RESEARCH COMMUNICATION
C8, a new member of the convertase family

Angela BRUZZANITI*‡, Katrina GOODGE*‡, Philippe JAY†, Sylvie A. TAVIAUX†, Mark H. C. Lam*, Philippe BERTA†, T. John MARTIN*, Jane M. MOSELEY* and Matthew T. GILLESPIE*§

*St. Vincent's Institute of Medical Research and The University of Melbourne, Department of Medicine, St. Vincent's Hospital, 41 Victoria Parade, Fitzroy, Vic. 3065, Australia, and †Centre de Recherches de Biochimie Macromoleculaire, CNRS UPR9008, INSERM U.249, Route de Mende, 34033 Montpellier, France

A novel subtilisin-like protein, PC8, was identified by PCR using degenerate primers to conserved amino acid residues in the catalytic region of members of the prohormone convertase family. PC8 was predicted to be 785 residues long and was structurally related to the mammalian convertases furin, PACE4, PC1 and PC2, sharing more than 50 % amino acid identity over the catalytic region with these family members. PC8 possessed the catalytically important Asp, His, Asn and Ser amino acids, the homo B domain of this family of enzymes and a C-terminal hydrophobic sequence indicative of a transmembrane domain.

Structurally, PC8 is more related to furin and PACE4 than to PC1 or PC2. Like furin and PACE4, PC8 mRNA was found to be widely expressed; this is in contrast with PC1 and PC2, which have a restricted distribution. Two transcripts, of 4.5 and 3.5 kb, were detected in both human cell lines and rat tissues. Unlike furin and PACE4, both of which map to chromosome 15, PC8 maps to chromosome 11q23–11q24, suggesting that this gene may have resulted from an ancient gene duplication event from either furin or PACE4, or conversely that these genes arose from PC8.

INTRODUCTION

Many biologically active proteins are synthesized as precursors which require specific endoproteolytic cleavage, commonly after dibasic residues, to generate their mature active forms. Over the last decade, a family of enzymes which is responsible for cleavage of pro-proteins to their biologically active forms has been identified [1–3]. This family of enzymes, known as 'prohormone convertases', are calcium-dependent serine proteinases exhibiting considerable sequence similarity to the bacterial subtilisins.

Kex2 was the first convertase to be described and was isolated and cloned from *Saccharomyces cerevisiae* [4,5]. Subsequently a further six members of this family, which exhibit highly conserved and similar catalytic domains, have been cloned from mammalian cells. These are furin [6,7], PC1 [8], also called 'PC3' [9], PACE4 [10,11], also identified as 'PC7' [12], PC2 [8,13], PC4 [14,15] and PC5 [16,17] also known as PC6 [18].

Furin, PACE4 and PC5 are expressed in a variety of tissues and cell lines. However, there are differences in their relative abundance [7,10–12,17–19]. Whereas PC1 and PC2 have more restricted distribution and are predominant in the anterior and intermediate pituitary respectively [8,20–22], PC4 is restricted to the testis [14,15].

The consensus motif for precursor cleavage by furin is marked by an R-X-K/R-R sequence. However, furin has also been shown to cleave at R-X-X-R and R-X-X-X-K/R-R sequences [1–3,19,23]. In contrast, PC1 and PC2 have been shown to direct precursor cleavage following dibasic amino acid pairs [20,24]; PC1 can also cleave at mono-arginyl sites with basic residues at the -4 or -6 position [25]. These sites are found in many pro-peptides and polypeptide precursors, and the mammalian pro-

hormone convertases have been found to cleave a number of these, including pro-opiomelanocortin [24], proinsulin [26], pro-insulin receptor [27], prorelaxin [28], proalbumin [29], pro-(von Willebrand factor) [19] and pro-(β-nerve growth factor) [23].

In the course of studies aimed at understanding the post-translational processing of parathyroid hormone-related protein, which elicits multiple biological actions and contains a number of potential sites for convertase action, including two furin target sequences [30,31], we have identified a new member of the convertase family. Reported here is the identification, tissue distribution and chromosomal assignment of this new convertase which we have designated 'PC8'.

MATERIALS AND METHODS

Cell culture

The human squamous cancer cell lines BEN, COLO 16 and the spontaeously immortalized human keratinocyte cell line HaCaT were cultured as previously described [31–33].

PCR oligonucleotides

oPC8 {5′-GGGCCCCA(G/A)CTNGC(G/A)CT(G/A)TA-3′, antisense strand oligonucleotide to amino acids 250–256 of human furin [7]}, oPC10 {5′-TANCC(A/G)TC(A/G)CA(A/G)T(C/T)(A/G)CA-3′, antisense strand oligonucleotide to amino acids 303–308 of human furin [7]}, oPC12 {5′-TGCTGCAT(A/G)TCNCGCCA-3′, antisense strand oligo-nucleotide to amino acids 390–395 of human furin [7]}, oPC13

{5′-CA(T/C)GG(A/C)AC(A/T)(A/C)G(A/C/G)TGTGC-(A/C/T)GG-3′, sense strand oligonucleotide to amino acids 194–200 of human furin [7]}, oPC33 [5′-CAACAACAGCT-TCTGTGC-3′, sense strand oligonucleotide to nucleotides 740–758 of human PC8 (Figure 1)], oPC34 [5′-TACCCGGA-TACCTGCGAT-3′, antisense strand oligonucleotide to nucleotides 784–801 of human PC8 (Figure 1)], oPC35 [5′-TGAT-CTGATAGTGCTTGT-3′, antisense strand oligonucleotide to nucleotides 845–861 of human PC8 9 (Figure 1)], T7 (5′-AATACGACTCACTATAG-3′), GAPDH-1, GAPDH-3, GAPDH-4 and GAPDH-5 have been described previously [33,34].

Northern-blot analysis and reverse transcription PCR (RT-PCR)

Total RNA was extracted using guanidinium chloride as described [35] and polyadenylated RNA was recovered from total RNA using a QuickPrep Micro mRNA purification kit (Pharmacia Biotech Inc., Uppsala, Sweden). Northern analysis and hybridizations were performed as previously reported [33]. Specifically bound probe was quantified by phosphorImager analysis (Molecular Dynamics, Sunnyvale, CA, U.S.A.).

RT-PCR was performed on 5 μg total RNA isolated from cells or tissues as previously described [33,34]. PCR cycling conditions were: 94 °C for 30 s, 45–55 °C for 60 s (depending on the oligonucleotide combinations), 72 °C for 30 s to 2 min, for 40 cycles followed by 10 min elongation time at 72 °C. PCR products resolved on a 2 % (w/v)-agarose gel were subsequently extracted using Qiaex DNA purification kit (Qiagen Inc., Chatsworth, CA, U.S.A.) or Wizard PCR Preps (Promega Corp., Madison, WI, U.S.A.) and PCR products were cloned into pCRScript II (Stratagene, La Jolla, CA, U.S.A.) or pGEMT (Pharmacia Biotech Inc.). DNA sequence analysis from both strands was performed using Sequenase sequencing (USB, Cleveland, OH, U.S.A.). Sequences were then compared with sequences in the GenBank™ databank.

Chromosomal localization

Metaphases were obtained from phytohaemagglutinin-stimulated lymphocytes from a healthy donor after thymidine synchronization and bromodeoxyuridine incorporation. The full-length cDNA HE6WCR40 was labelled with biotinyl-16-dUTP by random priming. Fluorescence *in situ* hybridization was performed as previously described [37].

RESULTS AND DISCUSSION

In order to identify convertase family members which may be involved in post-translational processing of parathyroid-hormone-related protein (PTHrP), we characterized PTHrP-producing cells for their convertase mRNA complement. Degenerate oligonucleotide primers (oPC2, oPC8, oPC10, oPC12 and oPC13) were designed to amino acid residues conserved between the catalytic regions of furin, PC1 and PC2. These primers, when used in RT-PCR, would allow the identification of known convertases and potentially new members of the convertase family. Total RNA from BEN cells was reverse-transcribed with random hexamers, and 30 cycles of PCR undertaken with the primers oPC8 and oPC13. Resultant PCR fragments of approx. 190 bp were subcloned into pCRScript and sequenced. By this approach, 18 clones showing sequence similarity to furin, two to PACE4 and two to PC2, were recovered. In addition, another clone, BC20, of 187 bp was detected, and this was highly related to, but distinct from, the other convertase cDNA clones.

The cDNA sequence of BC20 was extended over the catalytic domain by RT-PCR with a BC20-specific sense-strand oligonucleotide, oPC33, and antisense-strand degenerate primers (oPC10 and oPC12) to conserved amino acids in the catalytic domain between furin, PC1 and PC2, which resulted in fragments 343 and 617 bp respectively; the sequences of these fragments have been previously reported [38,39]. It was clear that BC20 formed a new member of the mammalian prohormone convertase family and we refer to it now as 'PC8', adopting the nomenclature of Seidah et al. [8]. 5′ Sequences were amplified by anchored PCR from a human liver cDNA library in pcDNAIneo (Stratagene) using the PC8-specific antisense oligonucleotide (oPC34) and a sense-strand-specific primer to the T7 RNA polymerase promoter. Combined, these strategies yielded overlapping sequences from the 5′ untranslated regions (UTR) to the end of the catalytic domain of PC8 (Figure 1, nucleotides 1–1320). During database searches a cDNA clone, HE6WCR40 (GenBank™ accession no. H58761), representing an N-terminally truncated PC8 cDNA, was identified. This clone provided confirmation of the sequence obtained by RT-PCR and the complete 3′ coding region of PC8. Subsequently, the entire coding region of PC8 (GenBank™ accession no. U40623) was amplified from BEN cell RNA by RT-PCR and the sequence was identical with that determined for HE6WCR40. The cDNA clones predicted PC8 to be encoded by a mRNA species of 3.3 kb, composed of a 21 bp 5′ UTR, an open reading frame of 2355 bp coding for a protein of 785 amino acids and a 3′ UTR of 925 nucleotides. The 3′ UTR contained a polyadenylation signal, AATAAA, which was followed 12 nucleotides later by a poly(A) stretch (not shown). The primary protein sequence predicts a 42-residue signal peptide at the N-terminus, six potential N-linked glycosylation sites (residues 167, 175, 200, 241, 511 and 763) and a 22-amino-acid transmembrane region between residues 666 and 687 (Figure 1). The Met residue at position 36 is within a consensus Kozak sequence [40]. However, if this was the initiating AUG codon, the protein would lack a signal peptide. The N-terminus of mature PC8 is assumed to begin at Ser-142, following the sequence R-A-K-R, which fits the R-X-K/R-R consensus motif cleavage site for furin. While this manuscript was in preparation, a cDNA clone, HSU33849, identical with PC8 (except for a 131-nucleotide extension of the 5′ UTR and the last nucleotide of HSU33849 corresponding to nucleotide 2559 in Figure 1) and a partial cDNA clone, 45900, sharing structural similarity with PC8, were deposited in the GenBank™ database (accession nos. U33849 and H09374 respectively).

The chromosomal location of the PC8 gene was determined by *in situ* hybridization using the HE6WCR40 cDNA clone. A total of 30 metaphase preparations were analysed, and a specific labelling of bands 11q23–q24 was observed on four (seven cells), three (four cells), two (12 cells) and one (four cells) chromatid of the chromosome homologues (Figure 2). In total, 27 non-specific hybridizations were scattered throughout the other chromosomes, and no secondary peak was detected.

Comparison of PC8 with other mammalian convertases

An amino acid alignment of PC8 to the known human convertases furin, PACE4, PC1 and PC2 is given in Figure 3. Each possesses a signal peptide, an N-terminal domain, a catalytic domain, a homo B domain and a C-terminal domain. PC8 shares its highest identity with furin (54 %), PACE4 (53 %), PC1 (53 %) and PC2 (52 %) within the 289 amino acid catalytic domain. In addition, PC8 possesses the conserved amino acids Asp (197), His (228), Asn (329; Asp in PC2) and Ser (406), which are catalytically important. Outside the catalytic domain, the con-
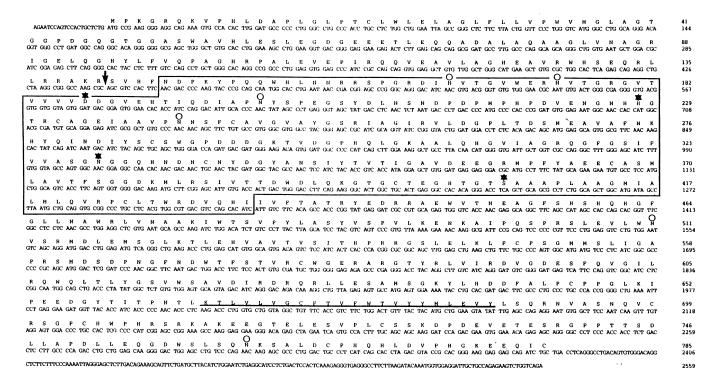
```
                    M   P   K   G   R   Q   K   V   P   H   L   D   A   P   L   G   L   P   T   C   L   W   L   E   L   A   G   L   F   L   L   V   P   W   V   M   G   L   A   G   T                41
AGAATCCAGTCCACTGCTCTG ATG CCG AAG GGG AGG CAG AAA GTG CCA CAC TTG GAT GCC CCC CTG GGC CTG CCC ACC TGC CTC TGG CTG GAA TTA GCC GGG CTC TTC TTA CTG GTT CCC TGG GTC ATG GGC CTG GCA GGG ACA         144

    G   G   P   D   G   Q   G   T   G   G   A   S   W   A   V   H   L   E   S   L   E   G   D   G   E   E   E   T   L   E   Q   Q   A   D   A   L   A   Q   A   A   G   L   V   N   A   G   R            88
GGT GGG CCT GAT GGC CAG GGC ACA GGG GGG GCG AGC TGG GCT GTG CAC CTG GAA AGC CTG GAA GGT GAC GGG GAG GAA GAG ACT CTT GAG CAG CAG GCG GAT GCC TTG GCC CAG GCA GCA GGG CTG GTG AAT GCT GGA CGC          285

    I   G   E   L   Q   G   H   Y   L   F   V   Q   P   A   G   H   R   P   A   L   E   V   E   P   I   R   Q   Q   V   E   A   V   L   A   G   H   E   A   V   R   W   H   S   E   Q   R   L            135
ATC GGA GAG CTT CAG GGG CAC TAC CTC TTT GTC CAG CCT GCT GGG CAC AGG CCG GCC CTG GAG GTG GAG CCC ATC CGG CAG CAG GTG GAG GCT GTG TTG GCT GGG CAT GAA GCT GTG CGC TGG CAC TCA GAG CAG AGG CTG          426

    L   R   R   A   K   R↓S   V   H   F  [N   D   P   K   Y   P   Q   Q   W   H   L   N   N   R   R   S   P   G   R   D   I   N   V   T   G   R   G   V   T                                          182
CTA AGG CGG GCC AAG CGC AGC GTC CAC TTC AAC GAC CCC AAG TAC CCG CAG CAA TGG CAC CTG AAT AAC CGA CGG AGC CCG GGC AGG GAC ATC AAC GTG ACG GGT GTG TGG GAA CGC AAT GTG ACT GGG CGA GGG GTG ACG          567

    V   V   V   V   D   D   G   V   E   H   T   I   Q   D   I   A   P   N   Y   S   P   E   G   S   Y   D   L   N   S   N   D   P   D   P   M   P   H   P   D   V   E   N   G   N   H   H   G            229
GTG GTG GTA GTG GAT GAC GGA GTG GAA CAC ACC ATC CAG GAC ATT GCA CCC AAC TAT AGC CCT GAG GGT AGC TAT GAC CTC AAC TCT AAT GAC CCT GAC CCC ATG CCC GAC CCG GAT GTG GAG AAT GGC AAC CAC CAT GGC          708

    T   R   C   A   G   E   I   A   A   V   P   N   N   S   F   C   A   V   G   V   A   Y   G   S   R   I   A   G   I   R   V   L   D   G   P   L   T   D   S   M   E   A   V   A   F   N   K            276
ACG CGA TGT GCA GGA GAG ATC GCG GCT GTG CCC AAC AAC AGC TTC TGT GCC GTG GGC GTG GCC TAC GGG AGC CGC ATC GCA GGT ATC CGG GTA CTG GAT GGA CCT CTC ACA GAC AGC ATG GAG GCA GTG GCG TTC AAC AAG          849

    H   Y   Q   I   N   D   I   Y   S   C   S   W   G   P   D   D   D   G   K   T   V   D   G   P   H   Q   L   G   K   A   A   L   Q   H   G   V   I   A   G   R   Q   G   F   G   S   I   F            323
CAC TAT CAG ATC AAT GAC ATC TAC AGC TGC AGC TGG GGA CCA GAT GAC GAT GGG AAG ACA GTG GAT GGC CCT CAT CAA CTT GGG AAG GCT GCC TTA CAA CAT GGG GTG ATT GCT GGT CGC CAG GGC TTT GGG AGC ATC TTT          990

    V   V   A   S   G   N   G   G   Q   H   N   D   N   C   N   Y   D   G   Y   A   N   S   I   Y   T   V   T   I   G   A   V   D   E   E   G   R   M   P   F   Y   A   E   E   C   A   S   M            370
GTG GTA GCC AGT GGC AAC GGA GGC CAA CAC AAC GAC AAC TGC AAC TAC GAT GGC TAC GCC AAC TCC ATC TAC ACC GTC ACC ATA GGA GCT GTG GAT GAG GAG GGA CGC ATG CCT TTC TAT GCA GAA GAA TGT GCC TCC ATG          1131

    L   A   V   T   F   S   G   G   D   K   M   L   R   S   I   V   T   T   D   W   D   L   Q   K   G   T   G   C   T   E   G   H   T   G   T   S   A   A   P   L   A   A   G   M   I   A            417
CTG GCA GTC ACC TTC AGT GGT GGG GAC AAG ATG CTT CGG AGC ATT GTG ACC ACT GAC TGG GAC CTT CAG AAG GGC ACT GGC TGC ACT GAG GGC CAC ACA GGG ACC TCA GCT GCA GCG CCT CTG GCA GCT GGC ATG ATA GCC          1272

    L   M   L   Q   V   R   P   C   L   T   W   R   D   V   Q   H   I  [I   V   F   T   A   T   R   Y   E   D   R   R   A   E   W   V   T   N   E   A   G   F   S   H   S   H   Q   H   G   F            464
TTA ATG CTG CAG GTG CGG CCC TGC CTC ACG TGG CGT GAC GTC CAG CAC ATC ATT GTC TTC ACA GCC ACC CGG TAT GAG GAT CGC CGT GCA GAG TGG GTC ACC AAC GAG GCA GGC TTC AGC CAT AGC CAC CAG CAC GGT TTC          1413

    G   L   L   N   A   W   R   L   V   N   A   A   K   I   W   T   S   V   P   Y   L   A   S   Y   S   V   P   V   L   K   E   N   K   A   I   P   Q   S   P   R   S   L   E   V   L   W   N            511
GGC CTC CTC AAC GCC TGG AGG CTC GTG AAT GCA GCC AAG ATC TGG ACA TCT GTC CCT TAC TTA GCA TCC TAC GTC AGT CCC GTG TTA AAA GAA AAC AAG GCG ATT CCG CAG TCC CCC CGT TCC CTG GAG GTC CTG TGG AAT          1554

    V   S   R   M   D   L   E   M   S   G   L   K   T   L   E   H   V   A   V   T   V   S   I   T   H   P   R   R   G   S   L   E   L   K   L   F   C   P   S   G   M   M   S   L   I   G   A            558
GTC AGC AGG ATG GAC CTG GAG ATG TCA GGG CTG AAG ACC CTG GAG CAT GTG GCA GTG ACA GTG TCC ATC ACT CAC CCA CGG CGC GAC AGC TTG GAG CTG AAG CTG TTC TGC CCC AGT GGC ATG ATG TCC CTC ATC GGC GCC          1695

    P   R   S   M   D   S   D   P   N   G   F   N   D   W   T   F   S   T   V   R   C   W   G   E   R   A   R   G   T   Y   R   L   V   I   R   D   V   G   D   E   S   F   Q   V   G   I   L            605
CCC CGC AGC ATG GAC TCG GAT CCC AAC GGC TTC AAT GAT TGG ACC TTC TCC ACT GTG CGA TGC TGG GGG GAG AGA GCC CGA GGG ACC TAC AGG CTT GTC ATC AGG GAT GTC GGG GAT GAG TCA TTC CAG GTC GGC ATC CTC          1836

    R   Q   W   Q   L   T   L   Y   G   S   V   W   S   A   V   D   I   R   D   R   Q   R   L   L   E   S   A   M   S   G   K   Y   L   H   D   D   F   A   L   P   C   P   P   G   L   K   I            652
CGG CAA TGG CAG CTG ACC CTA TAT GGC TCT GTG TGG AGT GCA GTA GAC ATC AGG GAC AGA CAA AGG CTG TTA GAG AGT GCC ATG AGT GGA AAA TAC CTG CAC GAT GAC TTT GCC CTG CCC TGC CCA CCG GGG CTG AAA ATT          1977

    P   E   E   D   G   Y   T   I   T   P   N   T   L   K   T   L   V   L   V   G   C   F   T   V   F   W   T   V   Y   Y   M   L   E   V   L   S   Q   R   N   V   A   S   N   Q   V   C            699
CCT GAG GAA GAT GGT TAC ACC ATC ACC CCC AAC ACC CTC AAG ACC CTG GTG CTG GTA GGC TGT TTC ACC GTC TTC TGG ACT GTT TAC TAC ATG CTG GAA GTA TAT TTG AGC CAG AGG AAT GTG GCT TCC AAT CAA GTT TGT          2118

    R   S   G   P   C   H   W   P   H   R   S   R   K   A   K   E   E   G   T   E   L   E   S   V   P   L   C   S   S   K   D   P   D   E   V   E   T   E   S   R   G   P   P   T   T   S   D            746
AGG AGT GGA CCC TGC CAC TGG CCC CAT CGG AGC AGG AAA GCC AAG GAG GAG GGC ACA GAG CTA GAA TCA GTG CCA CTT TGC AGC AGC AAG GAT CCA GAC GAA GTG GAA ACA GAG AGG GGC CCT CCC ACC ACC TCT GAC          2259

    L   L   A   P   D   L   L   E   Q   G   D   W   S   L   S   Q   N   K   S   A   L   D   C   P   H   Q   H   L   D   V   P   H   G   K   E   Q   I   C   *                                        785
CTC CTT GCC CCA GAC CTG CTG GAG CAA GGG GAC TGG AGC CTG TCC CAG AAC AAG AGC GCC CTG GAC TGC CCT CAT CAG CAC CTA GAC GTA CCG CAC GGG AAG GAG GAG CAG ATC TGC TGA CCTCAGGGCCTGACAGTGTGGGACAGG          2406

CTCTTCTTTCCCAAAATTAGGGAGCTCTTGACAGAAAGCAGTTCTGATGCTTACATCTGGAATCTGAGGCATCCTCTGACTCCACTCAAAGAGGGTGAGGGCCTTCTTAAGATACAAATGGTGGAGGATTGCTGCCAGAGAAGTCTGGTCAGA                              2559
```

Figure 1   Nucleotide and predicted amino acid sequences of PC8

The putative cleavage site to release the mature protein from the propeptide is indicated by the arrow (↓), and the putative transmembrane region is underlined. The subtilisin-like catalytic domain is boxed, and the catalytically important Asp, His, Asn, and Ser residues are indicated by ✱. Consensus sites for Asn-linked glycosylation are marked by the open circles (○).
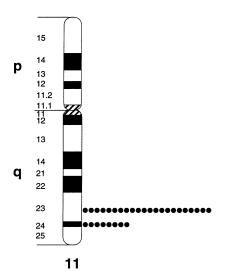

Figure 2   Idiogram of chromosome 11

Idiogram of the chromosome 11 homologue showing the unequivocal distribution of fluorescent spots on the 11q23–q24 bands for 15 metaphases.

served amino acid identity between PC8 and other convertases is dramatically reduced. From the initiating Met to the catalytic region (145 amino acids), PC8 shares 17 % identity with furin, 18 % with PACE4 and PC1 and 11 % with PC2. Within this region, the R-X-K/R-R consensus motif (PC8 residues 138–141)

for precursor cleavage catalysed by furin is conserved with other members of the prohormone convertase family (Figure 3). Sequences C-terminal to the catalytic domain (351 amino acids), which encompass the homo B domain, were marginally more conserved than those of the N-terminus: 24 % identity with PACE4, 20 % with furin and PC1 and 16 % with PC2. PC8 contains five Cys residues (232, 245, 337, 367 and 398), which are conserved with furin, PACE4, PC1 and PC2; the conservation of these residues indicates that they may be involved in disulphide bridges. The C-terminal Cys residue is only shared with furin and PACE4. However, PC8 does not contain a Cys-rich region like furin or PACE4. Only one of the six potential N-linked glycosylation sites of PC8 (Asn-241) was conserved with another family member (PACE4, Asn-199).

Unlike furin, PACE4, PC1 and PC2, PC8 does not contain an RGD sequence, which is a receptor recognition signal of extracellular-matrix proteins, but contains an RGS sequence (residues 539–541) at the position corresponding to the RGD motif in the other convertases.

Expression of PC8 mRNA

PC8 mRNA transcripts of 3.5 and 4.5 kb were found to be expressed in human cell lines (Figure 4A) and in all rat tissues examined (results not shown); these cell lines and tissues were also found to express mRNA for furin and PACE4 (results not shown). The 4.5 kb transcript may result from either alternative mRNA splicing, alternate promoter usage or use of a different polyadenylation motif. For confirmation of the Northern blots, RT-PCR was performed with the PC8-specific primers oPC33
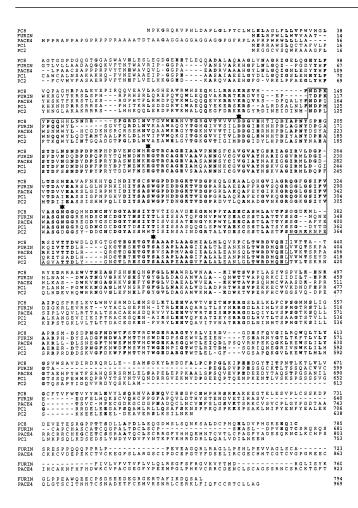
```
PC8                                              MPKGRQKVPHLDAPLGLPTCLWLELAGLFLLVPWVMGL  38
FURIN                                                           MELRPWLLWVVAAT--  14
PACE4   MPPRAPPAPGPRPPPRAAAATDTAAGAGGAGGAGGAGGPGFRPLAPRPWRWLLLLA----  56
PC1                                                             MERRAWSLQCTAFVLF  16
PC2                                                             MKGGCVSQWKAAAGFL  16

PC8     AGTGGPDGQGTGGASWAVHLESLEGDGEHETLEQQADALAQAAGLVNAGRIGELQGHYLF  98
FURIN   GTLVLLAADAQGQKVFTNTWAVRIP-GGPA----VANGVARKHGFLNLGQI-FGDYYHF  67
PACE4   L-LPAACSADYPRRFVYTNHWAVKIP-GGPA----EADRVAAKHGFLNLGQIGNLEDYYHF 109
PC1     CAWCALNSAKAKRQ-FVNEWAAEIP-GGPA----AASAIAEELGYDLLGQIGSLENHYLF  70
PC2     --FCVMVFASAERPVFTNHFLVELHKGQA----KARQVAAEHGFG-VRKLPFAEGLYHF  69

PC8     VQPAGHRPALEVEPIRQQVEAVLAGHEAVRWHSEQRLLRRAKRSVH---------FNDPK 149
FURIN   WHRGVTKRSLSPH---RPRHSRLQREPQVQWLEQQVAKRRTKRDVYQ--EP----TDPK 117
PACE4   YHSKTFRKRSTLSS---RGPHTFLRMDPQVKWLQQQEVKRRVKRQWRS--DFQALYFNDPK 164
PC1     KHKNHPRNSRRSA---FHITKRLSDDDRVIWAEQQYEKERSKRS--ALRDSALNLFNDPM 125
PC2     YHNGLAKAKRRRS---LHHKQQELQRDPRVKMALQQEGFDRKKRGYRDINEIDINM-WDPL 125

PC8     YPQQHHLNNRRR----SPGRDINVTGVMERNVTGRGVTVVVVVDGVEHTIQDIAPNYSPRG 205
FURIN   FPQQWYL------SGVTQRDLNVKAAWAQGYTGHGIVVSILDDGIEKNHPDLAGNYDPGA 171
PACE4   WSNMWYL-HCGDKNSRCRSEMNVQAAWKRGYTGKNVVVTILDDGIERNHPDLAPNYDSYA 223
PC1     WNQQWYLQDTRMTAALPKLDLHVIPVWQKGITGKGVVITVLDDGLEWNHTDIYANYDPEA 185
PC2     FTKQWYLINTGQADGTPGLDLNVAEAWELGYTGKGVTIGIMDDGIDYLHPDLASNYNAEA 185

PC8     SYDLNSNDPDPMPHPDVENGNHHGTRCAGEIAAVPNNSFCAVGVGVRDHGVRDVLDGP- 264
FURIN   SFDVNDQDPDPQPRYTQMNDNRHGTRCAGEVAAVANNNGVCGVGVAYNARIGGVRMLDGE- 230
PACE4   SYDVNGNDYDPSPRYDASNENKHGTRCAGEVAASANNSYCIVGIAYNAKIGGIRMLDGD- 282
PC1     SYDFNDNDHDPFPRYDPTNENKHGTRCAGEIAMQANNHKCGVGVAYNSKVAGIRMLD-GD 244
PC2     SYDFSSNDPYPYPRYTDDWFNSHGTRCAGEVSAAANNNICGVGVAYNSKVAGIRMLDQPF 245

PC8     LTDSMEAVAFNKHYQINDIYSCSWGPDDDGKTVDGPHQLGKAALQHGVIAGRGQFGSIFV 324
FURIN   VTDAVEARSLGLNPNHIHIYSASWGPEDDGKTVDGPARLAEEAFFRGVSQGRGGLGSIFV 290
PACE4   VTDAIEAKSLGISRPNYIDIYSASWGPDDDGKTVDGPGRLAQRAFEYGIKKGRGGLGSIFV 342
PC1     VTDAIEASSICGISRPNYIDIYSASWGPNDDGKTVDGPGRLAQRAFEYGVKQGRGQGKGSIYV 304
PC2     MTDIIEASSISHMPQLIDIYSASWGPTDNGKTVDGPRDVTLQXAMADGVNKGRGGKGSIYV 305

PC8     VASGNGGQHNDNCNYDGYANSIYTVTIGAVDEEGRMPFYAEECASMLAVTPSGGDKML-- 382
FURIN   WASGNGGREHDSCNCDGYTNSIYTLSISSATQFGNVPWYSEACSSTLATTYSSG--NQNE 348
PACE4   WASGNGGREGDYCSCDGYTNSIYTISVSSATRNGYKPWYLEECASTLATTYSSG--AFYE 400
PC1     WASGNGGRQGDRNCDCDGYTDSIYTISISSSAGLSPWYAEKCSSTLASTYSSGDTSYERKRNPE 362
PC2     WASGDGGSY-DDCNCDGYASSMNTISINSAINDGRTALYDESCSSTLASTFWNGNEKRKNPE 364

PC8     RSIVTTDWLQKGTGCTEGHTGTSAAAPLAAGMIALMLQVRPCLTWRDVQHI_LVFTA--T 440
FURIN   KQIVTTDLR----QKCTEGHTGTSASAPLAAGIIALTLEANKNLTWRDMQHLVVQTSKPA 404
PACE4   KRIVTTDLR----QRCTDGHTGTSVSAPMVAGIIALALEANLNLTWRDMQHLVVKTSRPA 456
PC1     QRITSADLH----NDCTETHTGTSASAPLAAGIFALALEANPNLTWRDMQHLVVWTSEYD 418
PC2     S-------GNCTLRHSGTSAAAPEAAGVFALALEANLNLTWRDMQHLTVLTS-KRN 420

PC8     RYEDRRAEWVTNEAGFSHSHQHGFGLLNMAWRLVNAA--KIWTSVPYLASYVSPVLK-ENK 497
FURIN   HLNAN--DWATNGVGRKVSHSYGYGLLDAGAMVALA--QNWTTVAPQRKCIIDILT-EPK 459
PACE4   HLKAS--DWATNGVGHKVSHFYGFGLVDAEALVVEA--RKWTAVPSQHMCVAASDK-RPR 511
PC1     PLANN-DGWKKNGAGLWNVSRFGFGLLNAKALVDLADPRTWRSVPEKKKCVVKDNDFEPR 477
PC2     QLHDEVHQWRRNGVGLEFNHLFGYGVLDAGAMVKMA--KDWKTVPQRFHCVGGSVQ-DPR 477

PC8     AIPQSPRSLEVLWNVSRMDLEMSGLKTLEHVAVTVSITHPRRGSLELKLFCPSGMMSLIG 557
FURIN   DIGKRLEVRKT--VTACLGEPNH-ITRLEHAQARLTLSYNRRGDLAIHLVSPMGTRSTLL 516
PACE4   SIPLVQVLRTTALTSACAEHSDQRVVYLEHVVVRTSISHPRRGDLQIYLVSPSGTKSQLL 571
PC1     ALKANGEVIIEIPTRACEGQENH-AIKSLEHVQFEATIEYSRRGDLHITLTSAAGTSTVLL 536
PC2     KIPSTGKLVLTLTTDACEGKEN-FVRYLEHVQAVITVNYNRRGDLHITLTSPMGTKSILL 536

PC8     APRSM-DSDPNGFNDWTFSTVRCWGERARGTVRLVIRDV---GDRSFGVGTLKRWQLTLY 613
FURIN   AARPH-DYSAIGFNDWAFMTTHSWDEDPSGEWVLEIEN---TSEANNYGTLTKFTLVLY 571
PACE4   ACRLL-DLSNEGFTNWEFMTVHCWGEKAEGQWTLEIQDLPSQVRNPEKQGKLKEWSLILY 630
PC1     AERER-DTSPNGFKNWDFMSVHTWGENPIGTWTLRITDM--SGRIQNEGRIVNWKLILY 592
PC2     SRRPRDDDSKVGFDKWPFMTTHTWGEDARGTWTLEL-GF---VGSAPQKGVLKEWTLMLH 592

PC8     GSVWSAVDIRDRQRLLE--SAMSGKYLHDDFALPCPPGLKIPEEDGYTITPNTLKTLVLV 671
FURIN   GTA-------------------------------------PEGLPVPPESSGCKTLSSQACVVC 599
PACE4   GTAEHPYHTFSAHQSRSRMLELSAPELEPPKKAALSPSQVEVPEDEEDYTAQSTPGSANIL 690
PC1     GTSSQPEHMKQPRVYTSYNTVQNDRRGVEKHVDPRGEEQTQENPKENTLVSRSPSSSSVG 652
PC2     GTQSAPYIDQVVRDYQSKLAM------------------------------------------- 613

PC8     GCFTVFWTVYYMLEVYLSQRNVASNQVCRSGPCHWPHRSRKAKEEGTELESVPLCSSKDP 731
FURIN   E-------EGFSLHQKSCVQHCPPGFAPQVLDTHYSTENDVETIRASV------------- 640
PACE4   Q--------TSVC-HPECGDKCDGPNADQCLNCVHFSLGSVWKTSRKCVSVCPLGYFGDTAA 743
PC1     PLANN-RDELEEQAPQAMLRLLQSAFEKNSPPKQSPKKSPSAKLNIPYENFYEALEK 701
PC2     S-------KKEELEEEL-DEAVERSRLKTLL 638

PC8     DEVETESRGGPPTTEDLLAPDLLEQGDNSLSQNKSALDCPHQHLDVPHGKEKGIC 785
FURIN   --CAPCHAASCATCCGVPALPDCLSCP------------------SHASL--DPVEQTCSRQSQS 681
PACE4   RRCRRCHKGCETCSSNANTATQCLSCRRGFYHLNMTCVTLCPAGFYADESQKNCLKCHPS 803
PC1     LNKPSQLKDSEDSLYNDYVDVFYNTKPYKKHRDDRLLQALVDILNEEN 753

FURIN   SRESPPQQQPPRLP------------PEVEAGQRLRAGLLPSHLPEVVAGLLSCA------ 723
PACE4   CKKCVDEPEKCTVCKEGFSLARGSCIPDCEPGTYFDSELIRCGECHHTCGTCVGPGREEC 863

FURIN   ---------FIVLVFVTVFLVLQLRSGFSRGVKVYTMD------------RGLISYK 760
PACE4   IHCAKNFHFHDWKCVPACGEGFYPEEMPGLPHKVCRRCDENCLSCAGSSRNCSRCKTGFT 923

FURIN   GLPPEAWQEECPSDSEEDEGRRERTAFIKDQSAL 794
PACE4   QLGTSCITNHTCSNADETFCEMVKSNRLCERKLFIQFCCRTCLLAG 969
```

Figure 3   Complete amino acid sequence predicted for PC8 aligned to furin, PACE4, PC1 and PC2

The subtilisin-like catalytic domains are boxed and the catalytically important Asp, His, Asn (Asp), and Ser residues are indicated by ✱. Residues of identity with PC8 are indicated in **bold**, with gaps introduced to maximize the alignments. Residue numbers are indicated at right.

Figure 4   Expression of PC8 in human cell lines and adult rat tissues

(**A**) A 5 μg portion of polyadenylated RNA from cell lines were electrophoresed through 1.2%-agarose/formaldehyde gels, transferred to nylon and hybridized with ^{32}P-labelled HE6WCR40 (PC8). The blots were reprobed with a cDNA for glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as an indicator of RNA loading. Lanes contain polyadenylated RNA isolated from human cell lines: 1, BEN; 2, COLO 16; 3, HaCaT. The positions of the 4.5 kb and the 3.5 kb mRNA species are indicated on the left. (**B**) RT-PCR analysis of PC8 mRNA was conducted under saturating PCR conditions using the PC8-specific primers oPC33 and oPC35, which amplify a 121 bp fragment of PC8. Lanes contain RT-PCR products from total RNA isolated from the human cell lines (1, BEN; 2, COLO 16; 3, HaCaT) and from rat tissues (4, brain; 5, thymus; 6, skin; 7, lung; 8, kidney; 9, spleen; 10, liver; 11, stomach; 12, bladder; 13, heart; 14, skeletal muscle; 15, ovary; 16, testis). Lane 17 is a PCR negative control lane in which the RNA subjected to PCR had not been reverse-transcribed. Products were verified with a ^{32}P-labelled internal PC8-specific oligonucleotide, oPC34. Amplification of a 415 bp fragment of GAPDH cDNA was achieved with GAPDH-specific primers GAPDH-3 and GAPDH-4 (human) or GAPDH-5 and GAPDH-4 (rat), and products were verified with the internal detection oligonucleotide GAPDH-1 [33,34].

and oPC35 under saturating PCR conditions, and resultant products were verified by hybridization analysis with an internal oligonucleotide, oPC34 (Figure 4B). All tissues were found to express PC8 mRNA. The ability of the primer pairs to specifically amplify PC8 mRNA from human and rat cDNA suggests that the sequences are well conserved between the two species; the primers did not amplify cDNA for PC1, PC2, PACE4 or furin from either species (results not shown). Since PC8 is widely expressed in endocrine and non-endocrine tissues, it is likely that PC8 is an enzyme of the constitutive pathway of protein secretion and is unlike PC1 and PC2 which are involved in the regulated pathway and have a distribution restricted to endocrine and neuroendocrine cells.

From a comparative perspective, PC8 shares more features in common with furin and PACE4 than with PC1 and PC2. These include: (i) high amino acid identity over the entire molecule with furin and PACE4, whilst the similarity to PC1 and PC2 is lower; (ii) like furin and PACE4, PC8 contains a long (42-amino-acid) signal peptide, whereas the signal sequences of PC1 and PC2 are relatively short; (iii) in common with furin, PC8 possesses a putative transmembrane domain, suggesting that it may be localized to the Golgi; and (iv) the mRNA for PC8 is ubiquitously expressed like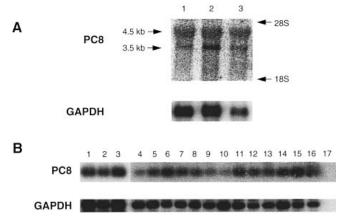 those for furin and PACE4, which is in marked contrast with the mRNAs for PC1 and PC2, which are restricted to endocrine and neuroendocrine cells. Combined, these data suggest that PC8 is more closely related to furin and PACE4 than to PC1 and PC2. As such, PC8, furin and PACE4 may have been derived from a common ancestral gene which has been duplicated on chromosome 15 for furin and PACE4 and on chromosome 11 for PC8. It remains to be established whether PC8 can fulfil the functions of furin, whether it does indeed act as a constitutive pathway enzyme and whether it has a substrate sequence specificity for cleavage of proteins similar to that of furin or other members of the convertase family.

REFERENCES

1   Barr, P. (1991) Cell **66**, 1–3
2   Seidah, N. and Chrétien, M. (1992) Trends Endocrinol. Metab. **3**, 133–140
3   Steiner, D. F., Smeekens, S. P., Ohagi S. and Chan, S. J. (1992) J. Biol. Chem. **267**, 23435–23438
4   Julius, D., Brake, A., Blair, L., Kunisawa, R. and Thorner, J. (1984) Cell **37**, 1075–1089
5   Fuller, R. S., Brake, A. and Thorner, J. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 1434–1438
6   Roebroek, A. J. M., Schalken, J. A., Leunissen, J. A. M., Onnekink, C., Bloemers, H. P. J. and Van de Ven, W. J. M. (1986) EMBO J. **5**, 2197–2202
7   van den Ouweland, A. M. W., van Duijnhoven, H. L. P., Keizer, G. D., Dorssers, C. J. and Van de Ven, W. J. M. (1990) Nucleic Acids Res. **18**, 664
8   Seidah, N. G., Gaspar, L., Mion, P., Marcinkiewicz, M., Mbikay, M. and Chrétien, M. (1990) DNA Cell Biol. **9**, 415–424

9   Smeekens, S. P., Avruch, A. S., LaMendola, J., Chan, S. J. and Steiner, D. F. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 340–344

10 Kiefer, M. C., Tucker, J. E., Joh, R., Landsberg, K. E., Saltman, D. and Barr, P. J. (1991) DNA Cell Biol. **10**, 757–769

11 Johnson, R. C., Darlington, D. N., Hand, T. A., Bloomquist, B. T. and Mains, R. E. (1994) Endocrinology (Baltimore) **135**, 1178–1185

12 Tsuji, A., Hine, C., Mori, K., Tamai, Y., Higashine, K., Nagamune, H. and Matsuda, Y. (1994) Biochem. Biophys. Res. Commun. **202**, 1452–1459

13 Smeekens, S. P. and Steiner, D. F. (1990) J. Biol. Chem. **265**, 2997–3000

14 Nakayama, K., Kim, W. S., Torii, S., Hosaka, M., Nakagawa, T., Ikemizu, J., Baba, T. and Murakami, K. (1992) J. Biol. Chem. **267**, 5897–5900

15 Seidah, N. G., Day, R., Hamelin, J., Gaspar, A., Collard, M. W. and Chrétien, M. (1992) Mol. Endocrinol. **6**, 1559–1570

16 Seidah, N. G., Day, R., Marcinkiewicz, M., Benjannet, S. and Chrétien M. (1991) Enzyme **45**, 271–284

17 Lusson, J., Vieau, D., Hamelin, J., Day, R., Chrétien, M. and Seidah, N. G. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 6691–6695

18 Nakagawa, T., Hosaka, M., Torii, S., Watanabe, T., Murakami, K. and Nakayama, K. (1993) J. Biochem (Tokyo). **113**, 132–135

19 Wise, R. J., Barr, P. J., Wong, P. A., Kiefer, M. C., Brake, A. J. and Kaufman, R. J. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 9378–9383

20 Seidah, N. G., Marcinkiewicz, M., Benjannet, S., Gaspar, L., Beaubien, G., Mattei, M. G., Lazure, C., Mbikay, M. and Chrétien, M. (1991) Mol. Endocrinol. **5**, 111–122

21 Joshi, D., Miller, M. M., Seidah, N. G. and Day, R. (1995) Endocrinology (Baltimore) **136**, 2721–2729

22 Scopsi, L., Gullo, M., Rilke, F., Martin, S. and Steiner, D. F. (1995) J. Clin. Endocrinol. Metab. **80**, 294–301

23 Bresnahan, P. A., Leduc, R., Thomas, L., Thorner, J., Gibson, H. L., Brake, A. J. and Thomas, G. (1990) J. Cell. Biol. **111**, 2851–2859

24 Benjannet, S., Rondeau, N., Day, R., Chretien, M. and Seidah, N. G. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 3564–3568

25 Nakayama, K., Watanabe, T., Nakagawa, T., Kim, W.-S., Nagahama, M., Hosaka, M., Hatsuzawa, K., Kondoh-Hashiba, K. and Murakami, K. (1992) J. Biol. Chem. **267**, 16335–16340

26 Bennet, D. L., Bailyes, E. M., Nielsen, E., Guest, P. C., Rutherford, N. G., Arden, S.D. and Hutton, J. C. (1992) J. Biol. Chem. **267**, 15229–15236

27 Bravo, D. A., Gleason, J. B., Sanchez, R. I., Roth, R. A. and Fuller, R. S. (1994) J. Biol. Chem. **269**, 25830–25837

28 Marriott, D., Gillece-Castro, B. and Gorman, C. M. (1992) Mol. Endocrinol. **6**, 1441–1450

29 Brennan, S. O. and Nakayama, K. (1994) FEBS Lett. **338**, 147–152

30 Diefenbach-Jagger, H., Brenner, C., Kemp, B. E., Baron, W., McLean, J., Martin, T. J. and Moseley, J. M. (1995) Eur. J. Biochem. **229**, 91–98

31 Moseley, J. M. and Gillespie, M. T. (1995) Crit. Rev. Clin. Lab. Sciences **32**, 299–343

32 Suva, L. J., Winslow, G. A., Wettenhall, R. E. H., Hammonds, R. G., Moseley, J. M., Diefenbach-Jagger, H., Rodda, C. P., Kemp, B. E., Rodriguez, H., Chen, E. Y., Hudson, P. J., Martin, T. J. and Wood, W. I. (1989) Science **237**, 893–896

33 Glatz, J. A., Heath, J. K., Southby, J., O'Keeffe, L. M., Kiriyama, T., Moseley, J. M., Martin, T. J. and Gillespie, M. T. (1994) Mol. Cell. Endocrinol. **101**, 295–306

34 Heath, J. K., Southby, J., Fukumoto, S., O'Keeffe, L. M., Martin, T. J. and Gillespie, M. T. (1995) Biochem. J. **307**, 159–167

35 Traianedes, K., Findlay, D. M., Martin, T. J. and Gillespie, M. T. (1995) J Biol. Chem. **270**, 20891–20894

36 Chomcznski, P. and Sacchi, N. (1987) Anal. Biochem. **162**, 156–159

37 Taviaux, S. A. and Demaille, J. G. (1993) Genomics **15**, 194–196

38 Bruzzaniti, A., Goodge, K., Martin, T. J., Moseley, J. M. and Gillespie, M. T. (1995) Bone **16** (suppl. 1), 126S #165

39 Gillespie, M. T., Bruzzaniti, A., Goodge, K., Moseley, J. M. and Martin, T. J. (1995) J. Cell. Biochem. **19B**, 243 #B7–102

40 Kozak, M. (1986) Cell **44**, 283–292