Human carboxypeptidase E

Isolation and characterization of the cDNA, sequence conservation, expression and processing *in vitro*

Edward MANSER,*† Denise FERNANDEZ,* Lynda LOO,* Phuy Yee GOH,* Clinton MONFRIES,† Christine HALL† and Louis LIM*†

*Institute of Molecular and Cell Biology, National University of Singapore, Kent Ridge, Singapore 0511, and †Department of Neurochemistry, Institute of Neurology, 1 Wakefield Street, London WC1N 1PJ, U.K.

Carboxypeptidase E (CPE), which cleaves C-terminal amino acid residues and is involved in neuropeptide processing, is itself subject to intracellular processing. Human CPE cDNA was isolated and sequence comparisons were made with those of a previously isolated brain cDNA (M1622) encoding rat CPE and of other human carboxypeptidases (M and N). Human (2.5 kb) and rat (2.1 kb) CPE cDNAs approximated to the size of their respective mRNAs; additional sequences were located in putative 5' and 3' untranslated regions of human CPE mRNA. There is 79% sequence similarity between human and rat CPE cDNAs, with greater similarity (89%) over the coding region and short sections of the non-coding sequence. The predicted 476-amino acid-residue sequences of human and rat preproCPEs are highly conserved (96% identity), with lower degree of similarity of the N-terminal signal peptide (76%). Human CPE showed 51% and 43% sequence similarity to human CPN and CPM respectively, with discrete regions of divergence dispersed between the highly conserved mechanistically implicated regions. Antiserum generated from a fusion protein, synthesized in *Escherichia coli* from constructs of the human cDNA, recognized an approx. 50 kDa membrane protein and a smaller soluble protein in rat and human brain preparations, corresponding to the two forms of native CPE. Human CPE mRNA transcripts directed the synthesis in reticulocyte lysate of a 54 kDa translation product, which in the presence of dog pancreas microsomal membranes was co-translationally processed with cleavage, insertion into membranes and glycosylation. Three processed forms were generated, the largest (56 kDa) and smallest (52 kDa) being equally glycosylated. The membrane association of the processed translation products and of native brain membrane CPE, detected immunologically, was resistant to moderate alkali but not pH 11.5 extraction. These results are consistent with secondary-structure predictions that CPE is a peripheral membrane protein. The dissimilar regions of human carboxypeptidases may provide information on sequences responsible for their different cellular disposition.

INTRODUCTION

One of the key enzymes involved in the biosynthesis of neuropeptides is carboxypeptidase E (CPE), originally designated enkephalin convertase (Fricker & Snyder, 1982) or carboxypeptidase H (Fricker, 1988), which specifically cleaves C-terminal basic amino acid residues. CPE activity is present in endocrine tissues, but highest concentrations are found in the brain and pituitary gland (Strittmatter *et al.*, 1984). The activity is localized to secretory granules in which CPE exists as membrane and soluble forms differing in molecular mass by 2–3 kDa (Supattapone *et al.*, 1984), with their relative distribution in the brain being region-specific (Strittmatter *et al.*, 1984). Separate forms of soluble and membrane CPE also occur in adrenal-gland chromaffin granules (Laslop *et al.*, 1986). Heterogeneity of CPE may be an important factor in neuropeptide processing; membrane and soluble CPE forms that have identical properties may fulfil different roles in neuropeptide processing as a consequence of their cellular location. The cDNAs for two other human carboxypeptidases, CPM (Tan *et al.*, 1989) and CPN (Gebhard *et al.*, 1989), have recently been characterized. These enzymes are respectively located in cell membranes and in the circulation, where this CPN catalytic subunit is found in association with a larger inactive subunit (Plummer & Hurwitz, 1978).

Many membrane proteins and peptide precursors are synthesized by polyribosomes bound to membranes of the endoplasmic reticulum, where the first stages of glycosylation and other post-translational processing events occur. In our previous investigations of this class of brain membrane proteins we isolated and characterized two unique developmentally regulated cDNAs (Hall *et al.*, 1987). One of these (M1622 cDNA), corresponding to a 2.2 kb mRNA expressed in brain and adrenal gland, can be identified as rat CPE on the basis of its sequence similarity to bovine and rat CPEs (Fricker *et al.*, 1986, 1989; Rodriguez *et al.*, 1989). Rat CPE mRNA is developmentally regulated in the brain (Hall *et al.*, 1987), being detectable at embryonic day 17 and increasing to maximal concentrations 10 days after birth, concordant with the period of cellular differentiation and establishment of functional synapses. Interestingly, CPE mRNA was expressed in primary glial as well as neuronal cultures (Hall *et al.*, 1987; Lowndes, 1988). CPE is a glycoprotein, which is synthesized as a larger inactive precursor (Fricker, 1988). It is exclusively synthesized by membrane-bound polyribosomes in the brain (Hall *et al.*, 1987). The subsequent post-translational processing steps that give rise to the active membrane and soluble forms of the enzyme and the sorting mechanism by which CPE is targeted to secretory granules have not been determined. In the present paper we describe the isolation of the

Abbreviations used: CPE, carboxypeptidase E; CPM, carboxypeptidase M; CPN, carboxypeptidase N.

† To whom requests for reprints should be sent.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers X51405 (human CPE) and X51406 (rat CPE).

cDNA of human CPE, its extensive sequence similarity to rat CPE, CPE membrane association and the primary events in its biosynthesis.

MATERIALS AND METHODS

Isolation of human CPE cDNA

A 1.9 kb *Pst*I fragment of the rat CPE clone M16225E (Hall *et al.*, 1987; Fig. 1) extending from nucleotide 295 to the 3'-end was used to screen a human retinal cDNA library in bacteriophage λ gt10 (Nathans *et al.*, 1986). Filters containing bacteriophage λ gt10 cDNAs were prehybridized at 60 °C for 4 h with $6 \times \text{SSC}/6 \times \text{Denhardt's}$ solution containing 100 μg of denatured salmon sperm DNA/ml. Hybridization was at 60 °C for 16 h and filters were washed in $2 \times \text{SSC}/0.1\%$ SDS at the same temperature, followed by $0.2 \times \text{SSC}/0.1\%$ SDS and autoradiographed. ($1 \times \text{SSC}$ is 0.15 M-NaCl/15 mM-sodium citrate buffer, pH 7; $1 \times \text{Denhardt's}$ solution is 0.02% Ficoll 400/0.02% polyvinylpyrrolidone/0.02% BSA.) In all 26 human cDNAs (designated A1622–Z1622) isolated from 5×10^4 plaques were purified, and cDNA inserts were excised by *Eco*RI digestion and subcloned into Bluescript SK(+/-). Restriction-enzyme digests of subcloned cDNAs were compared after electrophoresis on 1% agarose gels.

Nucleotide sequencing

Rat M16225E/CPE DNA sequence was determined (Sanger *et al.*, 1977) either after subcloning into M13 or by using Bluescript SKM(+/-) and exonuclease III and mung-bean-nuclease-nested deletions according to established protocols (Stratagene). Human CPE cDNA (P1622) was sequenced after random cloning of blunt-ended fragments generated by sonication or restriction-enzyme digestion (*Alu*I, *Hae*III or *Rsa*I) into M13. The sequence of G+C-rich regions at the 5'-end of the clones was confirmed by using the dG analogue 7-deaza-2'-dG (Boehringer Mannheim). In all cases sequences of both strands were determined. Compilation and analysis of the nucleotide sequence data were performed by using DNASTAR programs.

CPE- β -galactosidase fusion protein: purification of 17 kDa CPE peptide and its use as an immunogen

A CNBr-generated peptide from a human CPE fusion protein was used to raise specific rabbit antiserum. A human CPE- β -galactosidase fusion protein construct in pUR292 (Rüther & Müller-Hill, 1983) was produced by cloning a 1100 bp *Dra*I fragment from CPE cDNA (P1622) into the pUR292 *Bam*HI site filled in with Klenow polymerase. Human CPE- β -galactosidase fusion protein (150 kDa) was synthesized after isopropyl β -thiogalactoside induction in *Escherichia coli* BMH 71-18 transformed with recombinant CPE/pUR292 plasmid (Maniatis *et al.*, 1982). Fusion protein (or β -galactosidase control protein) was purified from *E. coli* after lysis in 0.2 M-Tris/HCl buffer, pH 7.5, containing 10 mM-magnesium acetate, 5% (v/v) glycerol, 10 mM-2-mercaptoethanol, 0.25 M-NaCl and 0.2 mg of lysozyme/ml at 4 °C for 30 min. Preparations were diluted with 5 vol. of 6 M-urea/1 mM-EDTA/1 mM-phenylmethanesulphonyl fluoride/1% Triton X-100 and sonicated. Inclusion bodies containing the fusion protein (or control β -galactosidase) resistant to this treatment were pelleted by centrifugation (30000 g_{av} for 20 min). The pellet was dissolved in 0.5 vol. of Laemmli (1970) sample buffer at 60 °C, and non-solubilized material was removed by centrifugation. Proteins were precipitated with acetone, dried and redissolved in 70% (v/v) formic acid. After cleavage with CNBr (10 mg/ml) by incubation

under N_2 at 20 °C for 20 h, peptides were resuspended in 3 M-urea/50 mM-ammonium formate buffer, pH 6.8, and purified by chromatography on a Mono Q Sepharose ion-exchange column (Pharmacia). A CNBr-cleavage peptide at 17 kDa from the C-terminus of CPE predicted from sequence data (see the Results section and Fig. 3) with the expected co-ordinates was detected on two-dimensional gel-electrophoretic analysis. This was absent from digests of β -galactosidase alone. The 17 kDa fragment eluted from Mono Q Sepharose between 0.5 M- and 1.0 M-NaCl was further purified by SDS/PAGE and electro-elution. Rabbits were injected subcutaneously with the purified peptide (50 μg) in Freund's complete adjuvant, followed by injection of peptide (50 μg) in incomplete Freund's adjuvant every 4 weeks. Rabbits were bled from an ear vein 1 week after the third and subsequent injections to obtain anti-CPE serum.

Fractionation of brain CPE and Western blotting

Human or rat brain homogenates (10%, w/v) in 20 mM-Tris/acetate buffer, pH 6.0, were centrifuged at 1000 g for 3 min, and the pellets were rehomogenized in an equal volume of the buffer and again centrifuged. This hypo-osmotic treatment results in the release of 'soluble' CPE from granules (Supattapone *et al.*, 1984). The combined supernatants were centrifuged at 40000 rev./min in a Beckman Ti50 rotor for 30 min to separate soluble and membrane fractions. Western-blot analysis with anti-(human CPE peptide) serum (1:250) was by established procedures (Towbin *et al.*, 1979); biotinylated anti-(rabbit immunoglobulin) second antibody and streptavidin-horseradish peroxidase (Bethesda Research Laboratories) were used according to the manufacturer's protocol.

Transcription *in vitro*

The production of 'capped' RNA transcripts of H1622 and P1622 cDNAs by using the T7 polymerase promoter of the Bluescript plasmid was according to Stratagene protocols. mRNA transcripts of the correct size were synthesized, as determined after electrophoresis on 4% polyacrylamide gels in parallel with RNA molecular-mass markers (Bethesda Research Laboratories) and ethidium bromide staining.

Northern-blot analysis of RNA and translation *in vitro*

Purification, gel electrophoresis and Northern-blot analysis of RNA, translation of mRNA and protein processing in reticulocyte lysates containing dog pancreas microsomal membranes *in vitro* and SDS/PAGE analysis of translation products were as described in previous publications (Hall *et al.*, 1984, 1987).

Isolation of membranes and alkaline extraction

Microsomal membranes were isolated from reticulocyte lysates containing CPE translation products (10 μl) after dilution with 0.1 M-Tris/HCl buffer, pH 7.5, and sedimentation through a cushion of 0.25 M-sucrose/50 mM-Tris/HCl buffer, pH 7.5 (40 μl), in a Beckman Airfuge (A-100/30 rotor) at a gauge pressure of 140 kPa (20 lbf/in²) for 5 min. Supernatants and pellet fractions in 1% Triton X-100 were analysed by SDS/PAGE after the addition of an equal volume of Laemmli (1970) sample buffer. Alternatively, for membrane-association assays, microsomal membranes containing CPE translation products or purified brain membranes (10 mg/ml) were resuspended in 0.1 M-sodium carbonate solution of appropriate pH (see the legends to Figs. 5 and 7), then incubated at 0–4 °C for 30 min, and 100 nl portions were centrifuged through a sucrose cushion (40 μl) in an Airfuge at a gauge pressure of 140 kPa for 5 min. The pellets were resuspended in 1% Triton X-100 and Laemmli (1970) sample buffer and analysed by SDS/PAGE.

RESULTS

Isolation of human CPE

A previously isolated developmentally regulated rat brain cDNA (M1622; Hall *et al.*, 1987) can now be identified as that for CPE on the basis of its similarity to published sequences (Fricker *et al.*, 1986; Figs. 1 and 3). With the use of this rat cDNA, human CPE cDNA was isolated from a human retinal cDNA library (Nathans *et al.*, 1986; see the Materials and methods section). Restriction-enzyme digests of 26 cDNAs with sizes ranging from 1.1 to 2.5 kb showed the shorter cDNAs to contain common 3'-ends with variable 5'-extensions; these cDNAs were probably derived from the same mRNA species. A single polyadenylation site has been identified by sequencing a number of cDNAs. The two longest clones, of 2.5 kb, differed only in the extreme 5'-region. One cDNA (H1622) was 133 nucleotide residues shorter at the 5'-end, contained a longer poly(A) tract at the 3'-end and differed only in the first 24 nucleotide residues from the other cDNA (P1622; Fig. 1).

cDNA and predicted amino acid sequence of human CPE and comparison with rat CPE

The size of both rat (2091 nucleotide residues) and human (2436 nucleotide residues) CPE cDNAs approximated to the size of their respective mRNAs (2.2 kb for the rat mRNA and 2.5 kb for the human mRNA; Fig. 2; Hall *et al.*, 1987). The sequence of the human CPE cDNA (P1622) is shown in Fig. 1; additional sequences in the human mRNA were located at both 5' and 3' non-coding regions. The 2.1 kb rat (M1622) cDNA sequence was virtually identical with published rat sequences (cf. legend to Fig. 1; Rodriguez *et al.*, 1989; Fricker *et al.*, 1989) but contained an additional 5'-end 32 nucleotide residues and extended 73 nucleotide residues at the 3'-end to include the polyadenylation signal (Fig. 1). Primer extension (Fricker *et al.*, 1989) and S1 nuclease protection data (G. Rajendran, unpublished work) indicates that the 5'-end of this cDNA was approximately 15 nucleotide residues downstream of the transcriptional initiation site of rat CPE mRNA. Human and rat DNA sequences shared 79% identity, with greater similarity (89%) over the putative coding region and in several short sections of the non-coding regions, including 60 nucleotide residues preceding the polyadenylation site (Fig. 1). A continuous open reading frame extends from the start of the P1622 human cDNA to the termination codon at position 1720. The extreme 5'-sequence codes for an additional 96 amino acid residues in frame upstream of the probable initiator methionine. An open reading frame also runs from the start of the rat cDNA. However, the first methionine residue at nucleotide-290 in the human cDNA is part of an initiation consensus sequence (Kozak, 1987) that is conserved in rat, human and bovine CPE cDNAs (Fricker *et al.*, 1989). The 476-amino acid-residue open reading frames extending from this start site in human and rat sequences share 96% identity (Fig. 3). The predicted molecular mass of 53156 Da is consistent with the size of the primary translation product of hybrid-selected rat brain mRNA (Hall *et al.*, 1987).

The *N*-terminal 25 amino acid residues of human CPE constitute a typical signal sequence (von Heijne, 1986). This has lower interspecies similarity (76%) than the remainder of the protein (97%), in which most of the amino acid substitutions are conservative. The 16-amino acid-residue 'pro' sequence that precedes the *N*-terminus of the mature protein (Fricker *et al.*, 1986) is 100% conserved between human and rat; a polyarginine stretch present in all CPE sequences is likely to be an endopeptidase-cleavage site (Fricker, 1988). Two potential *N*-glycosylation sites at amino acids 139 and 390 are also conserved.

Human CPE is distantly related to carboxypeptidases A and B (15–20% identity), but has greatest similarity to two carboxypeptidases with a more specialized distribution (Fig. 4a), CPN (51% identity; Gebhard *et al.*, 1989) and CPM (43% identity; Tan *et al.*, 1989). Comparison of CPE, CPM and CPN with carboxypeptidases A and B, for which the crystal structures are known (Schmid & Herriott, 1976; Rees *et al.*, 1983), reveals conservation of the following amino acid residues implicated mechanistically (numbered according to CPE, Fig. 4a; carboxypeptidase A numbering in parentheses): Arg-189 (145), which interacts with the terminal carboxy group, Tyr-320 (248), the proposed proton donor, Glu-342 (270), whose side chain is involved in peptide-bond cleavage, and the residues His-114 (69), Glu-117 (72) and His-248 (196), thought to interact with Zn²⁺. These relationships have been previously noted (Tan *et al.*, 1989; Gebhard *et al.*, 1989). The *C*-terminal region of CPE, beyond the paired amino acid potential cleavage sites (Fricker *et al.*, 1986), shows greatest sequence divergence. This and dispersed stretches with lower similarity may contribute to the specialized distributions of these three enzymes. Comparison of the hydrophathy profiles of all three human carboxypeptidases (Fig. 4b) confirms their close similarity, particularly of CPE and CPN, and also illustrates that the divergent *C*-terminals have markedly different predicted secondary structure.

Characterization of CPE forms in the brain: membrane association of native CPE

The native brain CPE forms were investigated by using human CPE polyclonal rabbit antiserum directed against a specific 17 kDa peptide corresponding to residues 327–464 of CPE (Fig. 3; see the Materials and methods section). The antiserum specifically recognized a 50 kDa membrane protein in human brain preparations (Fig. 5, lane 1), as well as the corresponding rat protein (Fig. 5, lanes 4, 5 and 7) because of the sequence conservation. It also detected the smaller (48 kDa) native soluble CPE form from rat brain (Fig. 5, lane 2) or human brain (not shown). On two-dimensional gel Western-blot analysis the 50 kDa brain membrane protein had the same *pI* (approx. 5.1) but was slightly smaller than the CPE primary translation product, which could be immunoprecipitated with the antiserum (E. Manser, unpublished work). The bands corresponding to native soluble and membrane-bound CPE are quite broad, as is common for glycoproteins. Cross-reacting 26, 80 and 120 kDa brain proteins (Fig. 5) have not been further characterized, and may represent in the case of the 26 kDa band a proteolytic fragment of CPE. The membrane-bound CPE was almost completely resistant to freeze-thaw and further hypo-osmotic treatment (Fig. 5, cf. lanes 3 and 4).

The membrane association of the native membrane 50 kDa CPE was investigated by alkaline treatment, which removes peripheral membrane proteins and luminal contents at higher pH (Fujiki *et al.*, 1982). Rat brain membranes were subjected to increasing pH, and association of the 50 kDa CPE with the membrane was determined by Western-blot analyses (Fig. 5, lanes 3–8). Even at pH 10.1 a significant proportion of CPE remained membrane-associated, despite a substantial release (Fig. 5, lanes 6 and 7). The other cross-reacting proteins were completely removed at this pH. CPE was fully released from membranes treated at pH 11.5 (Fig. 5, lane 8), consistent with its being a peripheral but not an integral membrane protein. Several integral transmembrane proteins, resistant to extraction at pH 11.5, were present in these membranes, as determined by Coomassie Blue staining (result not shown). The CPE released into the supernatant fraction (Fig. 5, lanes 3 and 6) from the membrane is larger than the native soluble CPE form (Fig. 5, lane 2).

AAATGGCGTCCCGTCTCTCCGCGGGCCCTGCCTCGCAGTGGTTTCTCTGCAGCTCCCTGGGCTCCGCGGCCAGTAGTGCAGCCGTGGAGCCGCG 100 - Human CPE (P1622)

[CCCCGGCTGAATGAGTGTACGCCGGGAAGGTGAGGCGAGTAGAGGCTGGTGGCGAACTTGCCGCCCC Human CPE 5' H1622]

AGACGAGAGGAGGCTGGTGTGAGCTCGCAATC - Rat CPE

GCTTTGCCGCTCTCTCTGGGTGGCCAGTGGCGGGGTGACACTCATTACGCCGGGAAGGTGAGGCGAGTAGAGGCTGGTGGCGAACTTGCCGCCCC 200 - Human CPE

CA-----CCCGGCTCGGGGAGACAAAAGAGCC-----GCGGTACAGCTCGCGGGGACGCG..... - Rat CPE

CAGCAGCGCCGGCGGGCTAAGCCAGGCTCGGGGAGACAAAAGAGCCCGCGGTAGGAAGGCACGGCGCGGGCGGAGCGCAGCGATGGCCGGG 300 - Human CPE

m a g r

.C..A..AC.G.G.....G....T.CC..G....TG...G.....GT.A.CG..T.....A....T....

GAGGGGCGAGCGCTGCTGGCTCTGTGCGGGGCACTGGCTGCTGCGGGTGGCTCTGGGCGCCGAAGCCAGGAGCCGGGGCGCCCGCGGGCAT 400

g g s a l l a l c g a l a a c g w l l g a e a q e p g a p a a g m

.C.....G.....T.....G.....G..A.....

GAGGGGCGCGGGGCTGACGAAGAGGACGCGATCTCTTCGAGTACCAAGCTACCCGAGCTGCGGAGGCGCTGCTGCTCGGTGGCTGAGTGC 500

r r r r r l q q e d g i s f e y h r y p e l r e a l v s v w l q c

.....C.....T.....C..G.....C....

ACCGCATCAGCAGGATTACAGGTTGGGCGCAGCTTCGAGGGCGGGAGCTCCTGGTTCAGAGCTGTCCGACAACCTGGCGTCCATGAGCTGGTG 600

t a i s r l y t v g r s f e g r e l l v i e l s d n p g v h e p g e

.A.....G.....C.....T.....G.....G..G.....G..T.....G..T.....G..

AGCCTGAATTTAAATACATTGGAATATGCATGGGAATGAGGCTGTGGAGGAGAACTGCTCATTTCTTGGCCAGTACCTATGCAACGAATACAGAA 700

p e f k y i g n m h g n e a v g r e l l i f l a q y l c n e y q k

A.....T.....C..A..A..C..T.....C...T.....C.....A.....A.....C...

GGGGAACGAGACAATTGTCAACCTGATCCACAGTACCCGATTACATCATGCTTCCCTGAACCCAGATGGCTTTGAGAAGGCAGCGTCTCAGCCTGGT 800

g n e t i v n l i h s t r i h i m p s l n p d g f e k a a s q p g

..G..G.....C.....C..C.....G.....C.....T.....C.....T..T.....

GAACCTCAAGGACTGGTTTGGGTGGAAGCAATGCCAGGGAATAGATCTGAACCGGAATTTCCAGACCTGGATAGGATAGTGTACGTGAATGAGAAAG 900

e l k d w f v g r s n a q g i d l n r n f p d l d r i v y v n e k e

...C....C..C..C..C..C..G..C.....C.....TT.....C..C.....C.....C..

AAGGTGGTCCAAATAATCATCTGTGAAAAATATGAAGAAATGTGGATCAAAACACAAAGCTTGCTCTGAGACCAAGGCTGTCTATTGGATTAT 1000

g g p n n h l l k n m k k i v d q n t k l a p e t k a v i h w i m

...C..C..A.....C.....C..G..C.....C.....T.....C..G.....C.....T

GGATATTCCTTTTGTGCTTTCTGCCAATCTCCATGGAGGAGACCTTGCGGCAATTATCCATATGATGAGACGCGGAGTGGTAGTGTACCAATACAGC 1100

d i p f v l s a n l h g g d l v a n y p y d e t r s g s a h e y s

...G...T.....A.....T.....C.....A..T.....A..T.....C.....

TCCTCCCGAGTACGCCATTTTCCAAAGCTTGGCCCGGCATCTCTCTTCAACCGGCCATGTCTGACCCCAATCGGCCACCATGTGCAAGAATG 1200

s s p d d a i f q s l a r a y s s f n p a m s d p n r p p c r k n d

.....T.....A.....T.....A.....C..C..T..A.....G.....C..

ATGATGACAGCAGCTTTGTAGATGGAACCAACCGTGGTGGTGTGGTACAGCGTACCTGGAGGATGCAAGACTTCAATTACCTTAGCAGCACTGTT 1300

d d s s f v d g t t n g g a w y s v p g g m q d f n y l s s n c f

.....T.....C..A..G.....A.....C..A.....G...

TGAGTACCCGTGGAGCTTAGCTGTGAGAAGTTCCACCTGAAGAGACTCTGAAGACCTACTGGGAGGATAACAAAACTCCCTCATTAGCTACCTTGAG 1400

e l t v e l s c e k f p p e e t l k t y w e d n k n s l i s y l e

.....T.....G.....T.....T..G.....C..C..A.....T.....T..G.....T.....C..C..

CAGATACACCGAGGAGTTAAGGATTGTCCGAGACCTCAAGGTAACCAATTCGGAATGCCACCATCTCCGTGGAAGGAATAGACCAGATGTACAT 1500

q i h r g v k g f v r d l q g n p i a n a t i s v e g i d h d v t s

.G..T.....G.....C.....GCG.....C.....C.....C.....

CCGCAAGGATGGTGATTCTGAGATTGCTTATACCTGGAACACTATAAACTTACAGCTCAGCTCCAGGCTATCTGCAATAACAAAGAAAGTGGCAGT 1600

a k d g d y w r l l i p g n y k l t a s a p g y l a i t k k v a v

.....T.....T.....G..C.....G.....T..C..A.....G..G.....G.....

TCCTTACAGCCCTGCTGCTGGGTGATTGTGAAGTGGAGTCAATTTCTGAAAGGAAGAGAGGAGGAAGGAATGATGGAATGGTGGAAATGATG 1700

p y s p a a g v d f e l e s f s e r k e e e k e e l m e w w k m m

.....G.....G.....GAAAGGCTCCTAATATGGCTTT-AATATATCCATAGACTATAGTAAGATGCAACGTGGCTCTTTATTTTAGGTTGTG 1800

TCAGAACTTTAAATTTTAA-AAAGGCTCTAGTATAGCTGCTTTAAATCTATCTATATAATGATAGTATGTAATGGTCTTTTT-TTAGATTGTG

s e t l n f .

CAGTTAATACTTAACATCGTTATTTTTC-ATCATTTAAGTAGTAGTTGCTAATCACTTAATACA-----CTTGACAGAACCTGTAGTGTCTGGAT

CAGTTAATACTTAACATCGTTATTTTTCATCATTTAAATATTAATCACTTTCTCTAAATAAATAGCCTCTTAGGTAATAATAAGAACTTGATA 1900

CGACTTCATTCTTTTCATGAACATTCGCTTT-----

TATTTTCATTCTCTTATATAGTATTCATTTTCTACCTATATTACACAAAAAGTATAGAAAAGATTAAAGTAATTTTGCCATCTAGGCTTAAATGCAAT 2000

-----AAAATCTAGGAAGATCTTTAACCTAGTTGGTGACAAATGTCACATGACA

ATTCTGGTATTATTACAAATGCAATTTTTTGGTAATTCAGCTTTCAAAAATTAGTGAAGTCTTTTACTGTAAATTGGTGACAAATGTCACATTAATG 2100

GATGCCA-TGAAAAGTCAACCGATGAGCTTGGATCTGTGAGCACTGTGTGTATGGCGAAGATCAGATAGTCCATATAAGTTGTCTTGTCTTGTGCTG

AATGCTATTGAAAAGGTTAACAGATACAGCTCGGAGTTGTGAGCACTTACTGCA-----AGACTTAAATAGTTCAGTATAAATGTGCTTTTTTCTTGTGCTG 2200

ATCCACTGTATAAGCATGATCTGTTAATGCGCTTCTGAGGGAAGGAAACGTCAGTCTTTTGCAA-AGGGGCTCTGAGCAGAAGGAAACCTAGTCTTGTG

ACTAATCT-ATAAGCATGATCTGTTAATGCATTTTTGATGGGAAGAAAGGTACATG-TTTCAAAGAGGTTTTATGAAAAGAAATAAAATGACTTCTTGTG 2300

GTGACTTTTGAAGAACAGAAGT-TATTAGTCAGCTGTAGTGCAACTTCAAGTTTGGGTTTTCTTCAATTGAGATTGGCCAGAAATTCATT

TTGACATATAGGAGCAATACATATTATATAGTCCGTTAACACTACTTAAAGTTTGGGTTTTCTTGTGTTAGATTGGCCAGAAATTCATT 2400

CTAATGAATAAAGGATTA Rat CPE

CTAATGAATAAAGGTTAAAAAAAATCCCCAGTAAAAAAA 2443 Human CPE

Fig. 2. Size of rat and human CPE mRNAs

Total RNA (25 µg) from rat brain (lane 1) and from human brain (lane 2) were electrophoresed on a 0.8% denaturing agarose gel. Purification and Northern-blot analysis of RNA were as previously described (Hall *et al.*, 1987) with a ³²P-labelled cDNA probe that included the 3' coding region of the human mRNA. The size of the mRNA (kb) was estimated from rRNA markers (18 S and 28 S).

Processing and membrane association of products of the translation of human CPE mRNA *in vitro*

As a first step in investigating the biogenesis of membrane and soluble forms of native brain CPE, we studied the synthesis and processing of human CPE *in vitro*. Translation in rabbit reticulocyte lysate *in vitro* of mRNA transcripts synthesized from CPE cDNAs by using the T7 promoter in Bluescript (see the Materials and methods section) resulted in the synthesis of a 54 kDa protein (Fig. 6, lanes 1 and 4). This is consistent with efficient initiation at the conserved ATG initiation consensus at nucleotide-290 in human CPE mRNA. Smaller translation products present in lower amounts may have resulted from the presence of some truncated mRNAs, internal initiation or partial proteolysis; none of these was inserted into microsomal membranes (see below).

In the presence of dog pancreas microsomal membranes two

processed forms of CPE (52 kDa and 56 kDa) were synthesized, in addition to variable amounts of a 54 kDa product (Fig. 6, lanes 2 and 5). All three molecular-mass species were inserted into microsomal membranes (Fig. 6, lanes 3 and 7, and Fig. 7, lanes 2, 3 and 6). In contrast, endogenous reticulocyte-lysate proteins and smaller-molecular-mass components remained in the supernatant fraction, as did 54 kDa CPE (presumably unprocessed) translation product (Fig. 6, lane 6). The possibility that different degrees of protein glycosylation could produce molecular-mass heterogeneity of partially processed forms was investigated. The carbohydrate was removed from the polypeptide chains by incubation with endoglycosidase H. Both major processed forms decreased in size by approx. 3 kDa (Fig. 6, lanes 8 and 9), indicating that the 52 kDa and 56 kDa processed forms were glycosylated to the same extent. Thus the 52 kDa processed CPE is not an incompletely glycosylated precursor to the 56 kDa form, although the variable middle processed band (54 kDa) may be an intermediate form since only two protein bands were present after endoglycosidase treatment (Fig. 6, lane 8). There was efficient signal-peptide cleavage in this system; in control experiments preprolactin was completely converted into the smaller pro form. Both deglycosylated proteins were smaller than the unprocessed primary product, suggesting that both lacked the signal peptide. The smaller (52 kDa) form appears to have undergone an additional processing step.

CPE precursors synthesized in reticulocyte lysate could be completely removed from microsomal membranes at pH 11.5 (Fig. 7, lane 4) in the same manner as native membrane CPE (cf. Fig. 5, lane 8). In hypo-osmotic buffer at pH 7.5, three processed forms of CPE were membrane-associated and a 54 kDa (unprocessed) form remained in the supernatant fraction (Fig. 7, lanes 1 and 2; cf. Fig. 6, lanes 6 and 7). Luminal proteins can be released by hypo-osmotic lysis; in control experiments prolactin (a secreted protein) was released from microsomal fraction under these conditions (results not shown). All three CPE forms remained in the membrane fraction after treatment at pH 9.1 (Fig. 7, lane 3). [By comparison, the endoplasmic-reticulum enzyme protein disulphide-isomerase is released at slightly alkaline pH (> 9.0; Paver *et al.*, 1989).] When purified microsomal membranes containing processed CPE were treated at pH 10.4, although there was significant release of CPE, a substantial proportion remained membrane-associated (Fig. 7, lanes 5 and 6). These results showed that all processing intermediates of CPE synthesized *in vitro* became co-translationally associated with the endoplasmic reticulum, and like native membrane CPE behaved as peripheral membrane proteins. The size of native brain membrane CPE (50 kDa) was less than the smallest processing intermediate synthesized *in vitro*.

Fig. 1. Nucleotide sequence of human and rat CPE cDNA and predicted human protein sequence

The nucleotide sequence of human CPE cDNA (P1622) and predicted amino acid sequence are shown. The 5' and 3' non-coding regions of the rat CPE cDNA (M16225E; Hall *et al.*, 1987) are shown in full above the human sequence; highly conserved non-coding regions are boxed. Over the putative coding region only the nucleotide residues that differ from the human sequence are shown; ., Identical nucleotides; —, gaps introduced for optimal alignment. The polyadenylation signal is underlined. The rat CPE sequence derived from M16225E cDNA differs from published rat sequences (Rodriguez *et al.*, 1989; Fricker *et al.*, 1989) in two base changes: T/C and C/A, which result in amino acid substitutions of V to A at position 415 and S to Y at position 453 (cf. Fig. 3). (For clarity in the Figure, lower-case letters are used for the one-letter code for amino acid residues.) Our rat sequence agrees with Rodriguez *et al.* (1989) but not Fricker *et al.* (1989) in encoding S (residue 82, Fig. 3) and not T. There are also two nucleotide additions in the published 3' non-coding section (Fricker *et al.*, 1989) and the sequence extends further at both ends.

DISCUSSION

There is considerable species variation in CPE mRNA size from 2.2 kb in the rat to 3.3 kb for the major bovine species (Fricker *et al.*, 1986). The species difference in size between human CPE and rat CPE is attributable to extra 5'- and 3'-sequences in the human sequence, with a high degree of sequence conservation over the region encoding a 476-amino acid-residue open reading frame in all three species.

The sequence relationship of human, rat and bovine CPEs, the efficient synthesis *in vitro* of a 54 kDa primary human CPE translation product (which is subsequently processed) and the conservation of an initiation consensus (Kozak, 1987) at the first methionine codon suggest that this represents the initiation site *in vivo*. The size of the primary translation product of rat CPE mRNA has been shown by hybrid-selected translation of brain mRNA to be 54 kDa (Hall *et al.*, 1987). However, in both rat and

Fig. 3. Amino acid sequence comparison of human, rat and bovine CPEs

Identical amino acids in rat and bovine sequences are indicated as a dot (·) above the predicted human sequence; difference are shown with non-conservative changes circled. The complete bovine sequence is from Fricker *et al.* (1986, 1989). The putative signal peptide is underlined and potential signal-peptide cleavage site arrowed, S. The *N*-terminal sequence of the mature bovine CPE (Fricker *et al.*, 1986) is marked by a horizontal arrow. The two potential *N*-linked glycosylation site are marked (◆). The predicted location of the 17 kDa CNBr-cleavage fragment is marked by upward arrows (†).

human (P1622) cDNAs an open reading frame extends from the extreme 5'-end. In the case of the human cDNA the sequence before the first methionine residue encodes an additional 96 amino acid residues. It remains a possibility that a larger human primary translation product could be derived if this cDNA were an incomplete copy of a longer mRNA with an initiator methionine residue further upstream. Hook & Affolter (1988) have reported a 75 kDa CPE primary translation product synthesized from human ileal carcinoma mRNA. There may also be 5'-end heterogeneity in CPE mRNA resulting from alternative exon usage, as is the case for other brain mRNAs (e.g. tyrosine hydroxylase; Le Bourdelles *et al.*, 1988). A second human CPE cDNA (H1622) showed sequence divergence at its extreme 5'-terminus that could result from either an alternatively spliced mRNA or transcriptional errors during cDNA synthesis and cloning.

Earlier studies on rat brain show CPE mRNA to be restricted to membrane-bound polyribosomes (Hall *et al.*, 1987). Membrane and soluble forms of CPE are thus both synthesized in the rough endoplasmic reticulum and apparently derive from a single mRNA species by post-translational processing. Another enzyme of neurosecretory granules, dopamine β -hydroxylase, is also found in membrane and soluble forms that differ in molecular mass and derive from a common precursor (Sabban *et al.*, 1987). Both membrane and soluble forms of CPE have the same *N*-terminus, and the smaller soluble form has been postulated to be derived by cleavage of a *C*-terminal domain involved in mem-

brane anchoring (Fricker *et al.*, 1986). However, the mechanism of membrane interaction is unclear, since this region of the protein is predicted to be an α -helical hydrophilic structure (in contrast with the hydrophobic *C*-terminal of CPM; cf. Fig. 4b). The hydropathy profile of CPE reveals a number of hydrophobic regions other than the signal sequence. None is of sufficient length to constitute a conventional membrane-spanning α -helix, but may provide a means of membrane interaction. Our studies on the membrane CPE showing its resistance to moderate alkali extraction and release from membranes at pH 11.5 indicate that it is a peripheral membrane protein, consistent with these sequence predictions. The human CPE translation product can be processed *in vitro* by cleavage, presumably of the *N*-terminal signal sequence, insertion into microsomal membranes and glycosylation. The generation of two major processed forms that were equally glycosylated and membrane-associated may represent the first steps in the generation of membrane and soluble forms. Nevertheless, our studies on the interaction of CPE with the membrane do not allow us to determine how the native soluble CPE form is generated. Later processing events may play a role in generating the native soluble CPE. Modification of the carbohydrate moiety in the Golgi apparatus and enzyme activation by cleavage of the 16-amino acid-residue putative 'pro' sequence in secretory granules, the final destination of CPE (Hook & Loh, 1984), could account for the smaller size of both native membrane and soluble CPE than their processing intermediates synthesized *in vitro*.

Fig. 4. Comparison of human carboxypeptidase sequences: (a) amino acid sequences of CPE, CPN and CPM; (b) hydropathy profiles of CPE, CPN and CPM

(a) Amino acid sequences of CPE, CPN and CPM. Human CPE (476 amino acid residues including pre-pro sequence) aligned with human CPN (438 amino acid residues; Gebhard *et al.*, 1989) by using the DNASTAR 'align' program and with human CPM (425 amino acid residues; Tan *et al.*, 1989). Numbers refer to CPE sequence. Residues implicated mechanistically (i.e. Zn^{2+} -binding site and substrate-binding active site) are indicated (▲). (b) Hydropathy profiles of CPE, CPN and CPM. Hydropathy plots were derived by using the algorithm of Kyte & Doolittle (1982) with a window of nine residues. Numbering is according to that of CPE and gaps have been introduced into the profile according to the alignment in (a). Most hydrophobic regions (+ values) are conserved between the three proteins; however, CPM has a hydrophobic C-terminal domain that may function as a membrane anchor. Regions that differ significantly in one protein compared with the other two are marked above with a bar.

Comparison of the three human carboxypeptidases (E, M and N) that have a common enzymic activity reveals conservation of domains, some of which are thought to contribute to the metal-ion-binding and catalytic sites, on the basis of comparative analysis with carboxypeptidases A and B. Dispersed between these conserved sections there are striking sequence differences,

which may represent regions of the protein responsible for specific enzymic characteristics, e.g. pH optima; alternatively, they may determine the final cellular destination of the different carboxypeptidases. The occurrence of specific sequences determining intracellular compartmentation of proteins is well established (Pfeffer & Rothman, 1987). Further studies on the

Fig. 5. Membrane association of mature CPE

Human or rat brain membrane fractions (see the Materials and methods section) were treated with 0.1 M-sodium carbonate buffers at various pH values; solubilized (S) and membrane fractions (P) were then separated by centrifugation through a sucrose cushion. After electrophoresis on SDS/polyacrylamide gels, CPE was detected by Western-blotting with human CPE antiserum. Lane 1, human brain CPE from purified membranes; lane 2, native rat brain soluble CPE (see the Materials and methods section and the Results section); lanes 3–8, after hypo-osmotic lysis and one cycle of freeze–thaw, rat brain membranes were resuspended (10 mg of protein/ml) in buffer at the appropriate pH for 30 min at 4 °C.

divergent sequences of the human carboxypeptidases may provide information on the nature of these targeting signals, particularly since CPE is found in secretory granules, CPN in serum and CPM ubiquitously in cell membranes.

Fig. 7. pH-dependent release from microsomal membranes of CPE processed *in vitro*

Reticulocyte lysates containing [³⁵S]methionine-labelled CPE synthesized in the presence of microsomal membranes were treated at pH 7.5, pH 9.1 or pH 11.5, as indicated above lanes 1–4, and membranes (P) were pelleted by sedimentation through a sucrose cushion. The supernatant (S) from mixtures treated at pH 7.5 is shown in lane 1. In lanes 5 and 6, microsomal membranes containing bound CPE were first purified by sedimentation through sucrose before treatment at pH 10.4, then re-purified by centrifugation through a sucrose cushion to assess the proportions of CPE solubilized (lane 5) or remaining membrane-bound (lane 6). Samples containing CPE translation products were analysed by SDS/PAGE and fluorography.

Fig. 6. Translation and processing of CPE *in vitro*

Human CPE mRNA transcripts of P1622 and H1622 cDNAs (see the Materials and methods section) were translated in rabbit reticulocyte lysate in the presence (+) or in the absence (–) of dog pancreatic microsomal membranes (PM) (lanes 1–5). Lanes 1 and 4, unprocessed translation products of H1622 and P1622 mRNA transcripts respectively; lanes 2 and 5, processed translation products of H1622 and P1622 mRNA respectively; lane 3, purified microsomal membranes containing processed translation products of H1622 mRNA. Lanes 6–9 show CPE translation products of P1622 mRNA synthesized in the presence of dog pancreas microsomal membranes; lane 7, pelleted microsomal membranes (P) purified by sedimentation through sucrose containing processed CPE; lane 6, supernatant fraction (S) containing CPE not incorporated into microsomal membranes; lanes 8 and 9, purified microsomal membranes containing processed CPE incubated in the presence (+, lane 8) or in the absence (–, lane 9) of endoglycosidase H (Endo H), as described in the Materials and methods section. [³⁵S]Methionine-labelled translation products were analysed by SDS/PAGE and fluorography.

We thank The Brain Research Trust for financial support of work undertaken at the Institute of Neurology.

REFERENCES

- Fricker, L. D. (1988) *Annu. Rev. Physiol.* **50**, 309–321
- Fricker, L. D. & Snyder, S. H. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 3886–3890
- Fricker, L. D., Evans, C. J., Esch, F. S. & Herbert, E. (1986) *Nature (London)* **323**, 461–464
- Fricker, L. D., Adelman, J. P., Douglass, J., Thompson, R. C., von Strandmann, R. P. & Hutton, J. (1989) *Mol. Endocrinol.* **3**, 666–673
- Fujiki, Y., Hubbard, A. L., Fouler, S. & Lazarow, P. B. (1982) *J. Cell Biol.* **93**, 97–102
- Gebhard, W., Schube, M. & Eulitz, M. (1989) *Eur. J. Biochem.* **178**, 603–607
- Hall, C., Mahadevan, L., Whatley, S., Biswas, G. & Lim, L. (1984) *Biochem. J.* **219**, 751–761
- Hall, C., Lowndes, C. M., Leung, T. K. C., Cooper, D. N., Goate, A. M. & Lim, L. (1987) *Biochem. J.* **244**, 359–366
- Hook, V. Y. H. & Affolter, H.-U. (1988) *FEBS Lett.* **238**, 338–342
- Hook, V. Y. H. & Loh, Y. P. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2776–2780
- Kozak, M. (1987) *Nucleic Acids Res.* **15**, 8125–8148
- Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- Laslop, A., Fischer-Colbrie, R., Hook, V., Obendorf, D. & Winkler, H. (1986) *Neurosci. Lett.* **72**, 300–304
- Le Bourdelles, B., Boularand, S., Boni, C., Horellou, P., Dumas, S., Grima, B. & Mallet, J. (1988) *J. Neurochem.* **50**, 988–991
- Lowndes, C. M. (1988) Ph.D. Thesis, University of London
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor
- Nathans, J., Thomas, D. & Hogness, D. S. (1986) *Science* **232**, 193–202
- Paver, J. L., Hawkins, H. C. & Freedman, R. B. (1989) *Biochem. J.* **257**, 657–663
- Pfeffer, S. R. & Rothman, J. E. (1987) *Annu. Rev. Biochem.* **56**, 829–852
- Plummer, T. H., Jr. & Hurwitz, M. Y. (1978) *J. Biol. Chem.* **253**, 3907–3912
- Rees, D. C., Lewis, M. & Lipscomb, W. N. (1983) *J. Mol. Biol.* **168**, 367–387
- Rodriguez, C., Brayton, K. A., Brownstein, M. & Dixon, J. E. (1989) *J. Biol. Chem.* **264**, 5988–5995
- Rüther, U. & Müller-Hill, B. (1983) *EMBO J.* **2**, 1791–1794
- Sabban, E. L., Kuhn, L. J. & Levin, E. (1987) *J. Neurosci.* **7**, 192–200
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Schmid, M. F. & Herriott, J. R. (1976) *J. Mol. Biol.* **103**, 175–190
- Strittmatter, S. M., Lynh, D. R. & Snyder, S. H. (1984) *J. Biol. Chem.* **259**, 11812–11817
- Supattapone, S., Fricker, L. D. & Snyder, S. H. (1984) *J. Neurochem.* **42**, 1017–1023
- Tan, F., Chan, S. J., Steiner, D. F., Schilling, J. W. & Skidgel, R. A. (1989) *J. Biol. Chem.* **264**, 13165–13170
- Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4350–4354
- van Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4690

Received 15 September 1989/3 November 1989; accepted 17 November 1989