Organization and chromosomal localization of the human *ECEL1* (*XCE*) gene encoding a zinc metallopeptidase involved in the nervous control of respiration

Olivier VALDENAIRE*[1,2], Elisabeth ROHRBACHER†, An LANGEVELD‡, Anja SCHWEIZER† and Carel MEIJERS‡

*INSERM U460, UFR X.Bichat, 75018 Paris, France, †Pharma Division, Preclinical Cardiovascular Research, F. Hoffmann-La Roche Ltd, 4070 Basel, Switzerland, and ‡Department of Cell Biology and Genetics/Pediatric Surgery, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

ECEL1 (endothelin-converting enzyme-like 1; previously known as XCE) is a putative zinc metalloprotease that was identified recently on the basis of its strong identity with endothelin-converting enzyme. Although the physiological function of ECEL1 is unknown, inactivation of the corresponding gene in mice points to a critical role of this protein in the nervous control of respiration. In the present study we have characterized the human *ECEL1* gene. It was located to region q36–q37 of chromosome 2 and shown to be composed of 18 exons spanning approx. 8 kb. The structure of the *ECEL1* gene displays some striking similarities with those of genes of related metallopeptidases, supporting the hypothesis that they are all derived from a common ancestor. A short phylogenetic study describing the relationship between the various members of this gene family is also presented.

Key words: endothelin-converting enzyme, metalloprotease, phylogeny.

INTRODUCTION

The superfamily of zinc metalloproteases has been shown to be involved in many crucial aspects of development and cellular physiology. Depending on the topology of their catalytic sites, these enzymes have been classified into several categories [1]. The gluzincins are thus characterized by the consensus sequence HEXXH, containing two zinc-co-ordinating His residues, and by possessing a Glu residue as their third zinc ligand. The gluzincin class is itself composed of several gene families. One of these, represented by neutral endopeptidase (NEP) or enkephalinase (EC 3.4.24.11), has been at the centre of growing attention in recent years. The six members of this family known so far are type-II integral membrane proteins, and share a common topology, with a short cytoplasmic N-terminal tail and a large C-terminal extracellular/intraluminal domain bearing the enzyme catalytic site.

NEP is identical to CALLA, the common acute lymphoblastic leukaemia antigen [2–5], and is known to metabolize many peptidergic substrates, including enkephalins, tachykinins, atrial natriuretic factor, bradykinin, endothelins and big endothelins [6]. NEP inhibitors (e.g. thiorphan) exhibit analgesic properties. Another antigen, the Kell blood group protein (Kell), was related to NEP after its molecular identification [7]. Kell antigen incompatibility can have severe consequences (e.g. haemolytic disease of the newborn). Very recently, Kell was shown to be able to convert big endothelin-3 into endothelin-3, although the physiological significance of this activity remains to be proved [8]. Endothelin-converting enzyme (ECE)-1 (EC 3.4.24.71) was identified in 1994 and belongs to the same family [9,10]. It is expressed abundantly in the endothelium and occupies a key position in the biosynthetic pathway of the endothelins, very potent vasoconstrictors, by processing their inactive precursors, or big endothelins. Thus inhibition of ECE-1 is viewed presently as a promising therapeutic tool for a number of cardiovascular pathologies. In addition, as a component of the endothelin system, ECE-1 contributes to the aetiology of a few cases of Hirschprung disease [11]. ECE-2 is remarkably similar to ECE-1 (59 % identity) and was also shown to convert big endothelins into endothelins, albeit at a low pH optimum (5.5 versus 7.0 for ECE-1) [12]. Targeted disruption of the *ECE-1* gene, however, suggested that the role of ECE-2 in endothelin metabolism may be questioned [13]. PEX, the expression of which is restricted to the bones and related tissues, was identified by a positional cloning approach due to its association with hypophosphataemic rickets, a disorder characterized by impaired renal tubular phosphate re-absorption [14], and was shown recently to be able to degrade parathyroid hormone *in vitro* [15].

ECEL1 (endothelin-converting enzyme-like 1), previously termed XCE (following the recommendation of the HUGO nomenclature committee, we will employ the term ECEL1 in place of the previously used XCE), is the newest member of the NEP-related family and was identified by homology cloning [16]. It is expressed specifically in the nervous system, particularly in the medulla oblongata and in the spinal cord, presumably by cholinergic neurons such as motor neurons or striatum interneurons [16]. Although no activity has been reported to date for ECEL1, the conservation in its amino acid sequence of residues that were shown to be crucial for NEP and ECE-1 catalytic activities suggests that ECEL1 is a functional metallopeptidase. In addition, disruption in mice of the *ECEL1* gene results in neonatal lethality due to respiratory failure shortly after birth. Analysis of the phenotype of ECEL1-deficient embryos together with the central nervous system-specific expression of ECEL1

Abbreviations used: ECEL1, endothelin-converting enzyme-like 1; ECE, endothelin-converting enzyme; NEP, neutral endopeptidase; Kell, Kell blood group protein; FISH, fluorescence *in situ* hybridization.

[1] Present address: Actelion Ltd., Gewerbestrasse 16 CH-4123 Allschwil, Switzerland.

[2] To whom correspondence should be addressed, at Actelion Ltd. (e-mail Olivier.Valdenaire@Actelion.Com).

The nucleotide sequence data reported will appear in the GenBank, EMBL, DDBJ and GSDB Nucleotide Sequence Databases under the accession numbers AJ130734, AJ250423 and AJ250424.

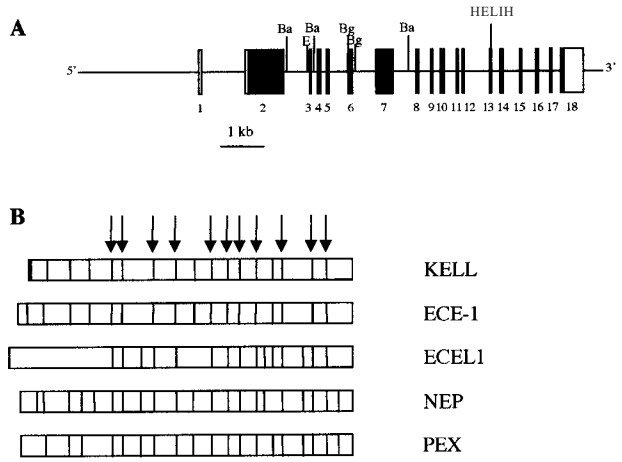

Figure 1 Genomic organization of the human *ECEL1* gene compared with those of related enzymes

(**A**) Structure of the *ECEL1* gene. Exons are boxed, and the coding sequence is represented in black. The exon displaying the characteristic HEXXH zinc-binding motif is indicated, along with the following restriction sites: *Bam*HI (Ba), *Bgl*II (Bg) and *Eco*RI (E). (**B**) Comparison of the exon/intron organization of the NEP-related gene family. Representations of the cDNA coding regions are aligned according to sequence identity, and are interrupted by the exon/intron boundaries. Arrows indicate boundaries that are exactly conserved in the five genes.

suggests a critical role for this enzyme in the nervous regulation of the respiratory system [17].

In this study, we have cloned the human *ECEL1* gene and determined its chromosomal localization by fluorescence *in situ* hybridization (FISH). We have also characterized its exon/intron structure and putative promoter region. Finally, we have analysed the evolutionary relationship that exists between the various members of the NEP family by comparing their sequences and gene structures.

MATERIALS AND METHODS

Cloning of the human *ECEL1* gene

Human ECEL1 cDNA was cloned as described previously [16]. A 2384-bp *Bcl*XI/*Bam*HI fragment of this cDNA was ^{32}P random-primed and used to screen a human genomic library constructed in λSp6/T7 (Clontech). All hybridization procedures were performed at 65 °C in 4 × SSPE [where 1 × SSPE = 0.15 M NaCl/10 mM sodium phosphate (pH 7.4)/1 mM EDTA], 10 × Denhardt's solution, 5% SDS and 100 μg/ml denatured herring sperm DNA. Screening of 10^6 clones revealed 54 positive signals. PCRs were performed on the selected clones immediately after the first round of screening, using two pairs of ECEL1-specific primers deduced from the 5′ and 3′ ends of the cDNA. Two clones that yielded signals for both amplifications, X5 and X32, were thus selected and purified.

Sequencing of the *ECEL1* gene

PCRs (30 cycles) were performed using as substrates the two purified phages X5 and X32, as well as human genomic DNA (100 ng), with various primers deduced from the ECEL1 cDNA sequence. The 5′ and 3′ regions of the gene were obtained using sets of primers including an ECEL1-specific primer and primers located on the phage short or long arms. Amplification products were subcloned into pCR2.1 (Invitrogen) and sequenced with the Thermo Sequenase cycle sequencing kit (Amersham) using a Licor automatic sequencer. At least three different clones were sequenced for each product, which enabled us to identify and discard errors due to *Taq* polymerase. The final sequence of the gene was obtained by combining overlapping sequences.

FISH

A 4732-bp genomic fragment was amplified by PCR out of the X5 genomic clone, using two ECEL1-specific primers (5′-GTGG-AAGTGGCTGCTAGACCAG-3′ and 5′-CAGCAAGAACAC-

Table 1 Exon/intron junctions of the human *ECEL1* gene

Exon sequences are in capital letters. Consensus splice sequences are highlighted in bold.

Exon	Size (bp)	5′ Splice donor	Intron	Size (bp)	3′ Splice donor
1	103	GCCGG/**gt**aagtggcg	1	958	cttcccgc**ag**/CTGCG
2	887	TCCGC/**gt**gagtgcgc	2	553	ccctcccc**ag**/ATTGA
3	69	AGAAG/**gt**gctggggc	3	107	cctggtgc**ag**/ATCCT
4	111	CCAAC/**gt**gagcagac	4	87	ctccccac**ag**/ATCAC
5	93	CCCAC/**gt**gagtgtgg	5	204	ggtctcac**ag**/TTGCG
6	125	CACCG/**gt**atggctgc	6	248	tggtcccc**ag**/GGTCC
7	222	CCAAG/**gt**gggttgcc	7	483	gaccctgc**ag**/GTGCA
8	99	CCAAG/**gt**gaggggggg	8	236	cggcccgc**ag**/CTCCA
9	75	ATGAG/**gt**gggccctg	9	149	cctcctgc**ag**/TTTGA
10	104	TCCAC/**gt**gggtgcct	10	242	ccctccac**ag**/GTGGC
11	59	GATGG/**gt**aaggggat	11	80	cttcttgt**ag**/TGTTC
12	52	CCACA/**gt**gagtacat	12	568	gtgggccc**ag**/GTCTC
13	68	CTGGG/**gt**gaggcctg	13	151	ctgcccac**ag**/GGGGC
14	125	AGCGG/**gt**gagacccc	14	349	ttcccggc**ag**/GTGAA
15	66	ACCAC/**gt**gagcggcc	15	277	gtgtcccc**ag**/GCCTA
16	96	CCCAG/**gt**gggctggg	16	242	ctggaccc**ag**/AACTG
17	77	TACAG/**gt**atgcccac	17	146	cccacccc**ag**/GGTGC
18	395				

AGCGAAGGTGGGG-3′), subcloned into the pCR2.1 plasmid, and used as a probe for FISH as described previously [18].

Dot-blot analysis of ECEL1 expression

A dot blot displaying 2 μg of poly(A)$^+$ from 50 human tissues (Clontech, ref. 7770-1) was hybridized against a ^{32}P-labelled random-primed probe synthesized using a PCR-generated 800-bp ECEL1 cDNA fragment (nucleotides 1009–1817 of the ECEL1 sequence; GenBank accession number Y16187). The membrane was hybridized at 65 °C for 12 h in ExpressHyb solution (Clontech) containing denatured herring sperm DNA (100 μg/ml) and the labelled probe (10^6 c.p.m./ml), washed at 65 °C for 15 min each in 2 × SSC/0.1 % (w/v) SDS (where 1 × SSC is 0.15 M NaCl/0.015 M sodium citrate), in 1 × SSC/0.1 % (w/v) SDS and in 0.1 × SSC/0.1 % (w/v) SDS, and then exposed for 48 h to an X-ray film.

Phylogenetic analysis

Phylogenetic analysis was performed with the GCG (Genetic Computer Group) package: distance matrix was calculated using the Kimura method, and the tree was built using the Neighbor-Joining method. The amino acid sequences that were used for every gene were those that could be aligned with the last 221 residues of ECEL1 (Arg-555 to Trp-775). GenBank/EMBL accession numbers of the sequences used are indicated in the legend of Figure 5.

RESULTS AND DISCUSSION

Isolation and characterization of the human *ECEL1* gene

A human genomic library was screened using an ECEL1 cDNA fragment. There were 54 signals detected, out of which two clones were selected by PCR and purified. Assuming that most of the introns would occur at similar positions in the *ECE-1* and *ECEL1* genes, sets of primers were deduced from the ECEL1 cDNA sequence so as to enable the amplification of ECEL1 intronic regions. The corresponding amplifications were then carried out using these two phages and human genomic DNA as substrates. A total of 10.65 kb, encompassing all *ECEL1* exons and 2.7 kb of the gene's upstream regions, was thus sequenced. As is already known to be the case for the human ECEL1 cDNA, the *ECEL1* gene sequence was characterized by a high GC content of more than 60 %. Alignment of the gene and the cDNA sequences revealed the presence of 18 exons ranging from 52 to 887 bp in size (Figure 1A and Table 1). As shown in Table 1, all the exon/intron boundaries displayed the consensus splice donor and acceptor sequences [19].

Analysis of the gene's upstream regions

The 3.7-kb sequence of genomic DNA localized upstream from the second exon was analysed using the FunSiteP 2.1 Promoter Recognition program (available on the worldwide web at http://transfac.gbf.de/) [20], which revealed that the region most likely to include the gene transcription start sites was located between nucleotides 923 and 940 (Figure 2A). Several cDNA clones were sequenced to get the ECEL1 cDNA 5′ end, and the longest of them began at position 948 [16], in line with this analysis. The genomic region upstream from this position, probably constituting the basal promoter of the *ECEL1* gene, displayed several potential binding sites for known transcription factors, including nuclear factor κB, activator protein 2 (AP-2), Sp1 sites and an E-box. It was shown in particular that Sp1 sites and E-boxes, alone or in combination, could be necessary but also sufficient to drive transcription [21–23]. It is therefore possible that such motifs, located just upstream from the first exon, constitute the core promoter of the *ECEL1* gene.

The genes *NEP* and *ECE-1* were shown to display several distinct promoters [24–27]. In the case of *ECE-1*, this multiplicity

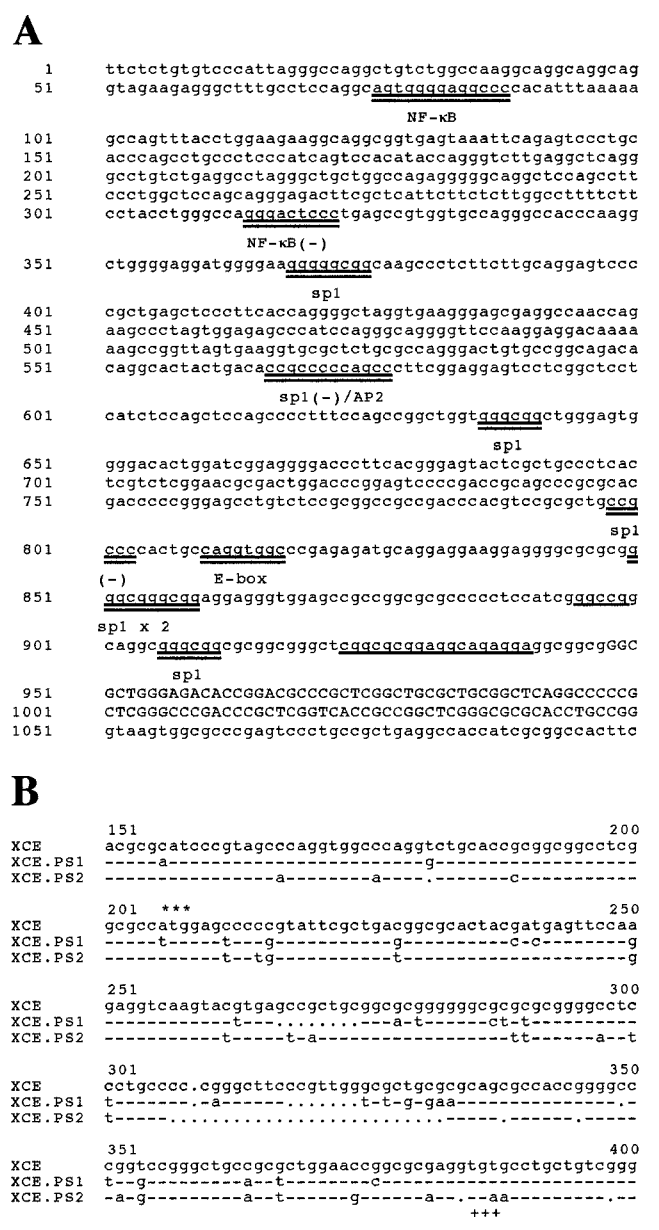

Figure 2 Partial sequences of the *ECEL1* gene 5′ regions (A) and of the two ECEL1 pseudogenes (B)

(**A**) Nucleotide sequence of the putative promoter of the *ECEL1* gene. The sequence of the first exon (i.e. that was displayed by at least one cloned ECEL1 cDNA) is represented by capital letters. Theoretical regions of initiation of transcription as determined using the FunSiteP 2.1 Promoter Recognition program [20] are singly underlined, whereas potential binding sites for known transcription factors are doubly underlined. (**B**) Alignment of short fragments of the two *ECEL1* pseudogenes (ECEL1.PS1 and ECEL1.PS2) with the ECEL1 cDNA sequence (ECEL1). Numbering refers to the previously published *ECEL1* sequence [16]. A dash represents a residue identical to the above ECEL1 cDNA nucleotide and a dot is a deletion. The ECEL1 cDNA start codon (disrupted in ECEL1.PS1) is indicated by ***, while a stop codon interrupting the ECEL1.PS2 reading frame is indicated by + + +.

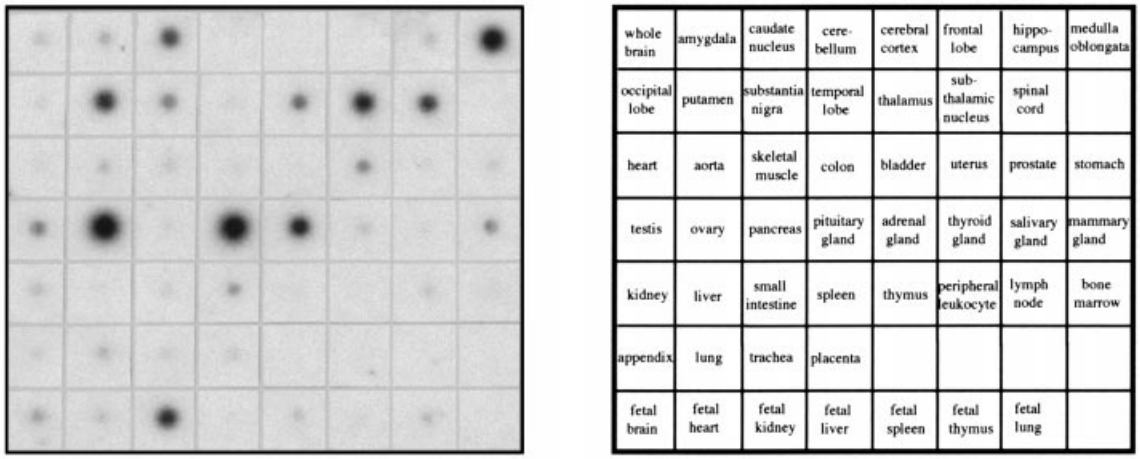

Figure 3 Tissue expression of ECEL1 mRNA as assessed by dot blot

Each dot corresponds to 2 μg of poly(A)$^+$. The left-hand panel corresponds to a 48-h exposure of the hybridized and washed membrane to an autoradiographic film. The human tissues that were assessed for ECEL1 expression, along with their position on the blot, are indicated on the right-hand panel.

leads to the existence of four iso-enzymes divergent in their N-terminal sequences [27]. In contrast, RNase-protection experiments performed on human tissues could only detect one type of ECEL1 mRNA, suggesting the presence of a unique promoter in the *ECEL1* gene [16].

Genomic structures of the NEP-related gene family

The gene structures of four members of the NEP-related family have been published previously [25,28–31], and have been compared with that of ECEL1 (Figure 1B). It is noteworthy that the locations of 11 exon/intron junctions are exactly conserved between the five genes. The corresponding introns interrupt the protein-coding sequences in their last 500 amino acid residues (approximately). This conservation probably reflects the fact that a number of residues located in the whole extracellular/intraluminal domain of the NEP-related enzymes are crucial for the proper folding of these enzymes, and therefore that the corresponding exons are placed under a strong selection pressure during evolution. In contrast, it can be seen that the genomic regions corresponding to the first 200 residues (including the enzyme transmembrane domain and cytoplasmic extremities) are differently organized from one gene to another. In addition, eight junctions, the positions of which are conserved only between two, three or four genes, might help to delineate the relationship between these five genes. Thus analysis of this partial conservation suggests that the most recent genomic duplications that took place to produce these five genes concerned ECE-1 and Kell on the one hand, and NEP and PEX on the other.

Identification of two related pseudogenes

In the course of PCR amplification of the gene intronic regions, we obtained several products displaying a sequence related to, but different from, the ECEL1-specific sequence. These products were amplified from genomic DNA and also from some genomic library clones, and corresponded to two genomic regions presenting a high identity with the second and largest exon of the *ECEL1* gene. As shown in Figure 2(B), the deduced sequences carry many base changes, insertions and deletions and therefore belong to pseudogenes unable to encode any ECEL1-related protein. Sets of primers located in other exons of the *ECEL1* gene only enabled amplification of fragments of the *ECEL1*gene. This suggests that the detected pseudogenes only correspond to partial duplications of the upstream regions of the *ECEL1* gene.

Tissue expression of the *ECEL1* gene

A dot blot displaying 2 μg of each of 50 human tissues was hybridized under high-stringency conditions with an 800-bp probe that was shown by Northern blotting to be specific for ECEL1 [16]. As described previously, high-level expression was detected in specific regions of the nervous system, including the caudate nucleus, the putamen, the subthalamic nucleus, the spinal cord and the medulla oblongata (Figure 3). In addition, endocrine tissues such as the ovary, but also the pituitary and adrenal glands, were found to express high levels of ECEL1 transcripts.

Chromosomal localization of the *ECEL1* gene

The chromosomal localization of the human *ECEL1* gene was determined by FISH analysis of metaphase spreads using a 5-kb genomic fragment as a probe. More than 20 metaphases were analysed, all of them yielding at least one strong signal at chromosome 2q36–q37. In 17 cases two copies of chromosome 2 were highlighted by the FISH. Figure 4 shows a typical example of a metaphase hybridized with the ECEL1 probe.

Most of the other members of the NEP-related gene family have been mapped on the human genome: the genes encoding NEP, Kell, ECE-1 and PEX are located on 3q25, 7q33, 1p36 and Xp22, respectively. Such a dispersion on the genome suggests that ancient genomic DNA duplications are at the origin of the diversity of NEP-related enzymes.

Phylogenetic analysis of the NEP-related gene family

Searching for sequences similar to that of ECEL1 in GenBank/EMBL databanks yielded 29 complete sequences. In particular, organisms whose genomes have been completely analysed revealed the presence of one or several NEP-related enzymes. This is the case for the nematode *Caenorhabditis elegans*, expressing up to six NEP-related enzymes, and of a few bacterial

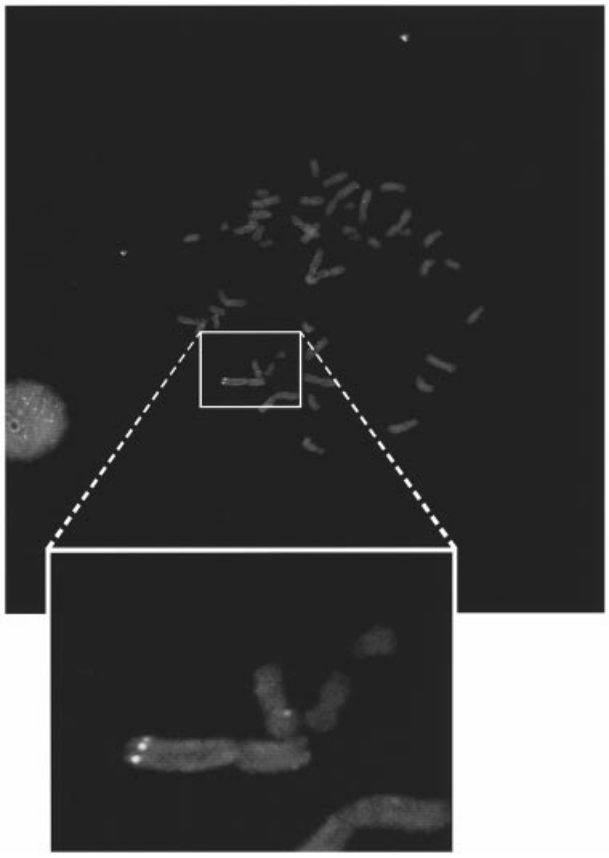

Figure 4 FISH mapping analysis of the human *ECEL1* gene

A typical example of a metaphase hybridized with a 5-kb ECEL1 probe is shown.

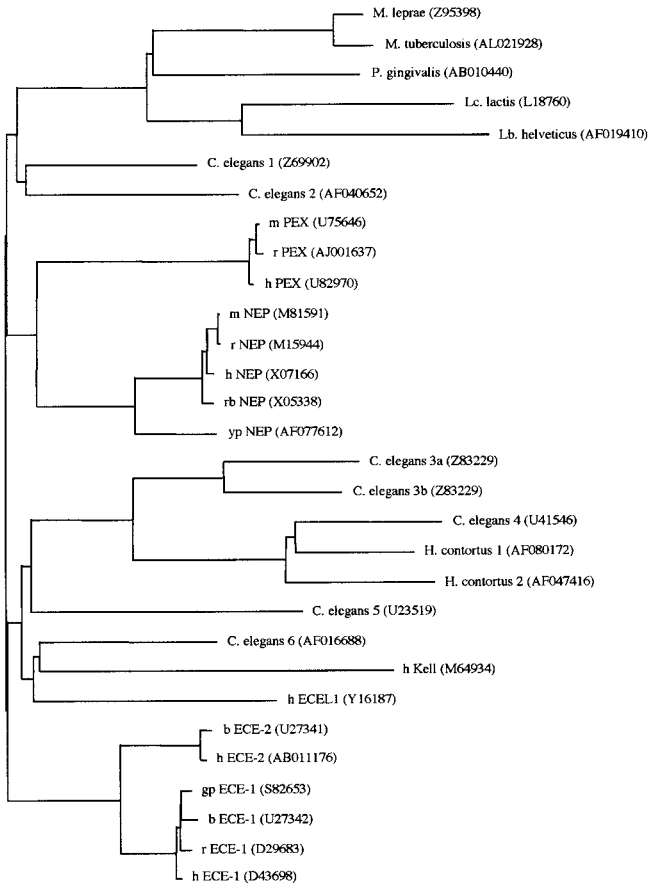

Figure 5 Phylogram of the NEP-related gene family

Analysis was performed with the GCG (Genetic Computer Group) package using the Kimura and Neighbor-Joining methods. The amino acid sequences that were used for every gene were those that could be aligned with the last 221 residues of ECEL1 (Arg-555 to Trp-775). The two halves of *C. elegans 3* (accesion no. Z83229, see text in the Results and discussion section) were each treated as a single gene (*3a* and *3b*). The distance between two genes corresponds to the added lengths of the branches separating the genes from their first common divergence node. This distance is inversely proportional to the degree of similarity shared by the two gene sequences. GenBank/EMBL accession numbers of sequences are indicated in parentheses. Abbreviations: M., *Mycobacterium*; P., *Porphyromonas*; Lc., *Lactococcus*; Lb., *Lactobacillus*; m, mouse; r, rat; h, human; rb, rabbit; yp, yellow perch; H., *Haemonchus*; b, bovine; gp, guinea pig.

strains, each possessing only one member of this family. In contrast, other bacteria represented in the GenBank/EMBL databanks and the yeast *Saccharomyces cerevisiae* do not express such enzymes. This raises the question of whether these few bacterial genes are derived directly from the ancestral gene of this metalloprotease family, or whether they are the consequence of the ancient integration of a eukaryotic gene into a prokaryotic genome. It is of note that the sequences of enzymes belonging to other gluzincin families (angiotensin-converting enzyme, aminopeptidases, etc.) could not be aligned with those of NEP-related enzymes. It is therefore most likely that the different gluzincin subfamilies did not diverge from a common ancestor, but rather converged during evolution towards a comparable zinc-binding site.

All 29 sequences correspond to NEP-like type-II integral membrane proteins, with two major exceptions. One concerns the absence of a putative transmembrane segment in the bacterial proteins. The second exception is given by one of the nematode enzymes (*C. elegans 3*, accession number Z83229 on Figure 5), which displays a single transmembrane domain and two zinc-binding signatures. Interestingly, and as is the case for angiotensin-converting enzyme, *C. elegans 3* results from the duplication of an ancestral NEP-like gene, followed by the fusion of the two copies.

The phylogram presented in Figure 5 merits several comments. It can be seen that the five bacterial enzymes form a subfamily: this indicates that their genes are orthologous (species counterparts). The same also applies to the three *PEX* genes, the five *NEP* genes, the two *ECE-2* genes and the four *ECE-1* genes. None of the invertebrate putative enzymes is clearly a homologue of a vertebrate enzyme, which could mean that the diversification of these two groups occurred after the divergence between the worms and the chordates. It is true however that branching at the base of the phylogram probably has little if any significance. Nevertheless, it is highly probable that contrary to some other classes of proteins (e.g. ion channels), studies in *C. elegans* would not help to clarify the function or even the number of the various mammalian enzymes.

Likewise, the relationship between vertebrate enzymes cannot be described completely. Assessment of the distances between their corresponding sequences indicates that ECE-1 and ECE-2 are closely related and result from a relatively recent gene duplication. Since on the phylogram the distance between these

two genes is equal to the distance separating mammalian NEPs from yellow perch NEP, one could suggest that this divergence corresponds approximately to the emergence of the amphibians. PEX and NEP, although more distantly related, also seem to share exclusively a common first ancestor, which is in line with the hypotheses deduced from the comparison of the family gene structures. On the other hand, the Kell antigen, the gene organization of which is highly conserved with that of ECE-1, is found to be very distant from all the members of the family. This probably corresponds to a higher mutation rate due to a decreased pressure of selection. ECEL1 appears to occupy an intermediate position in the family, being, however, slightly more closely related to ECE-1 than to the other members of the family. The identification of new sequences, especially non-mammalian ones, will certainly help to improve our understanding of the history of this metalloprotease family.

We thank A. Thouard for technical assistance and Dr M. Osborne-Pellegrin for critical reading of the manuscript. This work was supported by grants from the Fondation pour la Recherche Médicale (to O. V.) and the Sophia Foundation for Medical Research (to C. M.).

REFERENCES

1 Hooper, N. M. (1994) FEBS Lett. **354**, 1–6
2 Devault, A., Lazure, C., Nault, C., Le Moual, H., Seidah, N. G., Chretien, M., Kahn, P., Powell, J., Mallet, J., Beaumont, A. et al. (1987) EMBO J. **6**, 1317–1322
3 Letarte, M., Vera, S., Tran, R., Addis, J. B., Onizuka, R. J., Quackenbush, E. J., Jongeneel, C. V. and McInnes, R. R. (1988) J. Exp. Med. **168**, 1247–1253
4 Malfroy, B., Kuang, W. J., Seeburg, P. H., Mason, A. J. and Schofield, P. R. (1988) FEBS Lett. **229**, 206–210
5 Shipp, M. A., Richardson, N. E., Sayre, P. H., Brown, N. R., Masteller, E. L., Clayton, L. K., Ritz, J. and Reinherz, E. L. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 4819–4823
6 Roques, B. P., Noble, F., Dauge, V., Fournie-Zaluski, M. C. and Beaumont, A. (1993) Pharmacol. Rev. **45**, 87–146
7 Lee, S., Zambas, E. D., Marsh, W. L. and Redman, C. M. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 6353–6357
8 Lee, S., Lin, M., Mele, A., Cao, Y., Farmar, J., Russo, D. and Redman, C. (1999) Blood **94**, 1440–1450
9 Shimada, K., Takahashi, M. and Tanzawa, K. (1994) J. Biol. Chem. **269**, 18275–18278
10 Xu, D., Emoto, N., Giaid, A., Slaughter, C., Kaw, S., deWit, D. and Yanagisawa, M. (1994) Cell **78**, 473–485
11 Hofstra, R. M., Valdenaire, O., Arch, E., Osinga, J., Kroes, H., Loeffler, B. M., Hamosh, A., Meijers, C. and Buys, C. H. (1999) Am. J. Hum. Genet. **64**, 304–308
12 Emoto, N. and Yanagisawa, M. (1995) J. Biol. Chem. **270**, 15262–15268
13 Yanagisawa, H., Yanagisawa, M., Kapur, R. P., Richardson, J. A., Williams, S. C., Clouthier, D. E., de Wit, D., Emoto, N. and Hammer, R. E. (1998) Development **125**, 825–836
14 The HYP Consortium (1995) Nat. Genet. **11**, 130–136
15 Lipman, M. L., Panda, D., Bennett, H. P., Henderson, J. E., Shane, E., Shen, Y., Goltzman, D. and Karaplis, A. C. (1998) J. Biol. Chem. **273**, 13729–13737
16 Valdenaire, O., Richards, J. G., Faull, R. L. M. and Schweizer, A. (1999) Mol. Brain. Res. **64**, 211–221
17 Schweizer, A., Valdenaire, O., Köster, A., Lang, Y., Schmitt, G., Lenz, B., Blüthmann, H. and Rohrer, J. (1999) J. Biol. Chem. **274**, 20450–20456
18 Mulder, M. P., Wilke, M., Langeveld, A., Wilming, L. G., Hagemeijer, A., van Drunen, E., Zwarthoff, E. C., Riegman, P. H., Deelen, W. H., van den Ouweland, A. M. et al. (1995) Hum. Genet. **96**, 133–141
19 Breathnach, R. and Chambon, P. (1981) Annu. Rev. Biochem. **50**, 349–383
20 Kondrakhin, Y. V., Kel, A. E., Kolchanov, N. A., Romashchenko, A. G. and Milanesi, L. (1995) Comput. Appl. Biosci. **11**, 477–488
21 Ji, C., Casinghino, S., McCarthy, T. L. and Centrella, M. (1997) J. Biol. Chem. **272**, 21260–21267
22 Li, N. and Seetharam, B. (1998) J. Biol. Chem. **273**, 28170–28177
23 Vaishnaw, A. K., Mitchell, T. J., Rose, S. J., Walport, M. J. and Morley, B. J. (1998) J. Immunol. **160**, 4353–4360
24 Li, C., Chen, G., Gerard, N. P., Gerard, C., Bozic, C. R. and Hersh, L. B. (1995) Gene **164**, 363–366
25 Valdenaire, O., Rohrbacher, E. and Mattei, M. G. (1995) J. Biol. Chem. **270**, 29794–29798
26 Orzechowski, H. D., Richter, C. M., Funke-Kaiser, H., Kroger, B., Schmidt, M., Menzel, S., Bohnemeier, H. and Paul, M. (1997) J. Mol. Med. **75**, 512–521
27 Valdenaire, O., Lepailleur-Enouf, D., Egidy, G., Thouard, A., Barret, A., Vranckx, R., Tougard, C. and Michel, J.-B. (1999) Eur. J. Biochem. **264**, 341–349
28 D'Adamio, L., Shipp, M. A., Masteller, E. L. and Reinherz, E. L. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 7103–7107
29 Lee, S., Zambas, E., Green, E. D. and Redman, C. (1995) Blood **85**, 1364–1370
30 Francis, F., Strom, T. M., Hennig, S., Boddrich, A., Lorenz, B., Brandau, O., Mohnike, K. L., Cagnoli, M., Steffens, C., Klages, S. et al. (1997) Genome Res. **7**, 573–585
31 Holm, I. A., Huang, X. and Kunkel, L. M. (1997) Am. J. Hum. Genet. **60**, 790–797

Received 4 August 1999/8 November 1999; accepted 7 January 2000