

Amino acid sequence of rabbit kidney neutral endopeptidase 24.11 (enkephalinase) deduced from a complementary DNA

Alain Devault, Claude Lazure¹, Christiane Nault, Hervé Le Moual, Nabil G.Seidah¹, Michel Chrétien¹, Philippe Kahn², John Powell², Jacques Mallet², Ann Beaumont³, Bernard P.Roques³, Philippe Crine and Guy Boileau

Département de Biochimie, Université de Montréal, Montréal, Canada H3C 3J7, ¹Institut de Recherches Cliniques de Montréal, Montréal, Canada, ²Laboratoire de Neurologie Cellulaire et Moléculaire, CNRS 91190 Gif-sur-Yvette, and ³Département de Chimie Organique, INSERM U 266 et CNRS UA 498, U.E.R. des Sciences Pharmaceutiques et Biologiques, 75006 Paris, France

Communicated by J.P.Changeux

Neutral endopeptidase (EC 3.4.24.11) is a major constituent of kidney brush border membranes. It is also present in the brain where it has been shown to be involved in the inactivation of opioid peptides, methionine- and leucine-enkephalins. For this reason this enzyme is often called 'enkephalinase'. In order to characterize the primary structure of the enzyme, oligonucleotide probes were designed from partial amino acid sequences and used to isolate clones from kidney cDNA libraries. Sequencing of the cDNA inserts revealed the complete primary structure of the enzyme. Neutral endopeptidase consists of 750 amino acids. It contains a short N-terminal cytoplasmic domain (27 amino acids), a single membrane-spanning segment (23 amino acids) and an extracellular domain that comprises most of the protein mass. The comparison of the primary structure of neutral endopeptidase with that of thermolysin, a bacterial Zn-metallopeptidase, indicates that most of the amino acid residues involved in Zn coordination and catalytic activity in thermolysin are found within highly homologous sequences in neutral endopeptidase.

Key words: microvillar hydrolases/Zn-endopeptidase/enkephalinase/neuropeptide degradation

Introduction

Neutral endopeptidase (EC3.4.24.11) (NEP) is a membrane-bound Zn-metalloendopeptidase located in the plasma membrane of many tissues (Kenny, 1986). In mammalian brain, the enzyme has been shown to be involved in the inactivation of the opioid peptides, methionine- and leucine-enkephalins (Malfroy *et al.*, 1978; Almenoff *et al.*, 1981), and is therefore, often called 'enkephalinase'. The biological relevance of NEP as an enkephalin-degrading enzyme is supported by its distribution in rat brain which overlaps that of opioid receptors (Waksman *et al.*, 1986a), its neuronal localization (Matsas *et al.*, 1986; Waksman *et al.*, 1986b) and by the naloxone-reversible analgesic responses induced by inhibitors such as thiorphan (Roques *et al.*, 1980) and retrothiorphan (Roques *et al.*, 1983). Inhibitors of enkephalinase represent a new class of potential analgesic drugs (Chipkin, 1986; Roques and Fourmié-Zaluski, 1986). The design of these compounds was based on a generalized active site model for Zn-metallopeptidases already used for the develop-

ment of the angiotensin-converting enzyme inhibitor, captopril (Ondetti *et al.*, 1977). In order to design highly potent and orally active inhibitors, more precise information on the active site of NEP is required. For this purpose, we have elucidated the primary structure of the enzyme by cloning and sequencing DNA complementary to the mRNA coding for rabbit NEP.

Results

Purification and partial amino acid sequence of NEP

NEP was purified from rabbit kidney by immunoaffinity chromatography using a monoclonal antibody (Crine *et al.*, 1985) (Figure 1a, lane 1). The amino acid sequence determination of the purified native enzyme was attempted twice using the liquid-phase sequenator. On each occasion, analysis of the data proved extremely difficult because of exceedingly low initial yield (estimated at 10–20% based on the weighted amount loaded on the sequenator) and the presence within the sequence of a number of unstable phenylthiohydantoin PTH (-amino) acid derivatives (such as Ser and Thr) recovered in low yield. Furthermore, a rapid decrease in repetitive yield prevented the interpretation of the sequence data past 17 cycles. Nevertheless, in each run a single sequence corresponding to the one shown in Figure 1b was obtained. It is noteworthy that this putative NH₂-terminal sequence does not correspond to the one recently proposed by Fulcher *et al.* (1986) for porcine NEP. In fact, their sequence corresponds to a truncated form of NEP lacking the first 13 or 14 residues.

Considering the difficulties with the sequencing of the native enzyme, fragments from cyanogen bromide treatment were fractionated by SDS–polyacrylamide gel electrophoresis (SDS–PAGE) and isolated by electroelution (Figure 1a, lane 2). Four fractions were obtained after electroelution from polyacrylamide gel and directly submitted to sequence analysis on a gas-phase sequenator. As shown in Figure 1b, it was possible to obtain four distinct amino acid sequences which proved to be extremely helpful not only for the synthesis of the oligonucleotide probes but also to confirm in an independent manner the sequence deduced from the nucleotide analysis. Two unique single-stranded DNA probes of 63 (probe A) and 57 (probe B) nucleotides coding for amino acid sequences of CNBr-2 and CNBr-4 respectively were designed according to codon usage frequencies (Ikemura, 1985; Lathe, 1985) (Figure 1b).

Isolation and DNA sequencing of two overlapping cDNA clones for NEP

An oligo(dT)-primed λgt10 cDNA library was generated from rabbit kidney poly(A)⁺ RNA and 1 × 10⁵ recombinant phages were screened by plaque hybridization with both probes A and B. Of 10 positive recombinant phages obtained with probe A, clone λENK7 had the longest insert (2.2 kb). This insert was sequenced by the dideoxy chain termination procedure (Sanger *et al.*, 1977) (Figure 2). An open reading frame coding for 558 amino acids was found starting with the first nucleotide and end-

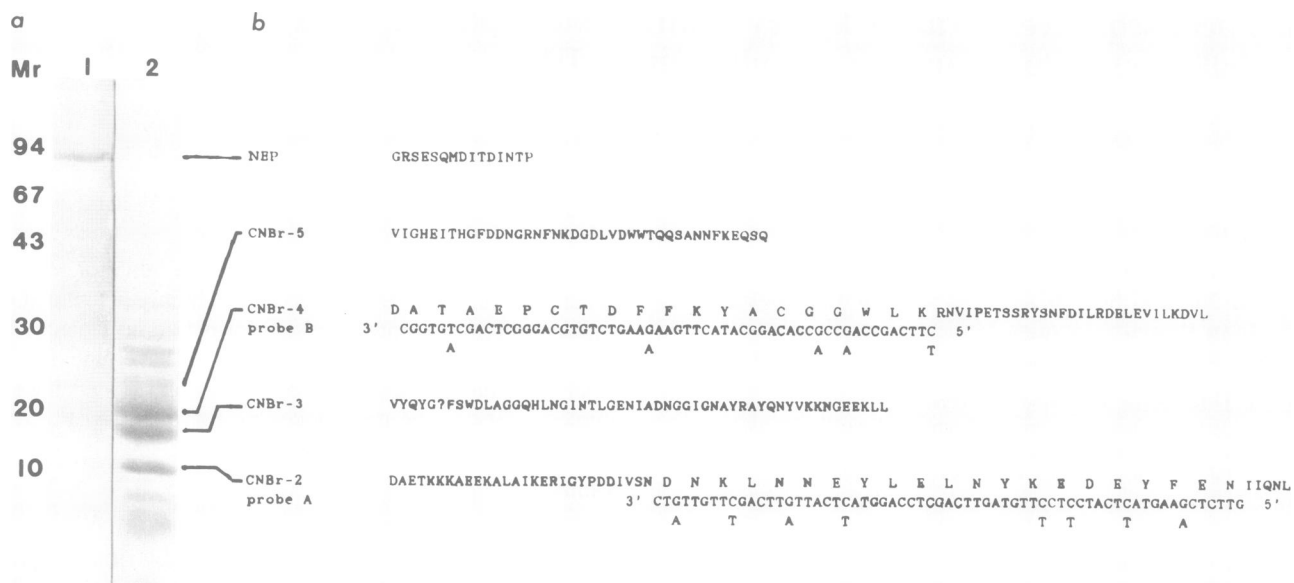


Fig. 1. Determination of NEP partial amino acid sequence. **(a)** Coomassie blue-stained SDS-polyacrylamide gel of purified NEP (**lane 1**) and NEP CNBr fragments (**lane 2**). The proteins used as standards are phosphorylase b, bovine serum albumin, ovalbumin, carbonic anhydrase, lactalbumin and trypsin inhibitor. (*M_r* refers to the relative mol. wt $\times 10^{-3}$). **(b)** NH_2 -terminal sequence of the intact NEP and some electroeluted CNBr fragments. The oligonucleotide probes A and B are complementary to the mRNA. The letters below the probe sequences represent those nucleotides that were found different from the cDNA sequence. The question mark at position 6 in CNBr-3 fragment indicates the position of a cycle where no PTH derivative could be detected.

ing 1674 nucleotides downstream. This open reading frame coded for peptides CNBr-2, CNBr-3 and CNBr-5.

Moreover, part of λ ENK7 coding sequence is found in the insert of a clone selected from a λ gt11 library for its ability to promote the synthesis of a fusion protein that binds immunoglobulins from a polyclonal antibody specific for rabbit NEP (not shown). The remaining 483 nucleotides are presumably part of the 3'-untranslated region of the NEP mRNA. However this sequence lacks both the canonical polyadenylation sequence AATAAA (Proudfoot and Brownlee, 1976) and the poly(A) tract. Undermethylation of the endogenous *Eco*RI site in the cDNA during library preparation most probably explains the absence of these regions and also most of the difference between the length of the cloned cDNA and that of the mRNA as measured by Northern blot hybridization (see below and Figure 3).

The lack of coding regions for CNBr-4 (probe B) and the N-terminal peptide (Figure 1b) of the protein indicates that clone λ ENK7 does not contain the total mRNA sequence. As further screening of the λ gt10 library with probe B did not yield any positive clone, we generated a new cDNA library in pUC19 using a synthetic oligonucleotide corresponding to a region proximal to the 5' end of clone λ ENK7 as a primer (Figure 2). Screening of the new library by colony hybridization with probe B allowed the isolation of four positive colonies. The sequence of the cDNA insert of one of the positive colonies (pENK8) showed the expected 117-nucleotide overlap with the 5' end of clone λ ENK7 (Figure 2). Furthermore, it included the sequences of peptide CNBr-4 and of the N terminus of the native protein.

Clones λ ENK7 and pENK8 contain overlapping DNA inserts which together spanned a stretch of 2.8 kb of DNA (Figure 2). This is shorter than the 3.6 kb determined by Northern blot hybridization for NEP mRNA (Figure 3). However, complete sequence analysis of the clones demonstrated an open reading frame of 2253 nucleotides starting at the first ATG codon encountered from the 5' end. We believe that this open reading frame codes for the total NEP primary structure for two reasons.

First, the N-terminal sequence of the protein determined by Edman degradation is identical to the protein sequence deduced from the cDNA (less the initiator methionine) and second, an in-phase TAG stop codon is found six nucleotides upstream from the initiator ATG.

Primary structure of NEP

The open reading frame encodes a protein of 750 amino acids (excluding the initiator methionine), with a calculated mol. wt of 85 452. This is close to the value of 94 000 determined by SDS-PAGE or that of 85 000 estimated after removing N-linked oligosaccharides with peptide: N-glycosidase (N-glycanase: Genzyme, Boston) (results not shown). The predicted protein primary structure contains five asparagine residues that are part of the consensus sequence Asn-X-Ser/Thr for N-glycosylation sites. Glycosylation of some of these asparagine residues could account for the difference between the mol. wt calculated from the amino acid composition and that estimated by SDS-PAGE of the native enzyme.

The failure to detect an amino acid residue in CNBr-3 at the position corresponding to Asn 628 of the protein sequence (Figure 1b) suggests that at least this residue is glycosylated. The primary structure of NEP does not contain serine- or threonine-rich domains that are believed to correspond to potential O-glycosylated sites such as those found in glycoporphin (Tomita *et al.*, 1978) the low density lipoprotein receptor (Yamamoto *et al.*, 1984) and sucrase-isomaltase (Hunziker *et al.*, 1986). Therefore it is not possible to infer this type of post-translational modification on the basis of the primary structure alone.

There are 12 cysteine residues in NEP, four of which are clustered in a 32-amino acid segment of the protein immediately following the putative transmembrane domain (see below). Such clustering of cysteine residues close to the anchoring point in the membrane has also been observed for sucrase-isomaltase and γ -glutamyl-transpeptidase, two other microvillar proteins (Hunziker *et al.*, 1986; Laperche *et al.*, 1986). It has been proposed

the C terminus of the molecule, protrudes on the extracellular luminal side.

Anchoring via the N-terminal region as suggested here for NEP is seen in other microvillar enzymes, namely aminopeptidases M and A, dipeptidylpeptidase IV and maltase-glycoamylase (for a review, see Semenza, 1986). A model for the mechanism of membrane insertion of the nascent NEP in the rough endoplasmic reticulum should therefore involve a variation of the initial helical hairpin model of Halegoua and Inouye (1979) as recently proposed by Wickner and Lodish (1985). The principal features of the primary structure of rabbit kidney NEP are summarized in Figure 5.

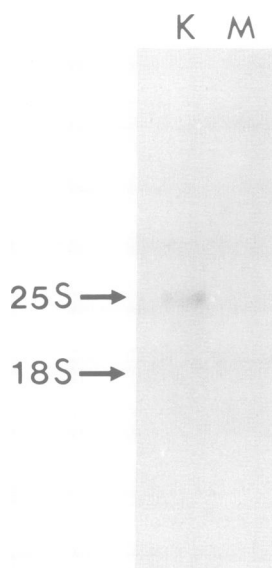


Fig. 3. Northern blot analysis of poly(A)⁺ RNA from rabbit kidney (lane K). Poly(A)⁺ RNA from skeletal muscle, which does not contain NEP, is shown as a control (lane M). Positions of 25S and 18S rRNA markers are indicated by arrows.

Discussion

There is very little overall homology between NEP and other Zn-metallopeptidases such as carboxypeptidase A, B and E (Fricker *et al.*, 1986) as well as thermolysin. However, most of the important amino acids present in the active site of thermolysin (Kester and Matthews, 1977) have been conserved in NEP. These include two of the Zn-coordinating residues His 585 and His 589 in NEP (which most probably correspond to His 142 and His 146 in thermolysin) and the essential amino acids involved in catalysis and binding (Glu 586 and His 639 in NEP versus Glu 143 and His 231 in thermolysin). In both enzymes, all of these amino acids are found within highly homologous sequences (Figure 6).

These results are in agreement with previous reports on the similar specificity of the two enzymes (Roques and Fournié-Zaluski, 1986; Pozsgay *et al.*, 1986; Hersh and Moribaza, 1986), and the presence of a critical His residue at the active site of NEP (Beaumont and Roques, 1986) in contrast to a tyrosine residue in carboxypeptidases (Fricker *et al.*, 1986; Quiocho and Lipscomb, 1971). The good correspondence between the active site sequences of thermolysin and NEP should allow the use of the known tertiary structure of thermolysin as a working model for the NEP active site. The 'docking' of inhibitors by computer graphics (Bush, 1984; Recanatini *et al.*, 1986), recently used in the design of ACE (Hangauer *et al.*, 1984) and renin inhibitors (Carlson *et al.*, 1985), could thus be applied to NEP. Clinical applications of NEP inhibitors would require knowledge of their effect on the transcription of the gene. Such information can now be easily obtained from *in situ* hybridization experiments using the cDNA which we have isolated.

Materials and methods

Purification and amino acid sequence determination of NEP and NEP fragments

NEP was purified from octyl glucoside-solubilized rabbit kidney cortex membranes by immunoaffinity chromatography essentially as described previously (Crine *et al.*, 1985; Aubry *et al.*, 1987). Purified NEP was cleaved with CNBr after reduction and carboxymethylation with iodoacetic acid (Glazer *et al.*, 1976) and the fragments were isolated by electrophoresis on a 10–20% SDS–polyacrylamide gel (Laemmli, 1970) and electroelution (Hunkapiller *et al.*, 1983).

The amino acid sequence determinations of the native NEP and of the reduced

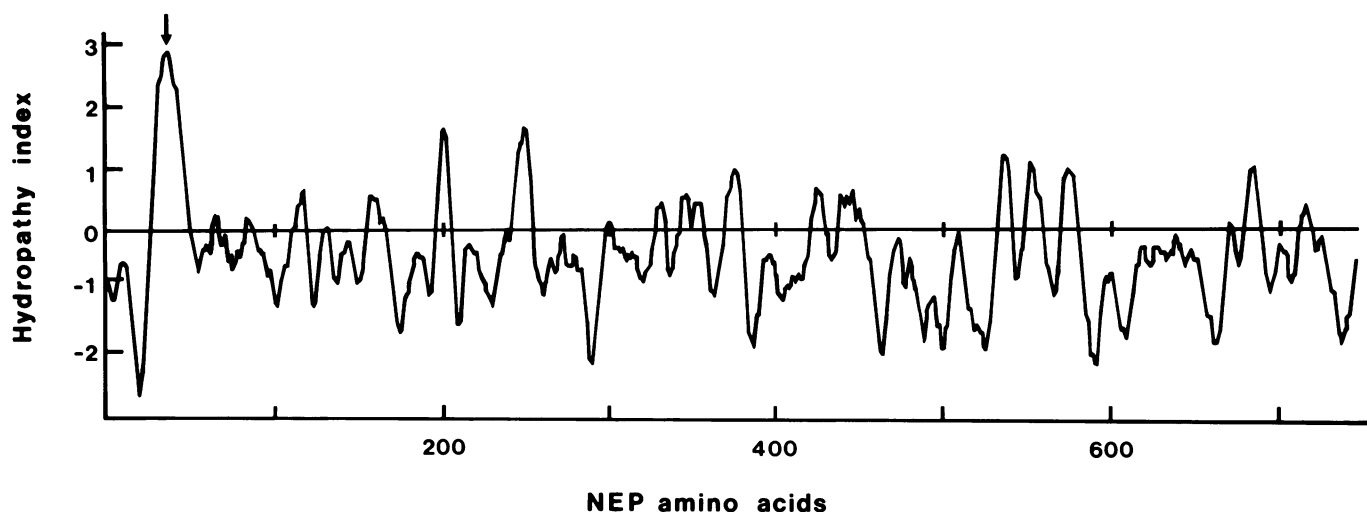


Fig. 4. Hydropathy analysis. The 750 amino acid long NEP sequence was scanned using the computer program of Kyte and Doolittle (1982). Numbers on the horizontal axis refer to the amino acid sequence. Negative values correspond to hydrophilic regions and positive values to hydrophobic regions. The window used in the scanning was nine amino acids. The arrow indicates the only potential membrane-spanning segment of NEP primary structure.

- Hersh,L.B. and Miribaza,R.J. (1986) *Biol. Chem.*, **261**, 6433–6437.
- Hunkapiller,M.W., Lujan,E., Ostrader,F. and Hood,L.E. (1983) *Methods Enzymol.*, **91**, 227–236.
- Hunziker,W., Speiss,M., Semenza,G. and Lodish,H.F. (1986) *Cell*, **46**, 227–234.
- Huynh,T.V., Young,R.A. and Davis,R.W. (1985) In Glover,D. (ed.), *DNA Cloning—A Practical Approach*. IRL Press, London and Washington, DC, pp. 49–88.
- Ikemura,T. (1985) *Mol. Biol. Evol.*, **2**, 13–34.
- Kenny,J. (1986) *Trends Biochem. Sci.*, **11**, 40–42.
- Kester,W.R. and Matthews,B.W. (1977) *J. Biol. Chem.*, **252**, 7704–7710.
- Kyte,J. and Doolittle,R.F. (1982) *J. Mol. Biol.*, **157**, 105–132.
- Laemmli,U.K. (1970) *Nature*, **227**, 680–685.
- Laperche,Y., Bulle,F., Aissani,T., Chobert,M.-N., Aggerbeck,M., Hanoune,J. and Guelaën,G. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 937–941.
- Lathe,R.J. (1985) *J. Mol. Biol.*, **183**, 1–12.
- Lazure,C., Seidah,N.G., Chrétien,M., Lallier,R. and St-Pierre,S. (1983) *Can. J. Biochem. Cell Biol.*, **61**, 287–292.
- Malfroy,G., Swerts,J.-P., Guyon,A., Roques,B.-P. and Schwartz,J.C. (1978) *Nature*, **276**, 523–526.
- Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, NY.
- Matsas,R., Kenny,J. and Turner,A.J. (1986) *Neuroscience*, **18**, 991–1012.
- Ondetti,M.A., Rubin,B. and Cushman,D.W. (1977) *Science*, **196**, 441–444.
- Pozsgay,M., Michaud,C., Liebman,M. and Orłowski,M. (1986) *Biochemistry*, **25**, 1292–1299.
- Proudfoot,N.J. and Brownlee,G.G. (1976) *Nature*, **263**, 211–214.
- Quioco,F.A. and Lipscomb,W.N. (1971) *Adv. Protein Chem.*, **25**, 1–78.
- Recanatini,M., Klein,T., Yang,C.-Z., MacLarin,J., Langridge,R. and Hansch,C. (1986) *Mol. Pharmacol.*, **29**, 436–447.
- Roques,B.P. and Fournié-Zaluski,M.C. (1986) *NIDA Research Monograph Series*, **70**, 128–154.
- Roques,B.P., Fournié-Zaluski,M.C., Soroca,E., Lecomte,B., Malfroy,B., Llorens,C. and Schwartz,J.C. (1980) *Nature*, **288**, 286–288.
- Roques,B.P., Lucas-Soroca,E., Chaillet,P., Costentin,J. and Fournié-Zaluski,M.C. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3178–3182.
- Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
- Semenza,G. (1986) *Annu. Rev. Cell Biol.*, in press.
- Speiss,M., Schwartz,A.L. and Lodish,H.F. (1985) *J. Biol. Chem.*, **260**, 1979–1982.
- Tam,L.-T., Engelbrecht,S., Talent,J.M., Gracy,R.W. and Erdős,E.G. (1985) *Biochem. Biophys. Res. Commun.*, **133**, 1187–1192.
- Tomita,M., Furthmayr,H. and Marchesi,V.T. (1978) *Biochemistry*, **17**, 4756–4770.
- Waksman,G., Hamel,P., Fournié-Zaluski,M.C. and Roques,B.P. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 1523–1527.
- Waksman,G., Hamel,P., Delay-Goyet,P. and Roques,B.P. (1986) *EMBO J.*, **5**.
- Wickner,W.T. and Lodish,H.F. (1985) *Science*, **230**, 400–407.
- Wood,W.I., Gitschier,J., Lasky,L.A. and Lawn,R.M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1585–1588.
- Yamamoto,T., Davis,C.G., Brown,M.S., Schneider,W.J., Casey,M.L., Goldstein,J.L. and Russell,D.W. (1984) *Cell*, **39**, 27–38.

Received on December 31, 1986; revised on February 16, 1987