995) Molecular cloning of human cDNA for cathepsin K: novel cysteine proteinase predominantly expressed in bone. *Biochem Biophys Res Commun* **206**, 89-96

25. Shi, G. P., Munger, J. S., Meara, J. P., Rich, D. H., and Chapman, H. A. (1992) Molecular cloning and expression of human alveolar macrophage cathepsin S, an elastinolytic cysteine protease. *J Biol Chem* **267**, 7258-62

26. Linnevers, C., Smeekens, S. P., and Bromme, D. (1997) Human cathepsin W, a putative cysteine protease predominantly expressed in CD8+ T-lymphocytes. *Febs Lett* **405**, 253-9

27. Seglen, P. O., Gordon, P. B., and Holen, I. (1990) Non-selective autophagy. *Semin Cell Biol* **1**, 441-8

28. Matsuura, A., Tsukada, M., Wada, Y., and Ohsumi, Y. (1997) Apg1p, a novel protein kinase required for the autophagic process in Saccharomyces cerevisiae. *Gene* **192**, 245-50

29. Funakoshi, T., Matsuura, A., Noda, T., and Ohsumi, Y. (1997) Analyses of APG13 gene involved in autophagy in yeast, Saccharomyces cerevisiae. *Gene* **192**, 207-13

30. Holen, I., Gordon, P. B., and Seglen, P. O. (1993) Inhibition of hepatocytic autophagy by okadaic acid and other protein phosphatase inhibitors. *Eur J Biochem* **215**, 113-22

31. Blommaart, E. F., Krause, U., Schellens, J. P., Vreeling-Sindelarova, H., and Meijer, A. J. (1997) The phosphatidylinositol 3-kinase inhibitors wortmannin and LY294002 inhibit autophagy in isolated rat hepatocytes. *Eur J Biochem* **243**, 240-6

32. Coux, O., Tanaka, K., and Goldberg, A. L. (1996) Structure and functions of the 20S and 26S proteasomes. *Annu Rev Biochem* **65**, 801-47

33. Palombella, V. J., Rando, O. J., Goldberg, A. L., and Maniatis, T. (1994) The ubiquitin-proteasome pathway is required for processing the NF- kappa B1 precursor protein and the activation of NF-kappa B. *Cell* **78**, 773-85

34. Rodriguez, M. S., Wright, J., Thompson, J., Thomas, D., Baleux, F., Virelizier, J. L., Hay, R. T., and Arenzana-Seisdedos, F. (1996) Identification of lysine residues required for signal-induced ubiquitination and degradation of I kappa B-alpha in vivo. *Oncogene* **12**, 2425-35

35. Ciechanover, A., DiGiuseppe, J. A., Bercovich, B., Orian, A., Richter, J. D., Schwartz, A. L., and Brodeur, G. M. (1991) Degradation of nuclear oncoproteins by the ubiquitin system in vitro. *Proc Natl Acad Sci U S A* **88**, 139-43

36. Udvardy, A. (1996) The role of controlled proteolysis in cell-cycle regulation. *Eur J Biochem* **240**, 307-13

37. King, R. W., Deshaies, R. J., Peters, J. M., and Kirschner, M. W. (1996) How proteolysis drives the cell cycle. *Science* **274**, 1652-9

38. Chen, P., Johnson, P., Sommer, T., Jentsch, S., and Hochstrasser, M. (1993) Multiple ubiquitin-conjugating enzymes participate in the in vivo degradation of the yeast MAT alpha 2 repressor. *Cell* **74**, 357-69

39. Miyazawa, K., Toyama, K., Gotoh, A., Hendrie, P. C., Mantel, C., and Broxmeyer, H. E. (1994) Ligand-dependent polyubiquitination of c-kit gene product: a possible mechanism of receptor down modulation in MO7e cells. *Blood* **83**, 137-45

40. Mori, S., Kanaki, H., Tanaka, K., Morisaki, N., and Saito, Y. (1995) Ligand-activated platelet-derived growth factor beta-receptor is degraded through proteasome-dependent proteolytic pathway. *Biochem Biophys Res Commun* **217**, 224-9

41. Roth, A. F., and Davis, N. G. (1996) Ubiquitination of the yeast a-factor receptor. *J Cell Biol* **134**, 661-74

42. Yu, H., Kaung, G., Kobayashi, S., and Kopito, R. R. (1997) Cytosolic degradation of T-cell receptor alpha chains by the proteasome. *J Biol Chem* **272**, 20800-4

43. Tierney, D. J., Haas, A. L., and Koop, D. R. (1992) Degradation of cytochrome P450 2E1: selective loss after labilization of the enzyme. *Arch Biochem Biophys* **293**, 9-16

44. Correia, M. A., Davoll, S. H., Wrighton, S. A., and Thomas, P. E. (1992) Degradation of rat liver cytochromes P450 3A after their inactivation by 3,5-dicarbethoxy-2,6-dimethyl-4-ethyl-1,4-dihydropyridine: characterization of the proteolytic system. *Arch Biochem Biophys* **297**, 228-38

45. Roberts, B. J. (1997) Evidence of proteasome-mediated cytochrome P-450 degradation. *J Biol Chem* **272**, 9771-8

46. Barrett, A. J., Brown, M. A., Dando, P. M., Knight, C. G., McKie, N., Rawlings, N. D., and Serizawa, A. (1995) Thimet oligopeptidase and oligopeptidase M or neurolysin. *Methods Enzymol* **248**, 529-56

47. Dauch, P., Vincent, J. P., and Checler, F. (1995) Molecular cloning and expression of rat brain endopeptidase 3.4.24.16. *J Biol Chem* **270**, 27266-71

48. Checler, F., Barelli, H., Dauch, P., Dive, V., Vincent, B., and Vincent, J. P. (1995) Neurolysin: purification and assays. *Methods Enzymol* **248**, 593-614

49. Serizawa, A., Dando, P. M., and Barrett, A. J. (1995) Characterization of a mitochondrial metallopeptidase reveals neurolysin as a homologue of thimet oligopeptidase. *J Biol Chem* **270**, 2092-8

50. Rennex, D., Hemmings, B. A., Hofsteenge, J., and Stone, S. R. (1991) cDNA cloning of porcine brain prolyl endopeptidase and identification of the active-site seryl residue. *Biochem.* **30**, 2195-203

51. Constam, D. B., Tobler, A. R., Rensing-Ehl, A., Kemler, I., Hersh, L. B., and Fontana, A. (1995) Puromycin-sensitive aminopeptidase. Sequence analysis, expression, and functional characterization. *J Biol Chem* **270**, 26931-9

52. Bromme, D., Rossi, A. B., Smeekens, S. P., Anderson, D. C., and Payan, D. G. (1996) Human bleomycin hydrolase: molecular cloning, sequencing, functional expression, and enzymatic characterization. *Biochem.* **35**, 6706-14

53. Kobayashi, K., Lin, L. W., Yeadon, J. E., Klickstein, L. B., and Smith, J. A. (1989) Cloning and sequence analysis of a rat liver cDNA encoding acyl-peptide hydrolase. *J Biol Chem* **264**, 8892-9

54.   Dando, P. M., Young, N. E., and Barrett, A. J. (1997) Aminopeptidase PS: a widely distributed cytosolic peptidase, in *Proteolysis in Cell Function* (Hopsu-Havu VK, Järvinen M, Kirschke H ed.) pp. 88-95, IOS Press, Amsterdam

55.   Adachi, H., Tsujimoto, M., Fukasawa, M., Sato, Y., Arai, H., Inoue, K., and Nishimura, T. (1997) cDNA cloning and expression of chicken aminopeptidase H, possessing endopeptidase as well as aminopeptidase activity. *Eur J Biochem* **245**, 283-8

56.   Dorr, R. T. (1992) Bleomycin pharmacology: mechanism of action and resistance, and clinical pharmacokinetics. *Semin Oncol* **19**, 3-8

57.   Tsunasawa, S. (1995) [Amino-terminal processing of nascent proteins: their rule and implication on biological function]. *Tanpakushitsu Kakusan Koso* **40**, 389-98

58.   Vanhoof, G., Goossens, F., De Meester, I., Hendriks, D., and Scharpe, S. (1995) Proline motifs in peptides and their biological processing. *Faseb J* **9**, 736-44

59.   Tisljar, U. (1993) Thimet oligopeptidase--a review of a thiol dependent metallo-endopeptidase also known as Pz-peptidase endopeptidase 24.15 and endo- oligopeptidase. *Biol Chem Hoppe Seyler* **374**, 91-100

60.   Healy, D. P., and Orlowski, M. (1992) Immunocytochemical localization of endopeptidase 24.15 in rat brain. *Brain Res* **571**, 121-8

61.   Serizawa, A., Dando, P. M., and Barrett, A. J. (1997) Oligopeptidase M (neurolysin). Targeting to Mitochondria and cytosol in rat tissues, in *Proteolysis in Cell Function* (Hopsu-Havu VK, Järvinen M, Kirschke H ed.) pp. 248-53, IOS Press, Amsterdam

62.   Camargo, A., Gomes, M. D., Reichl, A. P., Ferro, E. S., Jacchieri, S., Hirata, I. Y., and Juliano, L. (1997) Structural features that make oligopeptides susceptible substrates for hydrolysis by recombinant thimet oligopeptidase. *Biochem J* **324**, 517-22

63.   Cabrera, A., Gomes, M. D., Juliano, M., and Ferro, E. S. (1997) Susceptibility of neutral MHC-class I ligands to cytosolic oligopeptidases, in *Proceeding of the Biology of Proteolysis* (Craik, C. S., Gottesman, S., Hochstrasser, M., and Lazebnik, Y ed.) p24, Cold Spring Harbor Laboratory meeting, New York

64. Kato, A., Sugiura, N., Hagiwara, H., and Hirose, S. (1994) Cloning, amino acid sequence and tissue distribution of porcine thimet oligopeptidase. A comparison with soluble angiotensin-binding protein. *Eur J Biochem* **221**, 159-65

65. Kato, A., and Hirose, S., (1996) Targeting of neurolysin (oligopeptidase M, MOP) to different subcellular compartment by alternative promoter usage. *ICOP newsletter* Nov:4-5

66. Kato, A., Sugiura, N., Saruta, Y., Hosoiri, T., Yasue, H., and Hirose, S. (1997) Targeting of endopeptidase 24.16 to different subcellular compartments by alternative promoter usage. *J Biol Chem* **272**, 15313-22

67. Kato, A., Sugiura, N., Saruta, Y., Hosoiri, T., Yasue, H., and Hirose, S. (1997) Analysis of the genes of thimet oligopeptidase (EP 24.15) and oligopeptidase M (EP 24.16) and intracellular targeting of EP 24.16 by alternative promoter usage, in *Proteolysis in Cell Function* (Hopsu-Havu VK, Järvinen M, Kirschke H ed.) pp. 208-14, IOS Press, Amsterdam

68. Bond, J. S., and Beynon, R. J. (1985) Mammalian metalloendopeptidases. *Int J Biochem* **17**, 565-74

69. Pierotti, A., Dong, K. W., Glucksman, M. J., Orlowski, M., and Roberts, J. L. (1990) Molecular cloning and primary structure of rat testes metalloendopeptidase EC 3.4.24.15 [published erratum appears in Biochemistry 1994 Jan 18;33(2):622]. *Biochem.* **29**, 10323-9

70. Conlin, C. A., and Miller, C. G. (1995) Dipeptidyl carboxypeptidase and oligopeptidase A from Escherichia coli and Salmonella typhimurium. *Methods Enzymol* **248**, 567-79

71. Monnet, V. (1995) Oligopeptidases from Lactococcus lactis. *Methods Enzymol* **248**, 579-92

72. Isaya, G., Kalousek, F., and Rosenberg, L. E. (1992) Sequence analysis of rat mitochondrial intermediate peptidase: similarity to zinc metallopeptidases and to a putative yeast homologue. *Proc Natl Acad Sci U S A* **89**, 8317-21

73.   Isaya, G., Miklos, D., and Rollins, R. A. (1994) MIP1, a new yeast gene homologous to the rat mitochondrial intermediate peptidase gene, is required for oxidative metabolism in Saccharomyces cerevisiae. *Mol Cell Biol* **14**, 5603-16

74.   Isaya, G., and Kalousek, F. (1995) Mitochondrial intermediate peptidase. *Methods Enzymol* **248**, 556-67

75.   Buchler, M., Tisljar, U., and Wolf, D. H. (1994) Proteinase yscD (oligopeptidase yscD). Structure, function and relationship of the yeast enzyme with mammalian thimet oligopeptidase (metalloendopeptidase, EP 24.15). *Eur J Biochem* **219**, 627-39

76.   Wünsch, E., and Heidrich, H. G. (1963) Zur quantitativen bestimmung der kollagenase. *Hoppe Seylers Z Physiol Chem* **333**, 149-51

77.   Harper, E., and Gross, J. (1970) Separation of collagenase and peptidase activities of tadpole tissues in culture. *Biochim Biophys Acta* **198**, 286-92

78.   Camargo, A. C., Ramalho-Pinto, F. J., and Greene, L. J. (1972) Brain peptidases: conversion and inactivation of kinin hormones. *J Neurochem* **19**, 37-49

79.   Orlowski, M., Michaud, C., and Chu, T. G. (1983) A soluble metalloendopeptidase from rat brain. Purification of the enzyme and determination of specificity with synthetic and natural peptides. *Eur J Biochem* **135**, 81-8

80.   Barrett, A. J., and Brown, M. A. (1990) Chicken liver Pz-peptidase, a thiol-dependent metallo-endopeptidase [see comments]. *Biochem J* **271**, 701-6

81.   McKie, N., Dando, P. M., Rawlings, N. D., and Barrett, A. J. (1993) Thimet oligopeptidase: similarity to 'soluble angiotensin II-binding protein' and some corrections to the published amino acid sequence of the rat testis enzyme. *Biochem J* **295**, 57-60

82.   Pierotti, A., Dong, K. W., Glucksman, M. J., Orlowski, M., and Roberts, J. L. (1994) Molecular cloning and primary structure of rat testes metalloendopeptidase EC 3.4.24.15. *Biochem.* **33**, 622

83.   Thompson, A., Huber, G., and Malherbe, P. (1995) Cloning and functional expression of a metalloendopeptidase from human brain with the ability to cleave a beta-APP substrate peptide. *Biochem Biophys Res Commun* **213**, 66-73

84. Heidrich, H. G., Prokopova, D., and Hannig, K. (1969) The use of synthetic substrates for the determination of mammalian collagenases. Is collagenolytic activity present in mitochondria? *Hoppe Seylers Z Physiol Chem* **350**, 1430-6

85. Tisljar, U., and Barrett, A. J. (1990) A distinct thimet peptidase from rat liver mitochondria. *Febs Lett* **264**, 84-6

86. Sugiura, N., Hagiwara, H., and Hirose, S. (1992) Molecular cloning of porcine soluble angiotensin-binding protein. *J Biol Chem* **267**, 18067-72

87. Hagiwara, H., Sugiura, N., Wakita, K., and Hirose, S. (1989) Purification and characterization of angiotensin-binding protein from porcine liver cytosolic fraction. *Eur J Biochem* **185**, 405-10

88. Kiron, M. A., and Soffer, R. L. (1989) Purification and properties of a soluble angiotensin II-binding protein from rabbit liver. *J Biol Chem* **264**, 4138-42

89. Nakagawa, K., Kawabata, S., Nakashima, Y., Iwanaga, S., and Sueishi, K. (1997) Tissue distribution and subcellular localization of rabbit liver metalloendopeptidase. *J Histochem Cytochem* **45**, 41-7

90. Kawabata, S., and Davie, E. W. (1992) A microsomal endopeptidase from liver with substrate specificity for processing proproteins such as the vitamin K-dependent proteins of plasma. *J Biol Chem* **267**, 10331-6

91. Kawabata, S., Nakagawa, K., Muta, T., Iwanaga, S., and Davie, E. W. (1993) Rabbit liver microsomal endopeptidase with substrate specificity for processing proproteins is structurally related to rat testes metalloendopeptidase 24.15. *J Biol Chem* **268**, 12498-503

92. Kojima, N., Kawabata, S., Makinose, Y., Nishino, N., and Iwanaga, S. (1995) Substrate specificity of rabbit liver metalloendopeptidase and its new fluorogenic peptide substrates. *J Biochem* **118**, 855-61

93. Checler, F., Emson, P. C., Vincent, J. P., and Kitabgi, P. (1984) Inactivation of neurotensin by rat brain synaptic membranes. Cleavage at the Pro10-Tyr11 bond by endopeptidase 24.11 (enkephalinase) and a peptidase different from proline-endopeptidase. *J Neurochem* **43**, 1295-301

94.	Checler, F., Vincent, J. P., and Kitabgi, P. (1983) Degradation of neurotensin by rat brain synaptic membranes: involvement of a thermolysin-like metalloendopeptidase (enkephalinase), angiotensin- converting enzyme, and other unidentified peptidases. *J Neurochem* **41**, 375-84

95.	Checler, F., Vincent, J. P., and Kitabgi, P. (1986) Purification and characterization of a novel neurotensin-degrading peptidase from rat brain synaptic membranes. *J Biol Chem* **261**, 11274-81

96.	Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-7

97.	Breathnach, R., and Chambon, P. (1981) Organization and expression of eucaryotic split genes coding for proteins. *Annu Rev Biochem* **50**, 349-83

98.	von Heijne, G., Steppuhn, J., and Herrmann, R. G. (1989) Domain structure of mitochondrial and chloroplast targeting peptides. *Eur J Biochem* **180**, 535-45

99.	Davies, M. S., Henney, A., Ward, W. H., and Craig, R. K. (1986) Characterisation of an mRNA encoding a human ribosomal protein homologous to the yeast L44 ribosomal protein. *Gene* **45**, 183-91

100.	Williams, T., Admon, A., Luscher, B., and Tjian, R. (1988) Cloning and expression of AP-2, a cell-type-specific transcription factor that activates inducible enhancer elements. *Genes Dev* **2**, 1557-69

101.	Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R. J., Rahmsdorf, H. J., Jonat, C., Herrlich, P., and Karin, M. (1987) Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor. *Cell* **49**, 729-39

102.	Hyman, S. E., Comb, M., Pearlberg, J., and Goodman, H. M. (1989) An AP-2 element acts synergistically with the cyclic AMP- and phorbol ester-inducible enhancer of the human proenkephalin gene. *Mol Cell Biol* **9**, 321-4

103.	Kadonaga, J. T., Jones, K. A., and Tijan, R. (1986) Promoter specific activation of RNA polymerase II transcription by Sp1. *Trends Biochem. Sci.* **11**, 20-3

104.	Luscher, B., and Eisenman, R. N. (1990) New light on Myc and Myb. Part II. Myb. *Genes Dev* **4**, 2235-41

105. Ko, L. J., and Engel, J. D. (1993) DNA-binding specificities of the GATA transcription factor family. *Mol Cell Biol* **13**, 4011-22

106. Mizuguchi, G., Kanei-Ishii, C., Takahashi, T., Yasukawa, T., Nagase, T., Horikoshi, M., Yamamoto, T., and Ishii, S. (1995) c-Myb repression of c-erbB-2 transcription by direct binding to the c- erbB-2 promoter. *J Biol Chem* **270**, 9384-9

107. Kozak, M. (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* **15**, 8125-48

108. Danpure, C. J. (1995) How can the products of a single gene be localized to more than one intracellular compartment? *Trends Cell Biol.* **5**, 230-8

109. Ayoubi, T. A., and Van De Ven, W. J. (1996) Regulation of gene expression by alternative promoters. *Faseb J* **10**, 453-60

110. Natsoulis, G., Hilger, F., and Fink, G. R. (1986) The HTS1 gene encodes both the cytoplasmic and mitochondrial histidine tRNA synthetases of S. cerevisiae. *Cell* **46**, 235-43

111. Chatton, B., Walter, P., Ebel, J. P., Lacroute, F., and Fasiolo, F. (1988) The yeast VAS1 gene encodes both mitochondrial and cytoplasmic valyl- tRNA synthetases. *J Biol Chem* **263**, 52-7

112. Rosenberg, E., Russo, C., Kiron, M. A., and Soffer, R. L. (1988) Biochemical and immunological similarity of soluble angiotensin II- binding proteins in different organs. *Biochem Biophys Res Commun* **156**, 1230-6

113. Sugiura, N., Hagiwara, H., Ito, T., and Hirose, S. (1991) Biochemistry and molecular biology of angiotensin-binding protein, in *Peptide Regulation of Cardiovascular Function* (Imura H, Matsuo H, Masaki T ed.) pp. 157-66, Academic Press, New York

114. Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H., and Oliver, S. G. (1996) Life with 6000 genes [see comments]. *Science* **274**, 546, 563-7

115. Ibrahim-Granet, O., and D'Enfert, C. (1997) The Aspergillus fumigatus mepB gene encodes an 82 kDa intracellular metalloproteinase structurally related to mammalian thimet oligopeptidases. *Microbiology* **143**, 2247-53

116. Pfanner, N., Douglas, M. G., Endo, T., Hoogenraad, N. J., Jensen, R. E., Meijer, M., Neupert, W., Schatz, G., Schmitz, U. K., and Shore, G. C. (1996) Uniform nomenclature for the protein transport machinery of the mitochondrial membranes. *Trends Biochem Sci* **21**, 51-2

117. Witte, C., Jensen, R. E., Yaffe, M. P., and Schatz, G. (1988) MAS1, a gene essential for yeast mitochondrial assembly, encodes a subunit of the mitochondrial processing protease. *Embo J* **7**, 1439-47

118. Nunnari, J., Fox, T. D., and Walter, P. (1993) A mitochondrial protease with two catalytic subunits of nonoverlapping specificities. *Science* **262**, 1997-2004

119. Langer, T., and Neupert, W. (1996) Regulated protein degradation in mitochondria. *Experientia* **52**, 1069-76

120. Suzuki, C. K., Rep, M., Maarten van Dijl, J., Suda, K., Grivell, L. A., and Schatz, G. (1997) ATP-dependent proteases that also chaperone protein biogenesis. *Trends Biochem Sci* **22**, 118-23

121. Hayashi, M. A., Gomes, M. D., Reboucas, N. A., Fernandes, B. L., Ferro, E. S., and de Camargo, A. C. (1996) Species specificity of thimet oligopeptidase (EC 3.4.24.15). *Biol Chem Hoppe Seyler* **377**, 283-91

122. Ferro, E. S., Tambourgi, D. V., Gobersztejn, F., Gomes, M. D., Sucupira, M., Armelin, M. C., Kipnis, T. L., and Camargo, A. C. (1993) Secretion of a neuropeptide-metabolizing enzyme similar to endopeptidase 22.19 by glioma C6 cells. *Biochem Biophys Res Commun* **191**, 275-81

123. Vincent, B., Beaudet, A., Dauch, P., Vincent, J. P., and Checler, F. (1996) Distinct properties of neuronal and astrocytic endopeptidase 3.4.24.16: a study on differentiation, subcellular distribution, and secretion processes. *J Neurosci* **16**, 5049-59

124. Roelse, J., Gromme, M., Momburg, F., Hammerling, G., and Neefjes, J. (1994) Trimming of TAP-translocated peptides in the endoplasmic reticulum and in the cytosol during recycling. *J Exp Med* **180**, 1591-7

125. Conlin, C. A., Trun, N. J., Silhavy, T. J., and Miller, C. G. (1992) Escherichia coli prlC encodes an endopeptidase and is homologous to the Salmonella typhimurium opdA gene. *J Bacteriol* **174**, 5881-7

126. Henrich, B., Becker, S., Schroeder, U., and Plapp, R. (1993) dcp gene of Escherichia coli: cloning, sequencing, transcript mapping, and characterization of the gene product. *J Bacteriol* **175**, 7290-300