Gene organization and alternative splicing of human prohormone convertase PC8

Katrina A. GOODGE, Rachel J. THOMAS, T. John MARTIN and Matthew T. GILLESPIE[1]

Department of Medicine, The University of Melbourne and St. Vincent's Institute of Medical Research, 41 Victoria Parade, Fitzroy, Victoria, Australia, 3065

The mammalian Ca^{2+}-dependent serine protease prohormone convertase PC8 is expressed ubiquitously, being transcribed as 3.5, 4.3 and 6.0 kb mRNA isoforms in various tissues. To determine the origin of these various mRNA isoforms we report the characterization of the human *PC8* gene, which has been previously localized to chromosome 11q23–24. Consisting of 16 exons, the human *PC8* gene spans approx. 27 kb. A comparison of the position of intron–exon junctions of the human *PC8* gene with the gene structures of previously reported prohormone convertase genes demonstrated a divergence of the human *PC8* from the highly conserved nature of the gene organization of this enzyme family. The nucleotide sequence of the 5′-flanking region of the human *PC8* is reported and possesses putative promoter elements characteristic of a GC-rich promoter. Further supporting the potential role of a GC-rich promoter element, multiple transcriptional initiation sites within a 200 bp region were demonstrated. We propose that the various mRNA isoforms of PC8 result from the inclusion of intronic sequences within transcripts.

INTRODUCTION

Specific endoproteolytic processing of precursors at dibasic or multibasic amino acids is an essential event in the maturation and activation of many biologically active proteins and prohormones. One family of enzymes responsible for these endoproteolytic processing events is the prohormone convertase family. Prohormone convertases are Ca^{2+}-dependent serine proteinases with catalytic domains that exhibit sequence similarity to the bacterial subtilisins.

Yeast convertase Kex2 was the first convertase to be identified [1,2] and subsequently seven mammalian members of this enzyme family have been identified. The mammalian members of the prohormone convertase family include furin, also known as PACE [2–6]; PC2 [7,8]; PC1, also identified as PC3 [8–11]; PC4 [12,13]; PC5, also known as PC6 [14–16]; PACE4 [17], also identified as PC7 [18]; and most recently PC8, also referred to as LPC, rPC7 and SPC7 [19–22]. Furin, PACE4, PC5/6 and PC8 are expressed in a wide variety of cell lines and tissues, although variation in the relative abundance of these enzymes is observed [6,14,15,17,19,21,23]. In comparison, PC1 and PC2 have a more restricted distribution, primarily in cells of endocrine and neuroendocrine origin [7,8–11], and PC4 is restricted to the testis [12,13].

Previously we reported the identification of the seventh member of the prohormone convertase family, PC8. Northern hybridization analysis performed on poly(A)$^+$ RNA from a variety of rat tissues and human cell lines demonstrated at least two transcripts for PC8: a 3.5 kb transcript and a 4.3 kb transcript [19]; evidence for a larger transcript of approx. 6 kb in several rat tissues has been reported [21]. Despite the 4.3 kb transcript's seeming to be the most abundant mRNA species, all cDNA clones described for human, rat or mouse PC8 correspond to the 3.5 kb transcript. Explanations to account for the different-sized transcripts of PC8 include alternative polyadenylation, alternate promoter usage or alternative splicing, and such mechanisms have been described for prohormone convertase family members. To elucidate the nature of the different transcripts of PC8, and to gain insights to the potential evolution of the prohormone convertase family, we have investigated the structural organization of the human *PC8* gene.

In the present paper we have reported the gene structure and the nucleotide sequence of the 5′-flanking region of the human *PC8* gene. We propose that the 4.3 kb mRNA isoform of PC8, at least in the human gene, arises from an alternative splicing event that involves the incorporation of approx. 850 bp of intronic sequence, an event that disrupts the middle domain of human PC8. On the basis of the divergence of the gene structure of the human *PC8* gene from the highly conserved nature of the gene organization of the prohormone convertase family members, it has been proposed that PC8 is the most evolutionarily distant member of the prohormone convertase family.

MATERIALS AND METHODS

Cell culture

The human squamous cancer cell line BEN and the spontaneously immortalized human keratinocyte cell line HaCaT were cultured as described previously [24,25].

Isolation of the human (h)*PC8* gene

Sense strand oligonucleotide oPC72 (5′-GCCTTTGCTGAGTCAACACTAC-3′), corresponding to nt +54 to +75 of human PC8 mRNA, and anti-sense strand oligonucleotide oPC71 (5′-AGGGAACCAGTAAGAAGAGCC-3′), corresponding to nt +258 to +278 of human PC8 mRNA [19], were used to screen a commercial human P1 genomic library for the *hPC8* gene. Three independent P1 clones were obtained from Genome Systems. All three clones hybridized to oligonucleotide probes based on the 5′ untranslated region (UTR) and 3′ UTR of

Abbreviations used: EST, expressed sequence tag; h, human; m, mouse; RT–PCR, reverse transcriptase-mediated PCR; UTR, untranslated region; RACE, rapid amplification of cDNA ends.

[1] To whom correspondence should be addressed (e-mail m.gillespie@medicine.unimelb.edu.au).

human PC8 and therefore encompassed the entire coding region of the *hPC8* gene.

Intron–exon splice junctions of the *hPC8* gene

Intron–exon junctions were characterized by direct cycle sequencing (Life Technologies, Gaithersburg, MD, U.S.A.) with [γ-^{32}P]dATP end-labelled sequencing primers based on exonic and intronic sequences. Oligonucleotide sequences used in this direct sequencing strategy are available on request. Comparison of the products resulting from PCR amplifications with oligonucleotides based on flanking exonic sequences on genomic and cDNA templates was used to estimate the sizes of identified introns.

The nucleotide sequence of the 5′-flanking region of the human *PC8* gene was determined directly by double-stranded DNA cycle sequencing (Life Technologies) of the P1 genomic clones by using [γ-^{32}P]dATP end-labelled anti-sense strand oligonucleotides oPC80 (5′-CTCCCTGCGGAACTCG-3′), complementary to nt +6 to +22 of human *PC8* [19], and oPC114 (5′-CATGGTCCCTCTTCAGAAG-3′), corresponding to nt −226 to −208 of the human *PC8* gene (see Figure 3). To overcome the effects of the GC-rich nature of the 5′-flanking region, DMSO was added to the double-stranded DNA cycle sequencing reactions at a final concentration of 5 mM.

Rapid amplification of cDNA ends (RACE)

5′ RACE was performed to obtain *PC8* cDNA with extended 5′ ends by using the 5′ RACE kit (Gibco BRL) as recommended. Poly(A$^+$) RNA (1 μg) from the BEN cell line was reverse-transcribed with SuperScript II reverse transcriptase by using PC8-specific anti-sense strand oligonucleotide oPC34 (5′-TACCCGGATACCTGCGAT-3′), complementary to nt +941 to +958 of human *PC8* [19]. Resultant cDNA transcripts were C-tailed with terminal dideoxytransferase (TdT) and PCR amplification was performed with the supplied sense strand 5′ RACE adapter primer (Life Technologies) and anti-sense strand PC8-specific oligonucleotide oPC79 (5′-CTGGATTCTACATGAGGTTGAG-3′), corresponding to nt +145 to +166 of human PC8 mRNA (Figure 3). 5′ RACE products were cloned into pGEM®-T vector (Promega, Madison, WI, U.S.A.), and clones were confirmed by nucleic acid sequencing.

Primer extension analysis

Primer extension analysis was performed with standard methods [26]. Anti-sense PC8-specific oligonucleotide primer oPC79 was end-labelled with [γ-^{32}P]dATP by using T4 polynucleotide kinase (Boehringer Mannheim, Mannheim, Germany), annealed to 40 μg HaCaT cell line total RNA and extended with AMV reverse transcriptase (Promega). Extension products were analysed by electrophoresis on a 6% (w/v) polyacrylamide/8 M urea sequencing gel and detected by autoradiography at −70 °C for 5 days.

Detection of alternative splicing of the hPC8 gene

Total RNA was extracted with guanidinium thiocyanate as described previously [27]. Reverse transcription was performed on 5 μg of total RNA isolated from the BEN cell line, with $(dT)_{15}$ primer (Promega) and AMV reverse transcriptase (Promega), in accordance with the manufacturer's instructions. The inclusion of intronic sequences was confirmed by reverse transcriptase-mediated PCR (RT–PCR), with products resolved in a 2% (w/v) agarose gel then transferred to nylon membranes and authenticated by hybridization with [γ-^{32}P]dATP end-labelled oligonucleotide probes or by nucleic acid sequencing as described previously [19]. Oligonucleotides used for PCR and/or product detection were oPC54 (5′-CTCACCCTCTTTGAGTGGAGTCAG-3′, anti-sense strand oligonucleotide complementary to nt +2639 to +2662 of human PC8 mRNA [19]), oPC59 (5′-GTCATCCCTGATGACAAGCCT-3′, sense strand oligonucleotide complementary to nt +1943 to +1963 of human PC8 mRNA [19]), oPC67 (5′-AAAGTGTTGCTGGGGGAAGG-3′, sense strand oligonucleotide based on intron 12 of the human *PC8* gene), oPC99 (5′-TCAGCAGGATGGACCTGGAGATG-3′, sense strand oligonucleotide complementary to nt +1713 to +1735 of human PC8 mRNA [19]) and oPC118 (5′-CCTTCCCCCAGCAACACTTTC-3′, anti-sense strand oligonucleotide based on intron 12 of the human *PC8* gene). PCR amplification was performed for 40 cycles at an annealing temperature of 55 °C and extension for 3 min at 72 °C.

RESULTS AND DISCUSSION

Structure of the *hPC8* gene

A map of the *hPC8* gene was constructed by direct cycle nucleotide sequence analysis of three independent P1 genomic clones. The exonic regions of *hPC8* were fully sequenced to determine the exact locations of the intron–exon junctions and are shown in Table 1. Sequences of all intron–exon splice junctions were observed to conform to the GT/AG rule [28]. By using PCR, all intervening intron sequences were sized, with the exception of the intron 11, whose size was determined from the reported nucleotide sequence of cosmid clone HSU73638 from the GenBank expressed sequence tag (EST) database, and introns 12 and 13, which were sequenced. The human *PC8* gene was shown to span approx. 27 kb and consisted of 16 exons. A schematic representation of the human *PC8* gene is presented in Figure 1. Introns ranged in size from approx. 250 nt to 9412 nt. Exonic sequences ranged in size from approx. 46 nt to 1272 nt. Two introns were observed in the 5′ UTR of human PC8. As in other prohormone convertase genes, distinct structural features of *PC8* were encoded by separate exons. The catalytic domain of PC8 was encoded by exons 3–10 and the four catalytically important amino acids Asp^{187}, His^{228}, Asn^{329} and Ser^{406} are encoded by distinct exons, exons 5, 6, 8 and 10 respectively. The signal peptide is encoded by exon 3, the RGS motif by exon 12, and the transmembrane domain by exon 16.

Previous reports of the structural organization of the genes of mammalian prohormone convertases have demonstrated a high degree of conservation in the gene organization of this enzyme family [6,29–33]. A comparison of intron–exon splice junctions of *hPC8* with those of *hPC1*, *hPC2*, *hPACE*, mouse (*m*)*PC4* and human *furin* genes illustrated a distinctive divergence of the gene organization of human *PC8* from the striking conservation of splice boundaries observed in the gene organizations of other convertase family members. Figure 2 illustrates the divergence in conservation of the splice junctions of the *hPC8* gene with respect to the position of the splice junctions of *hPC1*, *hPC2*, *hPACE4*, *mPC4* and human *furin* genes. Divergence of the *hPC8* gene might implicate *PC8* as an ancestral gene of the mammalian prohormone convertase enzyme family.

The genes of *hfurin*, *hPC1/3*, *hPC2*, *hPACE4* and *mPC4* share nine conserved intron–exon splice boundaries, whereas *hPC8* shares only four of these conserved sites, indicating that *hPC8* differs substantially from the other family members. Further, two conserved splice boundaries are observed within the region encoding the prodomain of the human *furin*, *PC1*, *PC2*,

Table 1 Intron–exon splice junctions of the *hPC8* gene

Nucleotide sequences at the intron (lower-case letters) and exon (upper-case letters) boundaries are presented. Exons are numbered from the 5′ end of the hPC8 gene. Exon and intron sizes were determined either by PCR or nucleic acid sequencing (introns 12 and 13) or were derived from the EST database (intron 11). The amino acids interrupted are indicated in the standard single-letter code; n.a. indicates not applicable.

Exon	Exon size (bp)	5′ splice junctions		Intron size (bp)	3′ splice junctions		Residue(s) interrupted
1	46	GGGGCCG	gtgagttgggaactaggtgggggctggac	1750	cattatttctctgttcttggatcccacag GTCTGTA		n.a.
2	119	AGAATCCAG	gtgggcacgggggcctgttatgggaattag	500	ctcttccacatccttttttctttgtctgcag TCCACTGCT		n.a.
3	480	TGG CAC	CTG gtgagtccagtcctgcagtttgcagtttca	1050	tgtattcccatcctgtcctcgtcttacag AAT AAC	CGA	L^{156}/N^{157}
4	134	CCC AAC	TAT gtgagtggaccctggaacgacaccctcac	800	tgaccacatgtcacctccaaaatatttcag AGC CCT	GAG	Y^{201}/S^{202}
5	165	ATC GCA	G gtaacttgtgcctccagcctagccttggta	1250	gaggaggctgctgtggccttttattgcctag GT ATC	CGG	G^{257}
6	90	AGC TGC	AG gtgaggcagagctgcagatgagaggcaag	1250	tcccctctagggtatttacttttctcaacag C TGG	GGA	S^{287}
7	54	CTT GGA	AAG gtaactggcctgcacaaagactctgagaaa	500	aactgacaccctgatgttttggctcctcag GCT GCC	TTA	K^{305}/A^{306}
8	138	ACC ATA	G gtagtcaccaatgtttctgtgtcctgtg	800	ctcagctctaatcgccctcctctgtcctag GA GCT	GTG	G^{352}
9	104	CGG AGC	ATT gtgagtgcctccgggccctccacggcctg	3000	gcctacttcctttttcctgtttcacttccag GTG ACC	ACT	I^{385}/V^{386}
10	271	GCA GCC	AAG gtgaacaggtgttggcagggtgccctgtgt	450	gatttgccatcccctcctttcctttcatag ATC TGG	ACA	K^{477}/I^{478}
11	102	TGG AAT	G gtaagtaatcagcacaaaggggttaatct	9412	ttcccagctgcccagcctgactttctacag TC AGC	AGG	V^{512}
12	156	ATG GAC	TC gtatgtacaccgccgaggctgagcaacc	833	ccacctgtttctgctgtgttctttcctcag G GAT	CCC	S^{564}
13	94	GAT GTC	G gtgagtctcgtgccctcacacaggccctа	809	tggctcacgagctgccccttgctgccctag GG GAT	GAG	G^{596}
14	95	AGA CAA	AG gtgggtaaagccgcatgtgtcagagggaag	250	gggcatcctcacctctttcttctccacag G CTG	TTA	R^{627}
15	117	ACC CTC	AAG gtactagggccaagactcaagctgcgggta	350	gcctcacgtaggtgcttgcctcggccccag ACC CTG	GTG	K^{666}/T^{667}
16	1272						

PACE4 and mouse *PC4* genes, whereas these splice boundaries are absent from the *hPC8* gene. Thus the only conserved intron–exon boundaries are within the catalytic domain of the prohormone convertase family. Conservation of intron–exon splice boundaries is not apparent C-terminal to the catalytic domain.

Concurrent with our laboratory's identification of human *PC8*, Meerabux et al. [20] identified in high-grade lymphomas an identical transcript due to the close proximity of the human *PC8* gene to a chromosomal breakpoint. With the use of the experimental technique of exon trapping, Meerabux et al. reported the presence of an intron in the 3′ UTR of the human *PC8* gene. We have been unable to demonstrate such an intron–exon boundary by direct cycle sequencing techniques or by PCR amplification across the region of interest, in the three P1 clones and in genomic DNA prepared from a variety of human cell lines. We propose that the report of an intron in this region by Meerabux et al. [20] was either a result of a cryptic splice site within this region or translocation events or viral integration of a foreign sequence unique to the particular genomic clone isolated. Such integration events have been documented; in particular the viral integration of the HIV sequence has been found within the 3′ UTR of furin in HIV-associated lymphoma [34].

At the 5′ end of the *hPC8* gene

Nucleotide sequence analysis of the 5′-flanking region of the human *PC8* gene was obtained by direct cycle sequencing of *hPC8* genomic P1 clones. The sequence of the 5′-flanking region of *hPC8* demonstrated an absence of TATA or CAAT promoter elements [35]. However, this region was highly GC-rich, the first 200 nt displaying a G+C content of 76 % (Figure 3). Within this region there is a transcription factor Sp1 consensus binding site and several AP2 consensus binding elements [36–39]. Putative promoter elements identified in the investigation of the 5′-flanking region of the human *PC8* gene have been implicated in the transcriptional regulation of housekeeping genes.

5′ RACE from RNA extracted from the human cell line BEN revealed multiple start sites ranging from 68 nt to 178 nt upstream of the translational start site (Figure 3, +1 to +111). Some of these sites were confirmed by primer extension analysis on HaCaT total RNA as well as by the identification of others (Figure 4).

The GC-rich nature of the 5′-flanking region of the *hPC8* gene combined with the multiple transcriptional initiation sites is characteristic of the promoter regions of many housekeeping genes. In agreement with this proposition was the ubiquitous nature of the mRNA tissue distribution for PC8 [19,21], although the varying mRNA expression in certain tissues suggests that some cells might have altered transcriptional activation of *PC8*, or might confer altered mRNA stability for PC8. The absence of classical TATA or CAAT promoter elements is characteristic of many promoter regions reported for other prohormone convertase genes, including *PC2*, *PC4* and the P1A and P1B promoter region of *furin* [29,32,40].

Alternative splicing of hPC8

Evidence for alternative splicing of PC8 mRNA, transcriptional initiation or polyadenylation comes from results of Northern analyses that have detected transcripts of 3.5, 4.3 and 6.0 kb [19,21]; only the 3.5 kb transcript has been cloned and characterized. To elucidate potential mechanisms for the observation of the additional observed transcripts, 5′ RACE and 3′ RACE analysis were performed. These experimental approaches failed

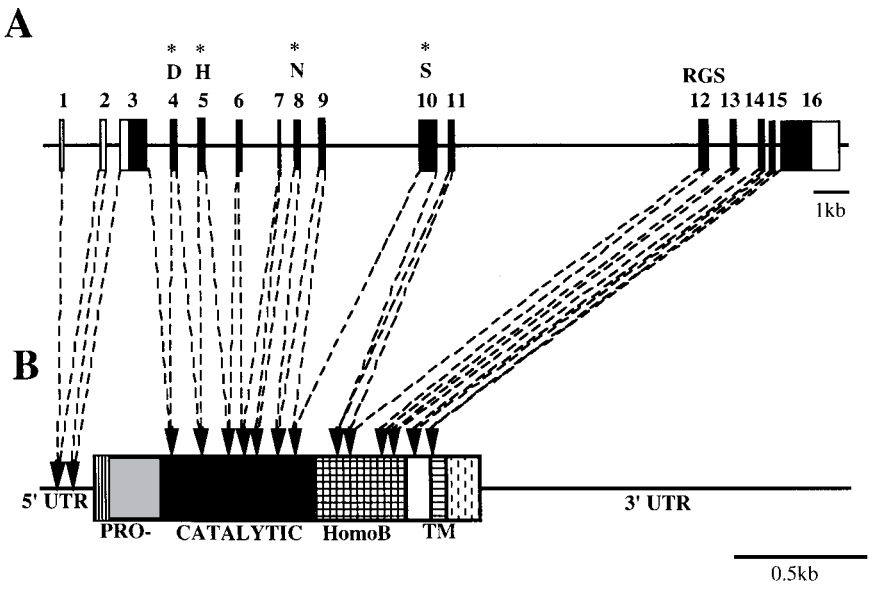

Figure 1 Gene organization of the *hPC8* gene

(**A**) Schematic representation of the intron–exon splice junctions of the *hPC8* gene. The 16 exons are indicated as boxes. Coding regions are shown in black except that the transmembrane domain is indicated by vertical lines, and the 5′ and 3′ UTRs are indicated by open boxes. Intronic sequences are represented by a solid horizontal line between the exons. The relative positions of the catalytically important residues are indicated by asterisks above the relevant exon. (**B**) Schematic representation of the structural domains of hPC8. The boxes indicate translated sequences and the horizontal lines represent untranslated sequences. The pro-domain, catalytic domain, homo B domain and transmembrane domain are indicated. The relative positions of the splicing boundaries are indicated by linking arrows above the diagram.

to identify cDNA clones encoding additional nucleotide sequences of the 4.3 or 6.0 kb transcripts, suggesting that alternative transcriptional initiation sites or polyadenylation did not account for these transcripts. Searches of the EST database for the full-length cDNA sequence, however, indicated a number of clones with disrupted PC8 cDNA sequences. The positions at which they occurred were at intron–exon boundaries and were found to comprise intronic sequences suggesting read-through introns. These read-through introns corresponded to introns 12 (EST H09374), 13 (EST AA127764, AA411647, A373787), 14 (EST AA127764, AA411647, A373787, AA373478, AA317286, AA193394) and 15 (EST AA317286, AA193394, T98068). The presence of intronic sequences in mRNA transcripts, however, might have resulted from prespliced mRNA intermediates. However, one identified EST clone, H09374, was unlikely to represent a prespliced mRNA intermediate as this cDNA clone possessed only the intron 12 intronic sequence, although it encompassed sequences encoded by exons 13–16.

We proposed that the nucleotide sequence of intron 12 was retained in human PC8 mRNA transcripts and that its retention resulted in the 4.3 kb transcript that we have detected in squamous cancer cell lines [19]. By using an RT–PCR strategy, we confirmed that intron 12 was transcribed in BEN cell mRNA species (Figure 5). RT–PCR of BEN cell RNA with PC8 exon 12 sense and exon 13 anti-sense primers (oPC99 and oPC59 respectively) amplified products of 229 and 1062 bp, and both products were detected by the exon 13-specific anti-sense oligonucleotide oPC120 (Figure 5, lane 1). The 1062 bp product resulted from amplification of exon 12, intron 12 and exon 13 sequences as confirmed by hybridization analysis with an intron 12 oligonucleotide oPC67 (Figure 5, lane 2), and nucleic acid sequencing confirmed the hybridization data for both the 229 and 1062 bp products. To determine whether the retention of intron 12 in PC8 transcripts was an exclusive event, in that no other intronic sequences were present, RT–PCR was performed with intron 12 primers. RT–PCR amplifications between exon 1 (oPC81) and intron 12 (oPC118) yielded one product of 2644 bp that was PC8-specific as detected by the oligonucleotide oPC99 (Figure 5, lane 3). Further, amplifications with oPC67 (intron 12) and oPC54 (exon 16, the ultimate exon) generated only one product of 865 bp; a PC8-specific oligonucleotide, oPC59, hybridized with this product (Figure 5, lane 4). Taken together, these results indicate that intron 12 is included in PC8 transcripts and its inclusion would account for the 4.3 kb transcript of PC8. Importantly, the RT–PCR analyses did not detect the inclusion of any other intronic sequences, suggesting that the retention of intron 12 was not the result of a prespliced intermediate. Nucleotide sequence analysis of intron 12 predicted that the inclusion of this intron would result in the production of an in-frame termination codon seven residues from the intron–exon boundary, thus truncating the translation of human PC8, and disrupting the catalytically important middle domain of PC8. This alternative splicing event might produce an inactive form of PC8 because the catalytically important homo B domain would be disrupted.

The occurrence of cDNA clones containing nucleotide sequences from consecutive introns (introns 12, 13, 14 and 15) could potentially result in the larger transcript of 6.0 kb as reported by Seidah et al. [21]. The inclusion of these intronic sequences in PC8 transcripts has the potential for both catalytic and structural effects. The read-through of intron 13, in addition to the disruption of the catalytically important homo B domain of PC8, would result in disruption of the open reading frame of hPC8 and thus might result in the translation of an hPC8 isoform lacking the putative transmembrane domain. The loss of the putative transmembrane domain of human PC8, if read-through of only intron 15 occurred, might result in a soluble form of the enzyme. However, it should be noted that we did not detect

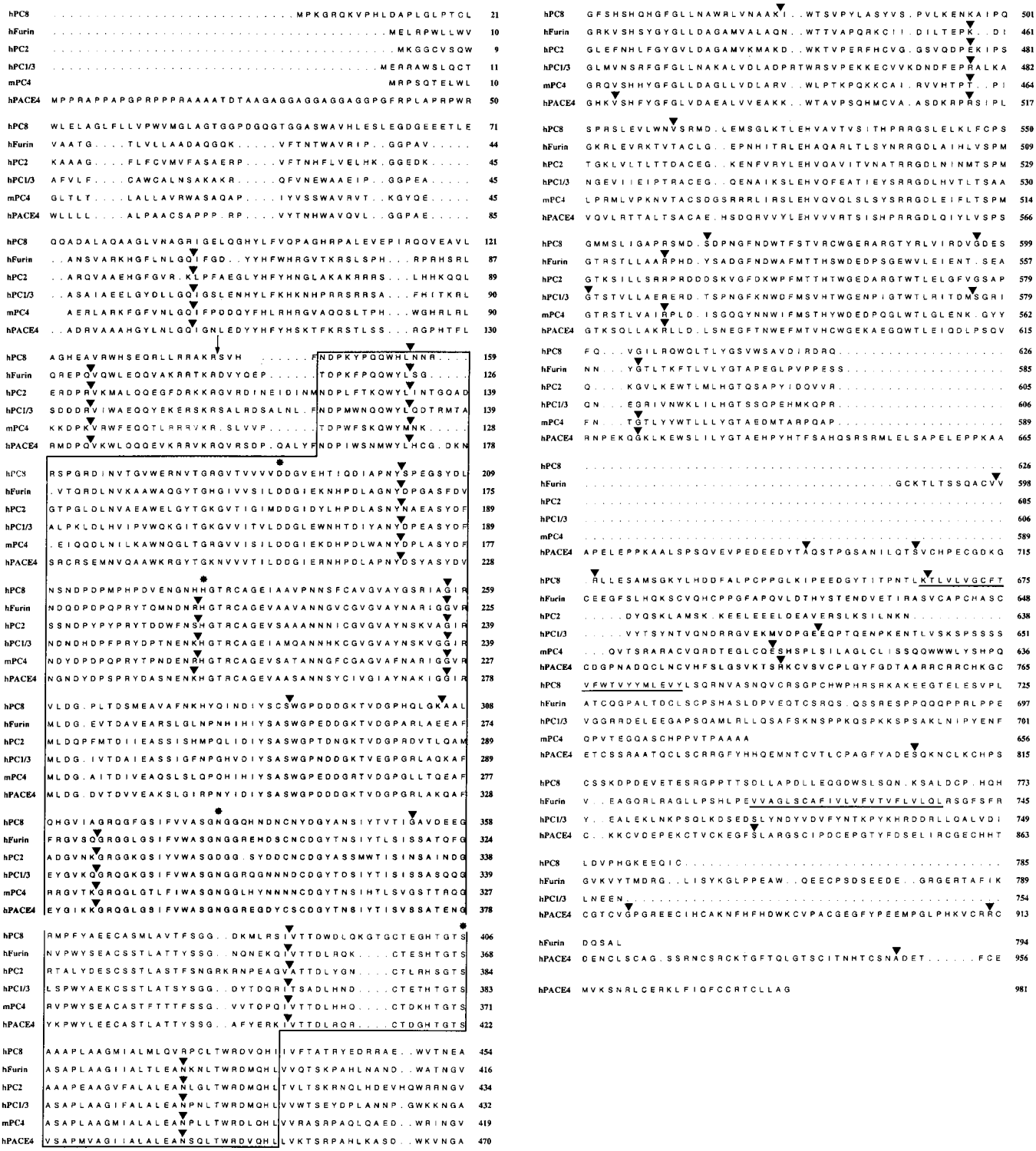

Figure 2 Comparison of the intron–exon splice junctions of *PC8* with those of other convertase genes

The figure is a schematic representation of the positions of intron–exon splice boundaries with respect to the predicted amino acid sequences of five human prohormone convertase genes *PC8*, *furin*, *PACE4*, *PC3/1* and *PC2* and the mouse *PC4* gene (in the single-letter amino acid code). The positions of the intron–exon splice junctions are indicated above the predicted amino acid sequence by a solid arrowhead. The catalytic domain of each family member is boxed and the four catalytically important residues of this domain are indicated by asterisks. The transmembrane domains of PC8 and furin are underlined.

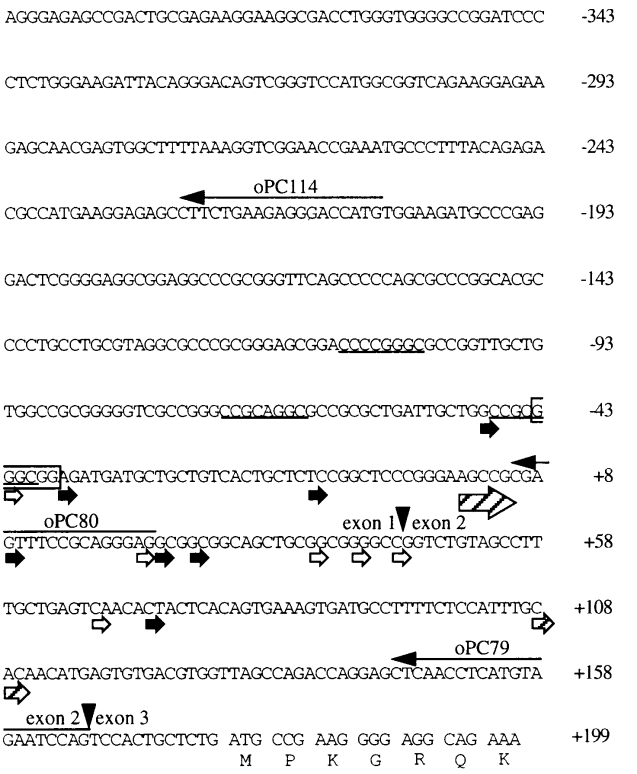

Figure 3 Nucleotide sequence of the 5′-flanking sequence of the *hPC8* gene

The nucleotide sequence of the 5′-flanking region of the *hPC8* gene, determined by direct cycle sequencing, is presented. Numbered in relation to the most extended 5′ RACE product identified, the nucleotide sequence is numbered at the right of the sequence. Transcriptional initiation sites determined by 5′ RACE (hatched arrows) and primer extension analyses are indicated by arrows below the nucleotide sequence, filled and unfilled arrows representing major and minor transcriptional initiation sites respectively. The large hatched arrow below the sequence represents the coincident transcriptional initiation sites determined by 5′ RACE and primer extension and is designated nt +1. The positions of intron–exon junctions are marked by vertical arrowheads. Oligonucleotides used in the direct determination of the nucleotide sequence, with the use of the DNA cycle sequencing technique, are indicated. The predicted amino acid sequence of the translated region is provided in the standard single-letter amino acid code. Putative promoter elements present in the nucleotide sequence of the 5′-flanking region of the human *PC8* gene are indicated as follows: Sp1 promoter elements are boxed and AP2 promoter elements are indicated by a solid underline.

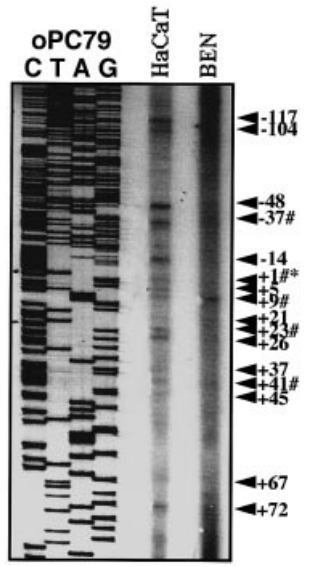

Figure 4 Primer extension analysis of the transcriptional initiation of human *PC8*

Primer extension analysis was performed on 40 μg of total RNA from the human keratinocyte cell line HaCaT [24], with [γ-^{32}P]dATP end-labelled anti-sense oligonucleotide oPC79 based on nt +145 to +166 of the reported cDNA sequence of human PC8 (see Figure 3). The sizes of extension products were determined by using a sequencing ladder generated with anti-sense oligonucleotide oPC79. Hash signs (#) indicate transcriptional initiation sites shared by the HaCaT and BEN cell lines, whereas the asterisk denotes the 5′ extension product resulting from 5′ RACE.

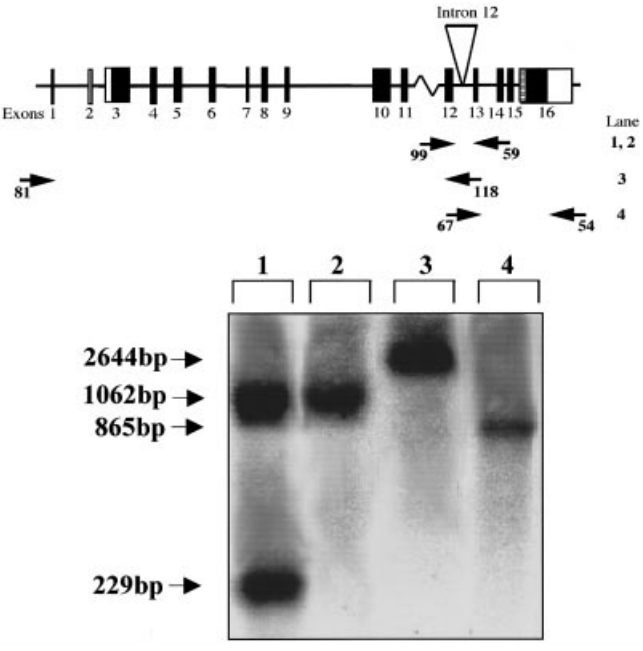

Figure 5 Identification of intron 12 sequences in *PC8*

Below the schematic representation of the intron–exon splice junctions of the *hPC8* gene are the positions of the oligonucleotides used in RT–PCR analysis: shown at the right are the lanes corresponding to the amplifications detected by Southern blotting. Shown is a Southern blot of RT–PCR with oPC99 (exon 12) and oPC59 (exon 13) detected with oPC120 (lane 1) and oPC67 (lane 2); oPC81 (exon 1) and oPC118 (intron 12) detected with oPC99 (lane 3); oPC67 (intron 12) and oPC54 (exon 16) detected with oPC59.

transcripts larger than 4.3 kb in the human cell lines that we have examined [19], and that we were unable to identify the inclusion of introns other than intron 12 in PC8 transcripts. The transcripts larger than 4.3 kb [21] might result from tissue/cell-specific alternative mRNA splicing or prespliced intermediates.

In conclusion, we report the characterization of the human *PC8* gene structure. The nucleotide sequence of the 5′-flanking region of the human *PC8* gene was determined and displays a number of promoter elements characteristic of GC-rich promoters. We propose that the human *PC8* gene undergoes alternative splicing events to produce various structurally unique isoforms via the inclusion of intronic sequences in human PC8 transcripts.

This work was supported by a Program Grant from the NHMRC, Australia (no. 963211) and grants from the Anti-Cancer Council of Victoria and Chugai Pharmaceutical Co. Japan. M.T.G. was a Research Fellow of the NHMRC, Australia.

REFERENCES

1 Fuller, R. S., Brake, A. and Thorner, J. (1989) Proc. Natl. Acad. Sci U.S.A. **86**, 1434–1438
2 Fuller, R. S., Brake, A. J. and Thorner, J. (1989) Science **426**, 482–486
3 Roebroek, A. J., Schalken, J. A., Leunissen, J. A., Onnekink, C., Bloemers, H. P. and Van de Ven, W. J. M. (1986) EMBO J. **5**, 2197–2202
4 van den Ouweland, A. M. W., van Duijnhoven, H. L. P., Keizer, G. D., Dorsers, L. C. J. and Van de Ven, W. J. M. (1990) Nucleic Acids Res. **18**, 664
5 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**, 9378–9382
6 Barr, P. J., Mason, O. B., Landsberg, K. E., Wong, P. A., Kiefer, M. C. and Brake, A. J. (1991) DNA Cell Biol. **10**, 319–328
7 Smeekens, S. P. and Steiner, D. F. (1990) J. Biol. Chem. **265**, 2997–3000
8 Seidah, N. G., Gaspar, L., Mion, P., Marcinkiewicz, M., Mbikay, M. and Chretien, M. (1990) DNA Cell Biol. **9**, 415–424
9 Seidah, N. G., Marcinkiewicz, M., Benjannet, S., Gaspar, L., Beaubien, G., Mattei, M. G., Lazure, C., Mbikay, M. and Chretien, M. (1991) Mol. Endocrinol. **5**, 111–122
10 Nakayama, K., Hosaka, M., Hatsuzawa, K. and Murakami, K. (1991) J. Biochem. (Tokyo) **109**, 803–806
11 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
12 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
13 Seidah, N. G., Day, R., Hamelin, J., Gaspar, A., Collard, M. W. and Chretien, M. (1992) Mol. Endocrinol. **6**, 1559–1570
14 Nakagawa, T., Hosaka, M., Torii, S., Watanabe, T., Murakami, K. and Nakayama, K. (1993) J. Biochem. (Tokyo) **113**, 132–135
15 Lusson, J., Vieau, D., Hamelin, J., Day, R., Chretien, M. and Seidah, N. G. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 6691–6695
16 Miranda, L., Wolf, J., Pichuantes, S., Duke, R. and Franzusoff, A. (1996) Proc. Natl. Acad. Sci. U.S.A. **93**, 7695–7700
17 Kiefer, M. C., Tucker, J. E., Joh, R., Landsberg, K. E., Saltman, D. and Barr, P. J. (1991) DNA Cell Biol. **10**, 757–769
18 Tsuji, A., Hine, C., Mori, K., Tamai, Y., Higashine, K., Nagamune, H. and Matsuda, Y. (1994) Biochem. Biophys. Res. Commun. **202**, 1452–1459
19 Bruzzaniti, A., Goodge, K., Jay, P., Taviaux, S. A., Lam, M. H., Berta, P., Martin, T. J., Moseley, J. M. and Gillespie, M. T. (1996) Biochem. J. **314**, 727–731
20 Meerabux, J., Yaspo, M. L., Roebroek, A. J., Van de Ven, W. J., Lister, T. A. and Young, B. D. (1996) Cancer Res. **56**, 448–451
21 Seidah, N. G., Hamelin, J., Mamarbachi, M., Dong, W., Tardos, H., Mbikay, M. and Day, R. (1996) Proc. Natl. Acad. Sci. U.S.A. **93**, 3388–3393
22 Constam, D. B., Calfon, M. and Robertson, E. J. (1996) J. Cell Biol. **134**, 181–191
23 Tsuji, A., Mori, K., Hine, C., Tamai, Y., Nagamune, H. and Matsuda, Y. (1994) Biochem. Biophys. Res. Commun. **202**, 1215–1221
24 Ellison, M., Woodhouse, D., Hillyard, C. J., Dowsett, M., Coombes, R. C., Gilby, E. D., Greenberg, P. B. and Neville, A. M. (1975) Br. J. Cancer **32**, 373–379
25 Findlay, D. M., DeLuise, M., Michelangeli, V. P., Ellison, M. and Martin, T. J. (1980) Cancer Res. **40**, 1311–1317
26 Ausubel, F. M., Brent, R., Kingston, R. E., Moore D. D., Seidman J. G., Smith, J. A. and Struhl, K. (1994) Current Protocols in Molecular Biology, John Wiley, New York
27 Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. **162**, 156–159
28 Mount, S. M. (1982) Nucleic Acids Res. **10**, 459–472
29 Ohagi, S., LaMendola, J., LeBeau, M. M., Espinosa, R., Takeda, J., Smeekens, S. P., Chan, S. J. and Steiner, D. F. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 4977–4981
30 Ohagi, S., Sakaguchi, H., Sanke, T., Tatsuta, H., Hanabusa, T. and Nanjo, K. (1996) Diabetes **45**, 897–901
31 Ftouhi, N., Day, R., Mbikay, M., Chretien, M. and Seidah, N. G. (1994) DNA Cell Biol. **13**, 395–407
32 Mbikay, M., Raffin-Sanson, M. L., Tadros, H., Sirois, F., Seidah, N. G. and Chretien, M. (1994) Genomics **20**, 231–237
33 Tsuji, A., Hine, C., Tamai, Y., Yonemoto, K., Mori, K., Yoshida, S., Bando, M., Sakai, E., Mori, K., Akamatsu, T. and Matsuda, Y. (1997) J. Biochem. (Tokyo) **122**, 438–452
34 Shiramizu, B., Herndier, B. G. and McGrath, M. S. (1994) Cancer Res. **54**, 2069–2072
35 McKnight, S. L. and Kingsbury, R. (1982) Science **217**, 316–324
36 Dynan, W. S. and Tjian, R. (1983) Cell **35**, 79–87
37 Briggs, M. R., Kadonaga, J. T., Bell, S. P. and Tjian, R. (1986) Science **234**, 47–52
38 Kadonaga, J. T., Carner, K. R., Masiarz, F. R. and Tjian, R. (1987) Cell **51**, 1079–1090
39 Imagawa, M., Chiu, R. and Karin, M. (1987) Cell **51**, 251–260
40 Ayoubi, T. A., Creemers, J. W., Roebroek, A. J. and Van de Ven, W. J. (1994) J. Biol. Chem. **269**, 9298–9303

Received 11 May 1998/1 September 1998; accepted 24 September 1998