Volume 335, number 3, 367–375 FEBS 13357 December 1993

Cloning of the PABA peptide hydrolase alpha subunit (PPHα) from human small intestine and its expression in COS-1 cells

Eric Dumermuth[a], Joyce A. Eldering[a], Jürgen Grünberg[a], Weiping Jiang[b], Erwin E. Sterchi[a,*]

[a]*Institute of Biochemistry and Molecular Biology, University of Berne, Bühlstrasse 28, CH-3012 Berne, Switzerland*
[b]*Department of Biochemistry and Molecular Biology, Pennsylvania State University College of Medicine, Hershey, PA 17033, USA*

Received 28 October 1993

PABA peptide hydrolase (PPH) from human enterocytes is comprised of two subunits, alpha and beta. PPHα is over 70% identical to meprin, a protease isolated from mouse and rat kidney. The enzyme shows a modular organization in that it contains an astacin protease domain, an adhesive domain, an EGF-like domain, and a putative C-terminal membrane spanning domain. Expression of a chimeric meprin-PPHα cDNA in COS-1 cells led to the synthesis of immature, transport-incompetent homodimers. In addition, complex glycosylated forms were detected in the culture medium, suggesting that the enzyme is secreted after proteolytic removal of the membrane anchor.

PABA peptide hydrolase; Meprin; Astacin; Zinc-metalloendopeptidase; Human; Enterocyte; COS-1

1. INTRODUCTION

PABA peptide hydrolase (PPH) has been identified as an enzyme activity in the human small intestinal mucosa capable of hydrolyzing *N*-benzoyl-L-tyrosine-*p*-amino benzoic acid (PABA peptide), a substrate used in an exocrine pancreatic function test [1]. It was immunopurified and biochemically characterized [2]. In addition to PABA peptide hydrolyzing activity, PPH cleaved a variety of biologically active peptides and was found to be dependent on bivalent metal ions. Biosynthetic labelling studies in organ cultured intestinal explants showed that the formation of dimers in the endoplasmic reticulum was an early post-translational event [3]. cDNA cloning led to the discovery of sequence homologies between PPH and meprin, a metalloendopeptidase from mouse kidney [4], and analysis of the DNA and protein sequence databanks found these two enzymes to be homologous to a protease from *Astacus fluviatilis*, a freshwater crayfish [5]. These studies led to the definition of a new family of metalloendopeptidases, the astacin family [6]. The astacin signature has been found in a variety of proteins or cDNA-deduced sequences, including the human bone morphogenic protein BMP1 [7], UVS.2 from *Xenopus laevis* embryos [8], tolloid from *Drosophila melanogaster* embryos [9], blastula protease 10 (BP10) and SPAN from sea urchin embryos [10,11] and LCE/HCE (low and high choriolytic enzyme) from *Oryzias latipes* (fish embryo) [12]. These proteins are involved in early morphogenic processes, but so far proteolytic activity has been demonstrated only for BMP1 and LCE/HCE. A characteristic structural feature of some of these proteins is the uniform presence of EGF-like domains. The cDNA and deduced amino acid sequences of the alpha and beta subunits of meprin from mouse [13,14] and rat [15–17] have recently been published. The meprin alpha subunits from mouse and rat are 92% similar/87% identical, and thus are isomeric forms of the enzyme (EC 3.4.24.18). Meprin has been shown to contain an EGF-like domain located close to a putative membrane anchoring domain near the C-terminus of the protein. It was found to contain a putative propeptide sequence which was preceded by a hydrophobic sequence compatible with a signal peptide sequence. In mouse kidney, two forms, meprin-A and -B, have been found. There are inbred mouse strains deficient in meprin-A but not meprin-B [18]. Mouse meprin has been reported to exist as homo- and/or hetero-oligomers (α_4, $\alpha_2\beta_2$, β_4) [19], whereas rat meprin appears to form ab heterodimers which form tetrameric structures [16].

Here, we report the cloned cDNA sequence of the α subunit of PPH from human small intestine and its comparison to meprin from mouse and rat kidney. The three enzymes are closely related and may be species-specific isomers. In order to study the postsynthetic events leading to the formation of dimeric and oligomeric structures, PPHα was expressed in COS-1 cells in a chimeric form (MPα) comprising the near full-length cDNA from PPHα and the 5′ cDNA portion from mouse meprinα (including the start codon for methionine). These expression data show, that PPHα forms ho-

*Corresponding author. Fax: (41) (31) 631 37 37.

Abbreviations: PPH, PABA peptide hydrolase; SDS-PAGE, sodium dodecylsulfate polyacrylamide gel electrophoresis; EGF, epidermal growth factor; ER, endoplasmic reticulum; NP40, Nonidet 40; DOC, deoxycholate; PBS, phosphate-buffered saline.

modimers and oligomers, and is retained in the ER of COS-1 cells in an immature membrane form. In addition, a fully glycosylated form is secreted in soluble form.

2. EXPERIMENTAL

2.1. *Materials*

The radiolabelled nucleotides ([α-^{35}S]dATP, [γ-^{32}P]dATP and [α-^{32}P]dCTP), [^{35}S]Methionine, and the GeneScreen membrane were from Dupont/NEN. Restriction endonucleases, *Thermus aquaticus* DNA- polymerase, Klenow fragment and T4 polynucleotide kinase were purchased from Boehringer Mannheim. The M-MLV Reverse transcriptase was from United States Biochemical Corp. The Zeta-Probe membranes were obtained from Bio-Rad. Calf intestinal alkaline phosphatase was from Promega. The pXT1 and pSG5 vectors were from Stratagene Cloning System. The pGEX-3X vector and the protein A Sepharose beads were from Pharmacia Biosystem. Cell culture media, penicillin and streptomycin were obtained from GIBCO BRL. Fetal calf serum (FCS) was from Biological Industries. Phenylmethanesulphonyl fluoride (PMSF), pepstatin, aprotinin, leupeptin, benzamidine, Freund's adjuvant and molecular weight standards were from Sigma Chemical Co. Immobilon membranes were from Millipore, ProBlot membranes from Applied Biosystems. All other chemicals were analytical grade from Merck.

2.2. *Enzyme purification*

PPH was immunopurified from Triton X-100 solubilized microvillus membrane vesicles of human small intestinal mucosa by a specific monoclonal antibody (HBB 3/716/36) as previously described [2] and analyzed by SDS-polyacrylamide gel electrophoresis, followed by Coomassie staining.

2.3. *Peptide sequencing*

Immunoprecipitated PPH was separated by SDS-PAGE, followed by electrophoretic transfer to nylon membranes. The subunits PPHα and PPHβ were visualized with Coomassie blue and cut out separately. N-terminal peptide sequences were obtained from blotted proteins in a gas-phase sequencer (Applied Biosystems, model 477).

2.4. *Isolation and characterization of PPH cDNA clones*

N-terminal peptide sequences of PPH subunits were used to design oligonucleotide probes using best 'guess' codons [20] (see Fig. 1B). The PPHα-specific probe had an expected homology of ≥ 82% to the target cDNA and was used to screen 150,000 to 300,000 plaques of a λgt11 human intestine cDNA library. Plaques were transferred to Zeta-Probe nylon membranes in duplicate using a standard procedure [21]. The oligonucleotides were 5′-^{32}P-labelled and hybridized overnight to the membranes 49°C. The hybridization solution contained 6 × SSC (1 × SSC is 150 mM NaCl, 15 mM trisodium citrate), 1% SDS, 1/6 vol. nonfat milk, and approx. 1 pmol/ml ^{32}P-labelled probe. Membranes were washed twice in 6 × SSC, 1% SDS, and then in 2 × SSC, 1% SDS for 15 min at the hybridization temperature. A single clone hybridized to the PPHα oligonucleotide probe. This isolate was termed PPHα1, its 210 bp insert was subcloned into pBluescript (Stratagene), and sequenced on both strands by the Sanger dideoxy termination method using Sequenase (Version 2.0, United Biochemical Corp.) and the M13 Universal and reverse primers. This clone was used to screen about 600,000 recombinants of a λZAPII human intestine cDNA library. Several positive clones were obtained and their inserts subcloned into pBluescript by in vivo rescue (Stratagene). These clones were all overlapping and extended the PPHα1 cDNA to both directions, including the poly A sequence at the 3′ end. The largest clone (pPPHα22.4) contained a 2.8 kb insert and was sequenced on both strands by subcloning restriction fragments, generation of 3′ deletions (Exonuclease III from Stratagene, Mung Bean Nuclease from New England Biolabs), and the remaining gaps were filled by extension with gene-specific primers based on previously determined sequences.

2.5. *Southern and Northern blot analysis*

For Southern blot analysis, 7 μg of genomic DNA from human adonocarcinoma cells, Caco-2, was digested with *Pvu*II, *Pst*I, *Hin*dIII, *Eco*RI, or *Bam*HI. The restriction fragments were separated by electrophoresis (1% agarose) and blotted by capillary transfer (20 × SSC, 16 h) onto Zeta Probe nylon membrane. The membrane was hybridized to a random-^{32}P-labelled PPHα1 cDNA probe in standard hybridization solution as above, at 60°C. The membranes were washed as above.

For Northern blot analysis, total RNA was prepared from human small and large intestinal mucosa (both normal and tumor-derived) and MDCK cells [22]. Polyadenylated mRNA (1 μg) was isolated from total RNA (≈ 75 μg) using the Dynabeads mRNA purification kit (Dynal), mRNA from human kidney was purchased from Clontech. mRNA was separated by electrophoresis in the absence of ethidium bromide in a 1% agarose-formaldehyde gel, followed by capillary transfer (0.05 M NaOH, 3 h) to Zetaprobe nylon membranes. Membranes were hybridized (50% formamide (deionized), 0.25 M NaH_2PO_4 (pH 7.2 with H_3PO_4), 0.25 M NaCl, 7% (w/v) SDS and 1 mM EDTA) to a random-^{32}P-labelled PPHα cDNA probe (42°C, 30 min) and washed at room temperature (2 × SSC, 0.1% SDS and 1 × SSC, 0.1% SDS, 15 min each). mRNA bands were visualized by autoradiography. To ensure the presence of intact mRNA on the blot, the membranes were stripped of the PPHα probe and hybridized with an actin oligonucleotide probe.

2.6. *Construction of meprin-PPHα expression vector (pSGMPα)*

All DNA manipulations were performed as described in Sambrook et al. [23]. To construct the pSGMPα expression vector, the pSKMPα was created by inserting a ~200 bp *Eco*RI–*Pst*I meprinα DNA fragment containing the AUG start codon (instead of the ~50 bp *Eco*RI–*Pst*I fragment from the PPHα subunit without a start codon) into pPPHα22.4. First pPPHα22.4 was digested with *Eco*RI and *Hin*dIII, dephosphorylated and gel purified. The ~200 bp *Eco*RI–*Pst*I meprinα cDNA fragment and the ~1.2 kbp *Pst*I–*Hin*dIII PPHα DNA fragment were ligated into this vector resulting pSKMPα. From this vector the ~2.9 kbp *Bam*HI–*Bgl*II MPα fragment was inserted into the *Bam*HI–*Bgl*II sites of the expression vector pSG5 creating pSGMPα (Fig. 6)

2.7. *Antibodies*

Antibodies against a fusion protein between Glutathione-*S*-transferase and a portion of the human PPHα (amino acids 1–421, counted from the coding sequence in pPPHα22.4) were raised in rabbits by standard procedures. The plasmid pPPG-N1 was constructed by inserting a *Bam*HI–*Kpn*I (blunt ended) fragment of PPHα in the *Bam*HI–*Sma*I site of pGEX-3X and recutting the vector with *Hin*dIII–*Xho*I. The restriction sites were filled using DNA polymerase I, Klenow fragment and religated. The fusion protein was expressed in *E. coli* [24] strain XL1-Blue (Stratagene) and was purified by two successive SDS-PAGE. After staining of the proteins with Coomassie R-250 prepared in water, the fusion protein was cut out of the gel and injected into the rabbits with Freund's adjuvant [25]. PPH-specific monoclonal antibodies used were HBB 3/716/36 [26].

2.8. *Transient expression of the pSGMPα in COS-1 cells*

COS-1 cells were grown in EMEM supplemented with 10% (v/v) FCS, 100 units/ml of penicillin and 100 μg/ml of streptomycin. 4 × 10^5 cells were seeded onto 6 cm cell culture dishes 24 h before transfection with DEAE-dextran. The method was as described by Berger and Kimmel [21]. The transfection cocktail (0.5 μg pSGMPα dissolved in 950 μl PBS and 50 μl DEAE-dextran stock solution (10 mg/ml PBS) was added to previously washed COS-1 cells. The cells were incubated for 30 min at 37°C. 4 ml tissue culture medium supplemented with 100 μM chloroquine was added and the cells were incubated again at 37°C for 2.5 h. After a shock with 10% (v/v) DMSO in EMEM for 2.5 min the cells were grown for 40–48 h in 5 ml EMEM containing 10% FCS before labelling.

2.9. *Metabolic labelling of COS-1 cells*

The cells were washed with warm PBS and were incubated in

methionine-free media supplemented with 10% (v/v) dialyzed FCS for 1 h. The cells were continuously labelled with 50 μCi [^{35}S]methionine (1 Ci = 37 GBq) for 5 h or over night [27].

2.10. *Immunoprecipitation*

The cells were washed three times with ice-cold PBS and scraped off in 1 ml ice-cold PBS. After collection by low speed centrifugation they were lysed in 300 μl NP40/DOC, 1% each in PBS. After removing the cell debris by high speed centrifugation 7 μl of 20% (w/v) SDS was added and the proteins were denatured by boiling for 5 min. 1 ml of HBT buffer (25 mM Tris-HCl, 50 mM NaCl, 1% (w/v) Triton X-100, pH 8.1) and 50 μl Sepharose A beads were added to preclear the probe. The immunoprecipitation with 20 μl anti-PPHα antisera was as described previously [27].

2.11. *Endoglycosidase H, endoglycosidase F and tunicamycin treatment*

Treatment of immunoprecipitated PPHα proteins with endo H and endo F, and metabolic labelling of cells in the presence of tunicamycin was as described earlier [3,27,28].

2.12. *Computer analysis of cDNA and proteins sequences*

For all computer assisted analysis of RNA, DNA and proteins, the GCG (Genetics Computer Group, Madison, Wisconsin) software package was used.

3. RESULTS

3.1. *Enzyme purification and NH_2-terminal amino acid sequence determination*

Analysis of the immunopurified PPH from small intestinal mucosa by SDS-PAGE under reducing conditions yielded two distinct Coomassie-stained polypeptide bands with apparent M_r = 90 kDa and 97 kDa, termed PPHα and PPHβ (Fig. 1A) respectively. Amino acid sequence analysis of these two polypeptides revealed two different NH_2-terminal sequences which, however, showed some sequence homology (Fig. 1B).

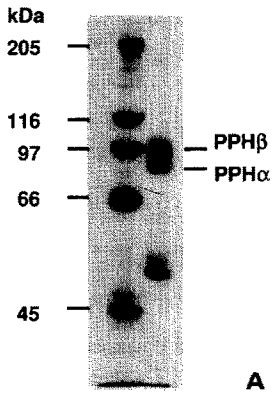

Fig. 1A. Analysis of immunoprecipitated PPH by SDS-PAGE. PPH was immunoprecipitated from Triton X-100 solubilized microvillus membrane vesicles prepared from human small intestinal mucosa using a specific monoclonal antibody. Electrophoresis was in a 7.5% DS-polyacrylamide gel under reducing conditions. Proteins were stained by Coomassie blue R-250. Markers are identified by M_r in the left lane.

3.2. *Cloning and sequence analysis of the alpha subunit of PPH cDNA*

Upon screening of a λgt11 library from human small intestine, using the PPHα-specific oligonucleotide probe, a cDNA fragment of 210 bp was isolated. This cDNA fragment, which contained an open reading frame of 67 amino acids, including the 24 amino acids sequenced from the NH_2-terminus of detergent-solubilized PPHα, was utilized to probe a λZapII human intestinal cDNA library. Among several clones isolated the largest had an insert of 2.8 kb, contained the entire 210 bp sequence of the probe, a stop codon and a polyA tail. The isolated 2.8 kb clone encoded 714 amino acids and included the NH_2 terminal amino acid sequence of mature PPHα (Fig. 2). Although the open reading frame extended from the NH_2 terminus, it did not contain a starting methionine. The NH_2 terminal amino acid sequence of the detergent-solubilized mature α subunit of PPH starts at position Asn^{34} and is indicated by a solid triangle in Fig. 2. A hydropathy analysis according to Kyte and Doolittle [29] revealed a 26 amino acid long hydrophobic peptide starting from His^{683} to Leu^{708}. This is followed by a short hydrophylic peptide (6 amino acids) containing three positively charged residues which represent the COOH-terminus of the translated sequence. These two peptide segments probably represent a membrane-spanning domain and a short cytosolic domain respectively. As previously reported [6], PPHα is a member of the newly defined family of metalloendopeptidases, the astacin family. The astacin domain in PPHα spans from Asn^{34} to His^{231}, and contains the extended consensus sequence, HEXXHXXGFXHE-(Q,H)XRXDRDX(Y,H)(V,I)X(I,V), found in all proteins of the astacin family. In analogy to astacin, the residues involved in the zinc-binding in PPHα are His^{123} (His^{92} in astacin), His^{127} (His^{96} in astacin), His^{133} (His^{102} in astacin), and Tyr^{182} (Tyr^{149} in astacin). Also conserved within the astacin signature of PPHα are four cysteine residues which have been shown to be involved in disulfide-bonding in astacin [5]. Thus, Cys^{75} is probably linked to Cys^{227} and Cys^{96} is linked to Cys^{115}.

The intervening sequence between the putative mem-

```
1              5                  10                  15
ASN SER ILE ILE GLY GLU LYS TYR ARG SER PRO HIS PRO ILE PRO TYR VAL LEU   PPHβ
 |   :          |                   |  |  |  |  :  |
ASN GLY LEU ARG ASP PRO ASN THR ARG TRP THR PHE PRO ILE PRO TYR ILE LEU   PPHα
                CTG GGG TTG TGG GCC ACC TGG AAG GGG TAG GGT ATG TA 5'     oPPHα
```

B

Fig. 1B. N-terminal amino acid sequences of the alpha and beta subunits of PPH. After SDS-PAGE, proteins were electroblotted onto ProBlot membranes, stained with Coomassie blue R-250 and excised for direct sequencing in an Applied Biosystems Sequencer model 477 fitted with a blotting cell. The PPHα-specific oligonucleotide probe used for initial screening of cDNA libraries is as shown.

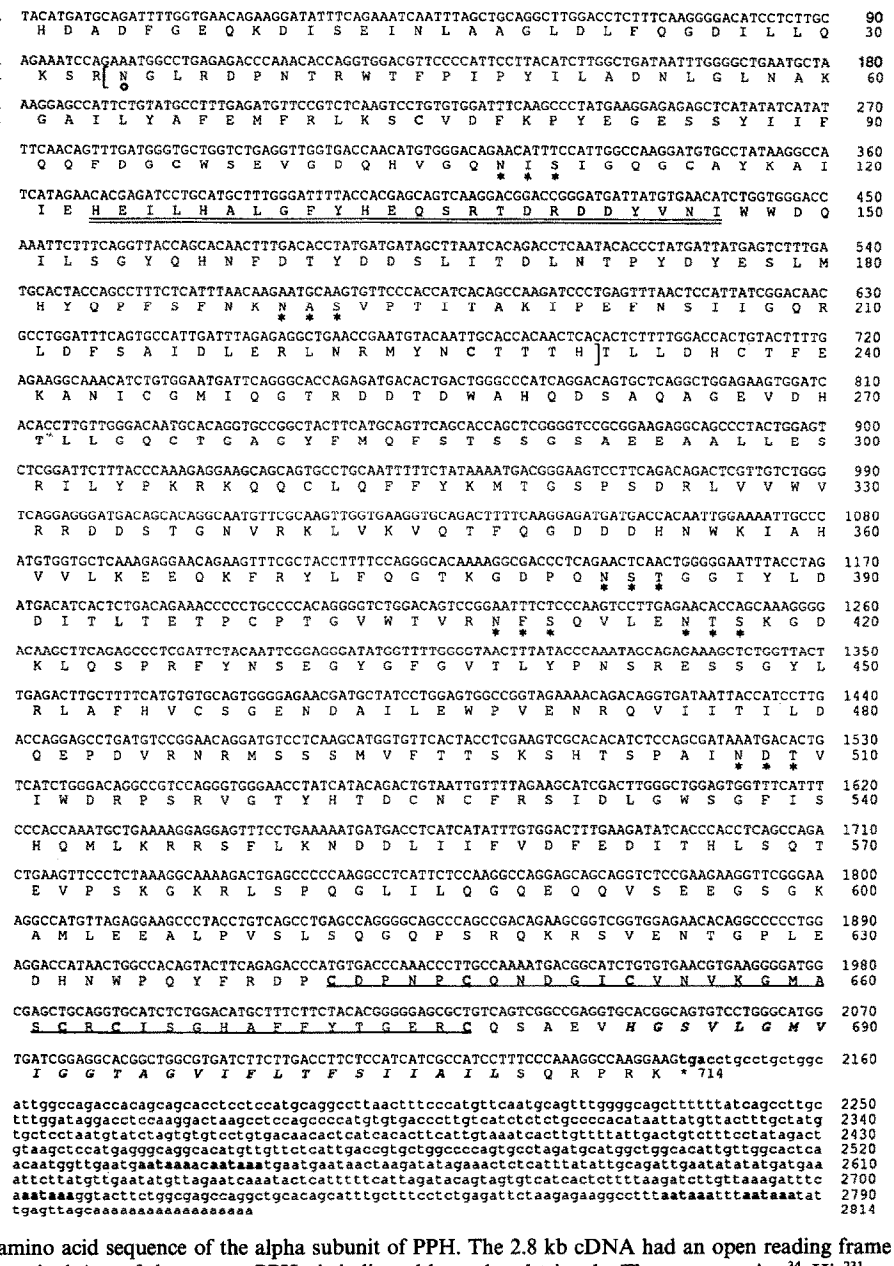

Fig. 2. cDNA and deduced amino acid sequence of the alpha subunit of PPH. The 2.8 kb cDNA had an open reading frame to the stop codon of 714 amino acids. The N-terminal Asn of the mature PPHα is indicated by a closed triangle. The sequence Asn^{34}–His^{231} represent the astacin domain (marked []). The extended astacin motif is double underlined. Cys^{642}–Cys^{677} (underlined by a solid line) show the EGF-like domain; the bold italic sequence is the putative membrane anchor domain. Conserved cysteines in the astacin and the EGF domain known to be involved in disulfide bonding are in bold type. Possible *N*-glycosylation sites are underlined by stars.

brane spanning domain and the astacin domain harbours other known motifs. One of these is an EGF-like domain spanning from Cys^{642} to Cys^{677} (36 amino acids), that contains six conserved cysteine residues. Another domain of about 170 amino acids immediately following the astacin domain has recently been described as an adhesive domain [30]. In addition, PPHα contains six potential *N*-glycosylation sites, two within the protease domain, and four within the intervening sequence. The domain structure of PPHα is depicted in Fig. 3.

3.4. *Southern blot analysis*

For Southern blot analysis, genomic DNA was isolated from the human colon carcinoma cell line Caco-2 and probed with PPHα1 after digestion with restriction nucleases. The results are shown in Fig. 4 and revealed only one or two bands per restriction digest, indicating that the gene encoding PPHα is a single copy gene.

3.5. *Expression of PPHα in human tissue*

The expression of PPHα mRNA in different human tissues was investigated using Northern blot analysis. Screening of mRNA from human heart, brain, placenta,

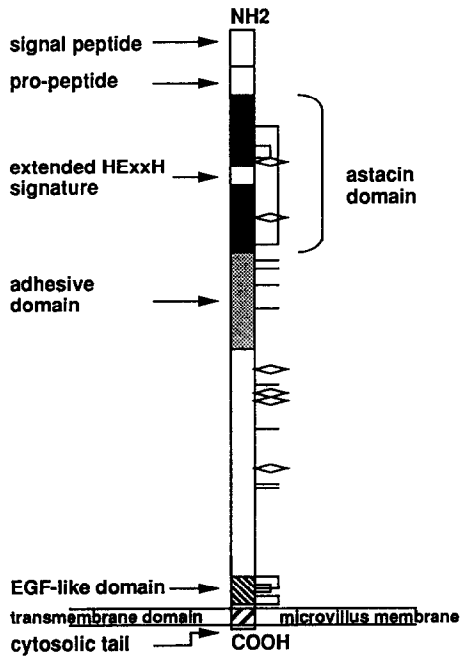

Fig. 3. Domain structure of the alpha subunit of PPH. The schematic representation of the proposed domain structure of PPHα is drawn to scale and the different domains indicated by arrows. <>, potential *N*-glycosylations sites; −, cysteine residues.

lung, liver, skeletal muscle, kidney and pancreas with a PPHα cDNA probe gave a negative result for all these tissues, whereas positive signals were obtained with an actin probe (data not shown). Northern blot analysis using mRNA prepared from human small intestinal mucosa, large intestinal mucosa, large intestinal tumor tissue, human kidney and MDCK cells is shown in Fig. 5. The strongest signal using a PPHα-specific probe (PPHα22.4) was obtained with mRNA from small intestinal mucosa. Weaker signals were obtained with mRNA from large intestinal mucosa and also large intestinal tumor tissue. No signal was obtained with mRNA from human kidney or MDCK cells. This suggests that the expression of PPHα is very tissue-specific being restricted to the gastrointestinal tract in human.

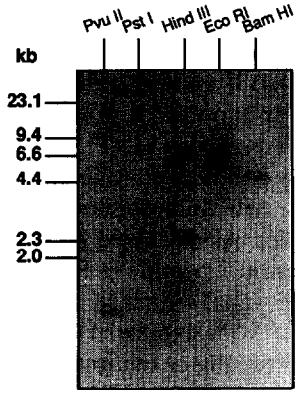

Fig. 4. Southern blot analysis. For this analysis, genomic DNA prepared from the human adenocarcinoma cell line Caco-2 was used. 7 μg DNA were digested with restriction enzymes indicated on the top of each lane and probed with PPHα1 cDNA. DNA sizemarkers are indicated on the left.

3.6. *Expression of meprin-PPHα (MPα) in COS-1 cells*

The expression of the PPHα subunit alone in a cell line was carried out to determine if homodimer formation occured. Due to the lack of the 5′-end in our PPHα cDNA we constructed a chimeric cDNA (MPα) which included the 5′-end from the alpha subunit of mouse meprin, encoding the starting methionine and a putative signal peptide sequence (Fig. 6). In vitro transcription/translation of this cDNA construct in rabbit reticulocytes resulted in a single product with an apparent M_r = 90 kDa (not shown). The products immunoisolated from transfected COS-1 cells after continuous labelling with [^{35}S]methionine for 5 h are shown in Fig. 7A and B. Analysis of the expressed protein by SDS-PAGE under reducing conditions revealed two bands of $M_r \approx$ 100 kDa and 90 kDa (Fig 7A, lane 1). Treatment of the immunoprecipitated MPα with endoglycosidase H (Fig. 7A, lane 2) and endoglycosidase F (Fig. 7A, lane 3) respectively, showed a shift of the upper band to the lower band (90 kDa). Metabolic labelling of transfected COS-1 cells in the presence of tunicamycin, an inhibitor of *N*-linked glycosylation, only resulted in the 90 kDa species being synthesized (Fig. 7A, lanes 4 and 5). These results indicate, that the two synthesized MPα proteins detected after 5 h of labelling correspond to a non-glycosylated form (M_r = 90 kDa) and an immature high-mannose form ($M_r \approx$ 100 kDa). Fig. 7B shows the SDS-PAGE analysis of the expressed MPα under reducing conditions (lane 1) and under non-reducing conditions (lane 2). It shows that a large amount of MPα, when analyzed under non-reducing conditions was in an apparent dimeric form. A considerable amount of proteins, however, did not enter the gel properly, indicating the presence of oligomers or aggregates. MPα could also be immunoprecipitated in an endo H resistent form from the culture medium of transfected COS-1 cells (data not shown).

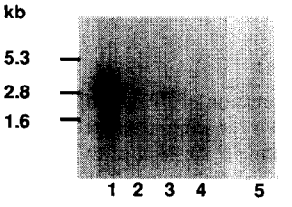

Fig. 5. Northern blot analysis. mRNA (≈ 1 μg/lane) from human small intestine (lane 1), human colon (lane 2), human colon tumor (lane 3), MDCK cells (negative control, lane 4), and human kidney (lane 5) was electrophoresed and blotted onto nylon membrane. Hybridization was with a random labelled ^{32}P-PPHα cDNA probe as described in methods. RNA sizemarkers are indicated on the left.

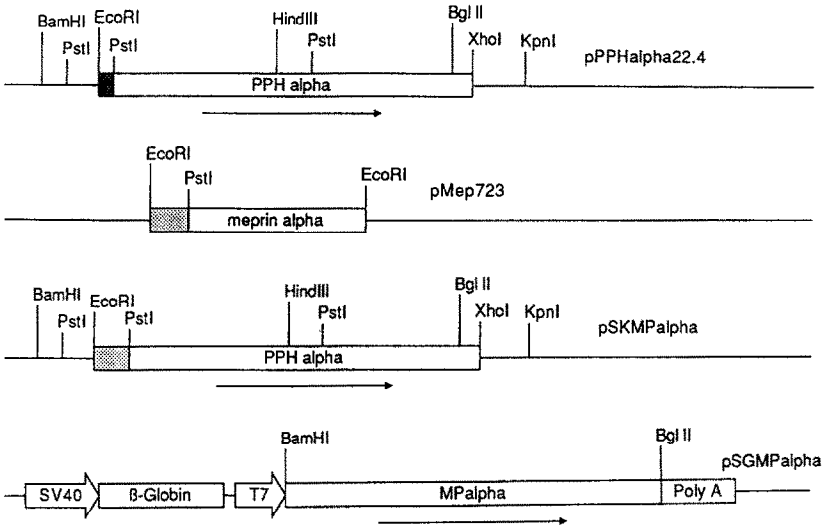

Fig. 6. Construction of a meprin-PPHα expression vector (pSGMPα). To construct this full-length chimera, a 50 bp *Eco*RI–*Pst*I fragment from pPPHα22.4 was replaced by a 200 bp *Eco*RI–*Pst*I fragment from the alpha subunit of mouse meprin cDNA (pMep723) (for details see experimental procedures). The chimeric cDNA (pSKMPα) was subcloned into the pSG5 vector, thus creating the pSGMPα expression vector.

4. DISCUSSION

The cDNA deduced amino acid sequence of the mature alpha subunit of PPH (amino acid 34–714 in Fig. 2) comprises 681 amino acids and has a calculated molecular weight of 80,565. As previously reported [6] it contains the astacin protease domain with an identity score of 31% to actacin and 82% to meprinα. Comparison to the recently published X-ray crystal structure of astacin [31,32] suggests that all the residues involved in the proposed pentavalent zinc-binding site are conserved in PPHα. This zinc-coordination site, involving

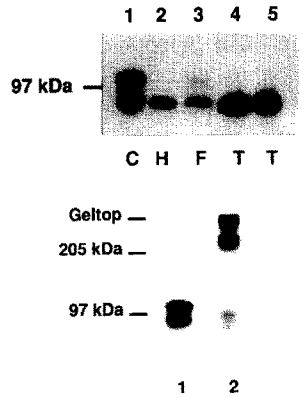

Fig. 7. Transient expression of PPHα (MPα chimera) in COS-1 cells. The pSGMPα vector was used to transfect COS-1 cells (for details see experimental procedures). 40–48 h after transfection, cells were labelled for 5 h with [^{35}S]methionine, PPHα was immunoprecipitated using a specific polyclonal antibody and analyzed by SDS-PAGE (7.5%). (A) SDS-PAGE under reducing conditions. Lane 1, control; lane 2, treatment of immunoprecipitate with endo H; lane 3, treatment of immunoprecipitate with endo F; lanes 4 and 5, metabolic labelling in the presence of 5 μg/ml and 10 μg/ml of tunicamycin respectively. (B) SDS-PAGE under reducing conditions (lane 1) and under non-reducing conditions (lane 2).

Tyr^{182}, (equivalent to Tyr^{149} in astacin) is different from other zinc-proteases such as thermolysin and carboxypeptidase A [33,34]. Based on the zinc-binding site, Jiang and Bond [35] have recently proposed a scheme for the classification of metalloproteases into five distinct families, namely thermolysin, astacin, serratia, matrixin and snake venom metalloproteinases. Also conserved in PPHα are four cysteine residues which have been shown to be involved in disulfide-bonding in astacin [5]. On the basis of the available X-ray crystal and sequence data, Stöcker et al. [32] have compared astacin with the alpha subunit protease domain of mouse meprin and have shown a close structural relationship of these two proteases.

The deduced amino acid sequence of PPHα reported here shows 85% similarity/75% identity to the alpha subunit of mouse meprin [13] and 86% similarity/78% identity to the alpha subunit of rat meprin [15] (Fig. 8). Compared to the primary sequences of the α subunits of meprin from mouse and rat, our protein is short of an estimated 44 [13] and 32 amino acids [15] respectively. Both rat and mouse meprin alpha subunits have a methionine in position 14 (Fig. 8). The additional methionine in the mouse enzyme (position 1, Fig. 8), suggests that the two alpha subunits are different in the two rodent species. As discussed by Jiang et al. [13] the missing NH_2-terminal peptide in PPHα probably represents a cleavable signal sequence. Amino acid sequence determination of MPα expressed in MDCK cells have clearly established this experimentally [36]. The 33 amino acids in front of the NH_2-terminal Asn in our sequence therefore constitute part of a potential propeptide sequence. The peptide bond Arg^{33}–Asn^{34} is compatible with a cleavage site for a trypsin-like protease. It has been shown that mouse meprin may be activated by

```
PPHalpha-H  .......... .......... .......... .......... ....H.DADF
Mepalpha-M  MARRLGRSSS FAIMLWIQPA CLLSLIFSAH IAAVSIKHLL NGSDH.DTDV
Mepalpha-R  .......... ...MLWTLPV CLLSLSFSAH IAAVSIQHLS TGHDHDDVDV
 Consensus  .......... ...mlw..p. cllsl.fsah iaavsi.hl. .g.dHdD.Dv  50

PPHalpha-H  GE.QKDISEI NLAAGLDLFQ GDILLQKSRN GLRDPNTRWT FPIPYILADN
Mepalpha-M  GE.QKDIFEI NLAAGLNLFQ GDILLPRTRN AMRDPSSRWK LPIPYILADN
Mepalpha-R  GEQQKDISEI NSAAGLNLFQ GDILLPRTRN ALRDPSSRWK PPIPYILADN
 Consensus  GE.QKDIsEI NlAAGLnLFQ GDILLprtRN alRDPssRWk .PIPYILADN  100

PPHalpha-H  LGLNAKGAIL YAFEMFRLKS CVDFKPYEGE SSYIIFQQFD GCWSEVGDQH
Mepalpha-M  LELNAKGAIL HAFEMFRLKS CVDFKPYEGE SSYIIFQKLS GCWSMIGDQQ
Mepalpha-R  LDLNAKGAIL NAFEMFRLKS CVDFKPYEGE SSYIIFQQFS GCWSMVGDQH
 Consensus  L.LNAKGAIL .AFEMFRLKS CVDFKPYEGE SSYIIFQqfs GCWSmvGDQh  150

PPHalpha-H  VGQNISIGQG CAYKAIIEHE ILHALGFYHE QSRTDRDDYV NIWWDQILSG
Mepalpha-M  VGQNISIGEG CDFKATIEHE ILHALGFFHE QSRTDRDDYV NIWWDQIITD
Mepalpha-R  VGQNISTGEG CDYKAIIEHE ILHALGFFHE QSRTDRDDYV NIWWNEIMTD
 Consensus  VGQNISIGeG CdyKAiIEHE ILHALGFfHE QSRTDRDDYV NIWWdqI.td  200

PPHalpha-H  YQHNFDTYDD SLITDLNTPY DYESLMHYQP FSFNKNASVP TITAKIPEFN
Mepalpha-M  YEHNFNTYDD NTITDLNTPY DYESLMHYGP FSFNKNESIP TITTKIPEFN
Mepalpha-R  YEHNFNTYDD KTITDLNTPY DYESLMHYGP FSFNKNETIP TITTKIPEFN
 Consensus  YeHNFnTYDD .tITDLNTPY DYESLMHYgP FSFNKNesiP TITtKIPEFN  250

PPHalpha-H  SIIGQRLDFS AIDLERLNRM YNCTTTHTLL DHCTFEKANI CGMIQGTRDD
Mepalpha-M  TIIGQLPDFS AIDLIRLNRM YNCTATHTLL DHCDFEKTNV CGMIQGTRDD
Mepalpha-R  AIIGQRLDFS ATDLTRLNRM YNCTRTHTLL DHCAFEKTNI CGMIQGTRDD
 Consensus  .IIGQrlDFS AiDL.RLNRM YNCT.THTLL DHC.FEKtNi CGMIQGTRDD  300

PPHalpha-H  TDWAHQDSAQ AGEVDHTLLG QCTGAGYFMQ FSTSSGSAEE AALLESRILY
Mepalpha-M  ADWAHGDSSQ PEQVDHTLVE QCKGAGYFMF FNTSLGARGE AALLESRILY
Mepalpha-R  ADWVHEDSSQ PGQVDHTLVG RCKAAGYFMY FNTSSGVTGE VALLESRILY
 Consensus  aDWaH.DSsQ pgqVDHTLvg qCkgAGYFM. FnTSsG..gE aALLESRILY  350

PPHalpha-H  PKRKQQCLQF FYKMTGSPSD RLVVWVRRDD STGNVRKLVK VQTFQGDDDH
Mepalpha-M  PKRKQQCLQF FYKMTGSPAD RFEVWVRRDD NAGKVRQLAK IQTFQGDSDH
Mepalpha-R  PKRKQQCLQF FYKMTGSPSD RLLIWVRRDD NTGNVRQLAK IQTFQGDSDH
 Consensus  PKRKQQCLQF FYKMTGSPsD Rl.vWVRRDD ntGnVRqLaK iQTFQGDsDH  400

PPHalpha-H  NWKIAHVVLK EEQKFRYLFQ GTKGDPQNST GGIYLDDITL TETPCPTGVW
Mepalpha-M  NWKIAHVTLN EEKKFRYVFL GTKGDPGNSS GGIYLDDITL TETPCPAGVW
Mepalpha-R  NWKIAHVTLN EEKKFRYVFQ GTKGDPGNSD GGIYLDDITL TETPCPTGVW
 Consensus  NWKIAHVtLn EEkKFRYvFq GTKGDPgNS. GGIYLDDITL TETPCPtGVW  450

PPHalpha-H  TVRNFSQVLE NTSKGDKLQS PRFYNSEGYG FGVTLYPNSR .ES.SGYLRL
Mepalpha-M  TIRNISQILE NTVKGDKLVS PRFYNSEGYG VGVTLYPNGR ITSNSGLLGL
Mepalpha-R  TIRNISQVLE NTVKGDRLVS PRFYNSEGYG FGVTLYPNGR ITSNSGYLGL
 Consensus  TiRNiSQvLE NTvKGDkLvS PRFYNSEGYG fGVTLYPNgR itSnSGyLgL  500

PPHalpha-H  AFHVCSGEND AILEWPVENR QVIITILDQE PDVRNRMSSS MVFTTSKSHT
Mepalpha-M  TFHLYSGDND AILEWPVENR QAIMTILDQE ADTRNRMSLT LMFTTSKNQT
Mepalpha-R  AFHLYSGDND VILEWPVENE QAIMTILDQE PDARNRMSLS LMFTTSKYQT
 Consensus  aFHlySGdND aILEWPVENr QaImTILDQE pD.RNRMSls lmFTTSK.qT  550

PPHalpha-H  SPAINDTVIW DRPSRVGTYH TDCNCFRSID LGWSGFISEQ MLKRRSFLKN
Mepalpha-M  SSAINGSVIW DRPSKVGVYD KDCDCFRSLD WGWGQAISEQ LLKRRNFLKG
Mepalpha-R  SSAINGSVIW DRPTKVGVYD KDCDCFRSID WGWGQAISEQ MLMRRNFLKD
 Consensus  SsAINgsVIW DRPskVGvYd kDCdCFRSiD wGWgqaISEQ mLkRRnFLK.  600

PPHalpha-H  DDLIIFVDFE DITHLSQTEV PSKGKRLSPQ GLILQGQEQQ VSEEGSGKAM
Mepalpha-M  DSLIIFVDFK DLTHLNRTEV PASARSTMPR GLLLQGQESP ALGESSRKAM
Mepalpha-R  DTLIIFVDFK DLTHLRQTEV PISSRSVIPR GLLLQGQEPL ALGD.SRIAM
 Consensus  D.LIIFVDFk DiTHL.qTEV P.s.rs..Pr GLiLQGQE..  alge.SrkAM  650

PPHalpha-H  LEEALPVSLS QGQPSRQKRS VENTGPLEDH NWPQYFRDPC DPNPCQNDGI
Mepalpha-M  LEESLPSSLG QRHPSRQKRS VENTGPMEDH NWPQYFRDPC DPNPCQNEGT
Mepalpha-R  MEESLPRRLD QRQPSRPKRS VENTGPMEDH NWPQYFRDPC DPNPCQNEGT
 Consensus  lEEsLP.sL. QrqPSRqKRS VENTGPmEDH NWPQYFRDPC DPNPCQNeGt  700

PPHalpha-H  CVNVKGMASC RCISGHAFFY TGERCQSAEV HGSVLGMVIG GTAGVIFLTF
Mepalpha-M  CVNVKGMASC RCVSGHAFFY AGERCQAMHV HGSLLGLLIG CIAGLIFLTF
Mepalpha-R  CVNVKGMASC RCVSGHAFFY TGERCQAMHV HGSLLGLLIG CITALIFLTF
 Consensus  CVNVKGMASC RCvSGHAFFY tGERCQamhV HGSlLGllIG ciaglIFLTF  750

PPHalpha-H  SIIAILSQRP RK
Mepalpha-M  VTFSTTNGKL RQ
Mepalpha-R  ITFSNTYQKL RQ
 Consensus  .tfs.t.qkl Rq    762
```

Fig. 8. Comparison of the cDNA deduced amino acid sequences from PPHα, and meprinα of mouse and rat. Complete primary sequence of the three enzymes, including a consensus sequence is shown. Abbreviations: PPH, PABA peptide hydrolase; Mep, meprin; and -H, human; R, rat; M, mouse.

the treatment with trypsin in vitro [18]. PPHα when expressed in COS-1 or MDCK cells is not enzymatically active but upon treatment with trypsin can be activated by conversion to the mature form [36]. It has been proposed that the NH_2-terminal alanine in astacin forms a salt-bridge with Glu^{103} [32,37]. In analogy, the Asn found at the NH_2-terminus of mature PPHα could form such a salt-bridge with Glu^{134}. The removal of the propeptide therefore constitutes an essential step in the activation process of these enzymes.

Another feature present in PPH and meprin is an EGF-like domain, comprising 36 amino acids (residues Cys^{642} to Cys^{677} in PPHα) and containing six conserved cysteine residues. As reviewed by Appella et al. [38], these cysteines are involved in the formation of disulfide bonds, resulting in a common structural folding shared by the EGF-like domains. EGF is involved in proliferative and developmental effects which are mediated by a specific surface receptor endowed with protein tyrosine kinase activity [39,40]. Single or multiple copies of

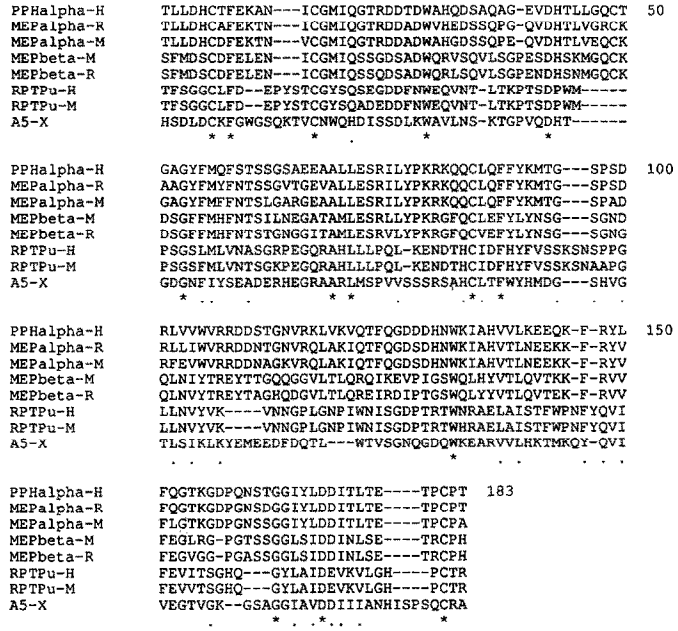

Fig. 9. Alignment of the adhesive (MAM) domains. The sequences were aligned using CLUSTAL V. The strictly conserved residues are indicated by asterisks and the similar residues by dots. Abbreviations: RPTP, receptor protein tyrosine phosphatase; PPH, PABA peptide hydrolase; Mep, meprin; and -H, human; R, rat; M, mouse; X, *Xenopus*.

EGF-like domains have been found in many different proteins, including most astacin family members [13,32]. Many proteins with EGF-like domains are growth factors, cell surface receptors or receptor-like proteins. Proteases containing an EGF-like sequence characterized to date have been shown to be involved in the proteloytic activation of zymogen or precursor molecules. Regulatory functions have also been suggested for some astacin proteases. The tight association of BMP-1 with BMP-2 and BMP-3, both peptide growth factors of the TGF-β family, strongly suggests that it directly or indirectly regulates the activities of these factors [7]. This idea is strengthened by the relation of BMP-2 to the decapentaplegic gene (dpp), which plays a crucial part in pattern formation in *Drosophila* and which in turn has been suggested to be regulated by tolloid, another member of the astacin family of proteases [9].

Recently, Beckmann and Bork have described another domain, which is present in diverse plasma-membrane-bound proteins such as protein tyrosine phosphatase μ, A5 protein, and the alpha and the beta subunits of meprin [30]. This MAM domain (μ, A5, meprin) comprises about 170 amino acids and can also be found in PPHα (Fig. 9). The exact function of MAM is not clear, but it has been suggested that it is likely to have a common adhesive function in these proteins, and it is quite likely that the conserved cysteine residues play an important role. Four cysteines are conserved in all of the proteins in Fig. 9 and one more cysteine is conserved in all meprin/PPH proteases. This additional Cys residue in meprin/PPH might be important for inter-subunit interaction.

Tissue expression of PPHα in human appears to be limited to the gastrointestinal tract. A notable finding was the apparent lack of enzyme in human kidney. The alpha subunit of mouse meprin has been found in the kidney of random bred mice but not in mice which are deficient in meprin A [41]. The mepriα subunit was not detected in the mouse intestine, while in the rat, it was expressed in kidney and intestine. In rat kidney, the alpha subunit has been shown by in situ hybridization to be expressed only in specific cells of the juxtamedullary region [15] providing further evidence for the highly specific expression of this enzyme. Taken together, these data indicate, that mepriα/PPHα expression is tissue-specific and differs from one species to another.

Expression of the MPα chimeric protein in COS-1 cells has led to the isolation of an immature enzyme as demonstrated by the persisting sensitivity to endoglycosidase H, suggesting that the protein is not able to exit from the ER compartment in COS-1 cells. The expression data clearly show that homodimers of MPα are formed in the ER, which is in agreement with organ culture data using human small intestinal mucosa [3]. A soluble and partially endo H resistant form was immunoisolated from the medium of transfected COS-1 cells. These data indicate that MPα expressed in COS-1 cells is synthesized as transport-incompetent membrane-bound form which cannot be exported from the ER compartment. The observation of a fully glycosylated soluble form of MPα in the culture medium, suggests

that proteolytic events lead to the secretion of a truncated form of MPα.

This work establishes the relationship of the α subunits of PPH as an isoenzyme of meprin (EC 3.4.24.18). As a member of the astacin family of proteases, in addition to a zinc-protease domain, PPHα contains other functional domains such as the EGF-like motif and the adhesion domain (MAM). The function of PPH in human small intestine is not known, but the complex modular organization distinguishes this enzyme from other proteases involved in simple digestive processes. Considering the continuous process of cell proliferation, migration and differentiation that occurs in the intestinal mucosa, it is conceivable that PPH is involved in the processing of precursor proteins or the regulation of growth factors. It might also be speculated, that PPH itself, due to its EGF-like domain, can function as a growth factor in intestinal morphogenesis. Detailed studies of the expression of PPH in different cells of the intestinal mucosa in combination with heterologous expression of the PPHα and PPHβ subunits are in progress to address these questions.

Acknowledgements: We thank Drs. Y. Edwards and D. Swallow for the generous gift of the λgt11 cDNA library; Dr. B.L. Nichols for the λZap cDNA library, for his valued contribution in discussions and for reading the manscript, and E. Mandac and U. Luginbühl for excellent technical assistance. This work was supported by Grant No. 32.31280.91 from the Swiss National Science Foundation.

REFERENCES

[1] Sterchi, E.E., Green, J.R. and Lentze, M.J. (1982) Clin. Sci. 62, 557–560.
[2] Sterchi, E.E., Naim, H.Y., Lentze, M.J., Hauri, H.P. and Fransen, J.A. (1988) Arch. Biochem. Biophys. 265, 105–118.
[3] Sterchi, E.E., Naim, H.Y. and Lentze, M.J. (1988) Arch. Biochem. Biophys. 265, 119–127.
[4] Bond, J.S. and Beynon, R.J. (1986) Curr. Top. Cell. Reg. 28, 263–290.
[5] Titani, K., Torff, H.J., Hormel, S., Kumar, S., Walsh, K.A., Rodl, J., Neurath, H. and Zwilling, R. (1987) Biochemistry 26, 222–226.
[6] Dumermuth, E., Sterchi, E.E., Jiang, W., Wolz, R.L., Bond, J.S., Flannery, A.V. and Beynon, R.J. (1991) J. Biol. Chem. 266, 21381–21385.
[7] Wozney, J.M., Rosen, V., Celeste, A.J., Mitsock, L.M., Whitters, M.J., Kriz, R.W., Hewick, R.M. and Wang, E.A. (1988) Science 242, 1528–1534.
[8] Sato, S.M. and Sargent, T.D. (1990) Dev. Biol. 137, 135–141.
[9] Shimell, M.J., Ferguson, E.L., Childs, S.R. and O'Connor, M.B. (1991) Cell 67, 469–481.
[10] Lepage, T., Ghiglione, C. and Gache, C. (1992) Development 114, 147–163.
[11] Reynolds, S.D., Angerer, L.M., Palis, J., Nasir, A. and Angerer, R.C. (1992) Develop. 114, 769–786.
[12] Yasumasu, S., Yamada, K., Akasaka, K., Mitsunage, K., Iuchi, I., Shimada, H. and Yamigami, K. (1992) Dev. Biol. 153, 250–258.
[13] Jiang, W., Corbea, C.M., Flannery, A.V., Beynon, R.J., Grant, G.A. and Bond, J.S. (1992) J. Biol. Chem. 267, 9185–9193.
[14] Gorbea, C.M., Marchand, P., Jiang, W., Copeland, N.G., Gilbert, D.J., Jenkins, N.A. and Bond, J.S. (1993) J. Biol. Chem. 268, 21035–21043.
[15] Corbeil, D., Gaudoux, F., Wainwright, S., Ingram, J., Kenny, A.J., Boileau, G. and Crine, P. (1992) FEBS Lett. 309, 203–208.
[16] Johnson, G.D. and Hersh, L.B. (1992) J. Biol. Chem. 267, 13505–13512.
[17] Johnson, G.D. and Hersh, L.B. (1993) J. Biol. Chem. 268, 17647.
[18] Kounnas, M.Z., Wolz, R.L., Gorbea, C.M. and Bond, J.S. (1991) J. Biol. Chem. 266, 17350–17357.
[19] Gorbea, C.M., Flannery, A.V. and Bond, J.S. (1991) Arch. Biochem. Biophys. 290, 549–553.
[20] Lathe, R. (1985) J. Mol. Biol. 183, 1–12.
[21] Berger, S.L. and Kimmel, A.R. (1987) Meth. Enzymol. 152, Academic Press, San Diego.
[22] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[23] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) in: Molecular Cloning, a Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, New York.
[24] Smith, D.B. and Johnson, K.S. (1988) Gene 67, 31–40.
[25] Harlow, E. and Lane, D. (1988) In: Antibodies, a Laboratory Manual, Cold Spring Harbour Laboratory, New York.
[26] Hauri, H.P., Sterchi, E.E., Bienz, D., Fransen, J.A. and Marxer, A. (1985) J. Cell Biol. 101, 838–851.
[27] Lottaz, D., Oberholzer, T., Bähler, P., Semenza, G. and Sterchi, E.E. (1992) FEBS Lett. 313, 270–276.
[28] Grünberg, J., Kruppa, A., Paschen, P. and Kruppa, J. (1991) Virol. 180, 678–686.
[29] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[30] Beckmann, G. and Bork, P. (1993) Trends Biochem. Sci. 18, 40–41.
[31] Gomis-Rüth, F.-X., Stöcker, W., Huber, R., Zwilling, R. and Bode, W. (1993) J. Mol. Biol. 229, 945–968.
[32] Stöcker, W., Gomis-Rüth, F.-X., Bode, W. and Zwilling, R. (1993) Eur. J. Biochem. 214, 215–231.
[33] Matthews, B.W. (1988) Acc. Chem. Res. 21, 333–340.
[34] Christiansen, D.N. and Lipscomb, W.N. (1989) Acc. Chem. Res. 22, 62–69.
[35] Jiang, W. and Bond, J.S. (1992) FEBS Lett. 312, 110–114.
[36] Grünberg, J., Eldering, J.A., Dumermuth, E. and Sterchi, E.E. (1993) FEBS Lett. 335, 376–379.
[37] Bode, W., Gomis, R.F.X., Huber, R., Zwilling, R. and Stocker, W. (1992) Nature 358, 164–167.
[38] Appella, E., Weber, I.T. and Blasi, F. (1988) FEBS Lett. 231, 1–4.
[39] Carpenter, G. and Cohen, S. (1979) Ann. Rev. Biochem. 48, 193–216.
[40] Heldin, C.-H. and Westermark, H. (1984) Cell 37, 9–20.
[41] Jiang, W., Sadler, P.M., Jenkins, N.A., Gilbert, D.J., Copeland, N.G. and Bond, J.S. (1993) J. Biol. Chem. 268, 10380–10385.