Cloning, sequencing and expression of rat liver pyruvate carboxylase

Sarawut JITRAPAKDEE*, Grant W. BOOKER*, A. Ian CASSADY† and John C. WALLACE*‡

*Department of Biochemistry, University of Adelaide, Adelaide, South Australia, 5005, Australia, and †Center for Molecular and Cellular Biology, University of Queensland, Brisbane, Queensland 5072, Australia

Overlapping clones encoding rat liver pyruvate carboxylase (PC) have been isolated by screening a liver cDNA library and by performing rapid amplification of cDNA ends polymerase chain reaction on total liver RNA. The sequence of rat PC cDNA contains an open reading frame of 3537 nucleotides encoding a polypeptide of 1178 amino acids with a calculated M_r of 129848. This is flanked by a 5′ untranslated region of 66 bp and a 3′ untranslated region of 421 bp including the poly(A) tail. The inferred protein sequence is 96.6% identical with mouse and 96.3% identical with human PCs, 68.4% identical with mosquito PC and 53.5% identical with yeast PC isoenzymes PC1 and PC2. On the basis of partial proteolysis and sequence homology with PC from other organisms (yeast, mosquito, mouse and human) and with other biotin enzymes, three functional domains, namely the biotin carboxylation domain, the transcarboxylation domain and the biotinyl domain, have been identified. Comparison with the known structure of the biotin carboxylase subunit of *Escherichia coli* acetyl-CoA carboxylase [Waldrop, Rayment and Holden (1994) Biochemistry **33**, 10249–10256] highlights the functional importance of 11 highly conserved residues. Northern analysis revealed that PC mRNA is highly expressed in rat liver, kidney, adipose tissue and brain, moderately expressed in heart, adrenal gland and lactating mammary gland, and expressed at a low level in spleen and skeletal muscle.

INTRODUCTION

Pyruvate carboxylase (PC) [pyruvate:carbon dioxide ligase (ADP-forming), EC 6.4.1.1] is a biotin-containing enzyme that catalyses the carboxylation of pyruvate to form oxaloacetate. The native enzyme consists of four identical subunits arranged in a tetrahedron-like structure [1]. Each subunit (M_r approx. 130000) contains a covalently bound biotin moiety that serves as a carrier for transferring CO_2 between the two reaction subsites [2]. Mammalian PC plays an important role in various metabolic pathways including gluconeogenesis, lipogenesis, porphyrinogenesis and the biosynthesis of neurotransmitter substances, by replenishment directly or indirectly of tricarboxylic acid cycle intermediates on which these pathways depend [3]. The expression of PC in mammals is tissue-specific, with its catalytic activity being highest in liver, kidney, adipose tissue, brain, lactating mammary gland and adrenal gland [4]. Changes in the total amount of PC protein, through alterations in the rate of synthesis, constitute an important mechanism for the long-term regulation of metabolic flux from pyruvate to oxaloacetate in the liver, kidney and adipose tissue [3]. Changes in the total amount of PC activity are also associated with alterations in the concentrations of other key gluconeogenic and lipogenic enzymes [3].

The complete amino acid sequences of PC from different organisms including yeast [5,6], mosquito [7], mouse [8] and human [9,10] have been reported. These data show a high degree of amino acid sequence similarity among species. Comparison of the primary structure of PC from a number of organisms allows the identification of structurally important residues that might play a role in the catalytic reaction. Here we report the complete nucleotide sequence and inferred amino acid sequence derived from overlapping rat PC cDNA clones, together with the tissue-specific expression of PC mRNA. The functional importance of particular residues is discussed within the context of the structural domains of the enzyme.

EXPERIMENTAL

cDNA library screening

A cDNA library in λgt10 prepared from the livers of Wistar rats was kindly supplied by Dr. G. W. Howlett, Russell Grimwade School of Biochemistry, University of Melbourne, Australia. A total of 2.5×10^7 plaque-forming units from this library were plated on *Escherichia coli* C600 Hfl$^-$ and screened by hybridization *in situ* with a 1.1 kbp *Bgl*II–*Stu*I fragment of a human PC cDNA clone, pPC 34.1 [11] as a probe. The probe was radiolabelled with ^{32}P by nick translation [12]. Hybridization was carried out in 5 × SSC (1 × SSC is 0.15 M sodium chloride, 0.015 M sodium citrate, pH 7.0), 5 × Denhardt's solution [0.1% Ficoll, 0.1% poly(vinylpyrrolidone), 0.1% BSA], 50 mM potassium phosphate, pH 6.8, 25 μg/ml salmon sperm DNA, 1% (w/v) glycine and 50% (v/v) formamide at 42 °C for 18 h. The filters were washed in 1 × SSC containing 0.1% SDS at 37 °C. The positive clones were purified and characterized by restriction-enzyme digestion. Screening of other overlapping clones was performed by replating the library and hybridization with insert DNA from positive clones under the same conditions as described above except that the hybridization was performed at 42 °C and washing was in 0.5 × SSC containing 0.1% SDS at 65 °C.

RACE polymerase chain reaction

The rapid amplification of cDNA ends polymerase chain reaction (RACE-PCR) and reverse transcriptase PCR (RT-PCR) were employed to obtain the full length of cDNA by using the 5′ AmpliFinder RACE kit (Clontech). The cDNA synthesis primers were designed either from the 5′ end of existing clones or from sequences conserved between mouse [8] and human [9,10]. Briefly, 10 μg of total liver RNA was hybridized to primers and reverse transcribed by AMV reverse transcriptase at 52 °C for 1 h. The

Abbreviations used: ACC, acetyl-CoA carboxylase; BC, biotin carboxylase; ODC, oxaloacetate decarboxylase; PC, pyruvate carboxylase; PCC, propionyl-CoA carboxylase; TC, transcarboxylase; RACE, **r**apid **a**mplification of **c**DNA **e**nds; RT, reverse transcriptase.

‡ To whom correspondence should be addressed.

The nucleotide sequence reported will appear in DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession number U36585.

purified cDNA was ligated to a 5′ Anchor adaptor followed by PCR with 5′ Anchor primer and nested primer. The PCR was performed in 50 μl of reaction mixture containing 10 mM Tris, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM each of dNTP, 0.01 % gelatin, 100 ng/ml of each primer, 1 μl of a 1 : 100 dilution of cDNA ligation reaction and 2 units (1 unit is defined as 10 nmol of dNTP incorporated in 30 min at 74 °C) of Taq polymerase (Promega). The PCR profile consisted of 40 cycles of amplification. Each cycle consisted of denaturation at 94 °C for 45 s, annealing at 65 °C for 45 s and extension at 72 °C for 2 min (7 min for the last cycle). The products were confirmed by Southern blot hybridization with ^{32}P-end-labelled internal oligonucleotide probes. The authentic products were purified from agarose gel with BRESAclean® (Bresatec, Adelaide, South Australia) and essentially re-amplified under the same conditions as described above except that 30 cycles of PCR were performed. The RT-PCR conditions were exactly the same as for the RACE-PCR except that the ligation of the adaptor sequence to the 3′ end of cDNA was omitted.

DNA sequencing

The rat cDNA inserts from positively hybridizing λgt10 clones were subcloned and sequenced as previously described [5] with T7 Sequenase (USB) and M13 universal and reverse sequencing primers.

The RT-PCR and RACE-PCR products were purified from agarose gel with BRESAclean® (Bresatec, Adelaide) and directly sequenced with the fmole® DNA sequencing system (Promega) using the same set of PCR and internal primers for the sequencing reactions.

Sequence analyses by computer

BLAST searches [13] were made courtesy of the NCBI (Bethesda, MD, U.S.A.) of the non-redundant database comprising the Protein Data Base (PDB) plus GenBank (release no. 91) plus GenBankupdate plus EMBL (release no. 43) plus EMBL update by TBLASTN, and on the non-redundant database comprising PDP plus SwissProt plus SPupdate plus PIR plus GenPept plus GPupdate by BLASTP, with an overlapping series of 11 fragments, each of approx. 200 residues, from the rat PC primary structure as the 'query' sequences.

Purification of rat PC

PC was purified from freeze-dried rat liver mitochondria by the method of Goss et al. [14], except that DEAE-Sephacel replaced DEAE-Sephadex as the first chromatographic matrix. Avidin-Sepharose affinity chromatography, as described previously [5], was then used to produce homogeneous enzyme.

Partial proteolysis and N-terminal sequencing

Rat PC (specific activity 14 μmol/min per mg of protein at 30 °C) at a concentration of 1 mg/ml in 100 mM *N*-ethylmorpholine, pH 8.0, was treated with chymotrypsin at a substrate-to-protease mass ratio of 150. After 0, 5, 10, 20 and 40 min of incubation at 37 °C, the reactions were stopped by adding PMSF to a final concentration of 1.0 mM. A subsample (5 μg) of each was fully reduced and denatured, and run on SDS/PAGE [12.5 % (w/v) gel] as previously described [5]. Proteolytic fragments of the enzyme were isolated from polyacrylamide gels by electrotransfer to 3-amino-propyltriethoxysilane-derivatized glass-fibre sheet by the method of Aebersold et al. [15]. The derivatized glass-fibre sheets were then placed directly into the reaction chamber of an Applied Biosystems gas-phase automated peptide sequencer (Model 470A). The resultant phenylthiohydantoin-amino acids evolved during each cycle were identified and quantified by comparison with derivatized standards chromatographed under the same conditions, as previously described [5].

Northern blot analysis

Total RNA of various tissues from normal, fed Hooded-Wistar rats was extracted as described by Chomczynski and Sacchi [16] with the minor modifications of including 1 % (w/v) SDS in the denaturing solution (4 M guanidinium thiocyanate, 25 mM sodium citrate, pH 7, 0.5 % sarcosyl, 0.1 M 2-mercaptoethanol) and performing LiCl precipitations [17]. Total RNA (40 μg) was denatured before electrophoresis on 2.2 M formaldehyde gel and transfer to a nylon membrane as described by Sambrook et al. [12]. The 1.6 kbp RACE-PCR product was labelled with [α-^{32}P]dATP by the random priming method with a Giga prime labelling kit (Bresatec, Adelaide). The filter was prehybridized in 20 ml of 5 × SSPE (1 × SSPE is 0.18 M sodium chloride, 10 mM sodium phosphate, pH 7.7, 1 mM EDTA), 5 × Denhardt's solution, 1 % (w/v) SDS, 50 % (v/v) formamide and 200 μg/ml salmon sperm DNA at 42 °C with gentle shaking. After 6 h of incubation, the hybridization buffer was replaced with 10 ml of fresh solution containing heat-denatured probe. Hybridization was performed at 42 °C for 12–16 h. The filter was washed twice with 2 × SSC containing 0.1 % SDS at 42 °C for 30 min and with 0.1 × SSC containing 0.1 % SDS at 42 °C for 30 min. The intensities of hybridization signals were compared by using 18 S and 28 S ribosomal RNA as internal standards.

RESULTS AND DISCUSSION

Isolation of PC cDNA clones from liver cDNA library

The initial screening of the rat liver cDNA library with a human cDNA probe yielded four positive clones, and digestion with *Eco*RI revealed that the longest insert (0.9 kbp) was in λRL1.1. The library was then rescreened with the λRL1.1 insert as probe to obtain clones representing regions nearer to the 5′ end. Two further overlapping PC cDNA clones were found, λRL2.35 and

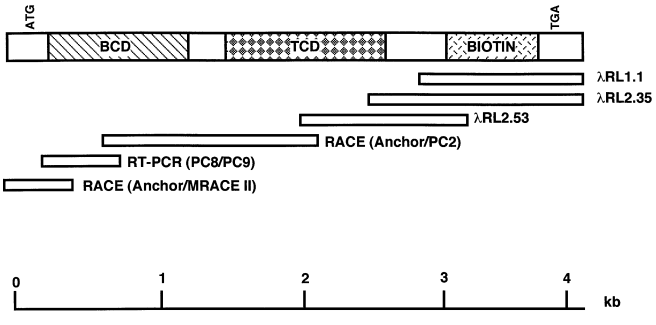

Figure 1 Schematic diagram of isolation of rat PC cDNA clones

Three overlapping clones, λRL1.1, λRL2.35 and λRL2.53, which represent the C-terminal part of PC, were obtained by screening a liver cDNA library. The three overlapping PCR clones, RACE (Anchor/PC2), RT-PCR (PC8/PC9) and RACE (Anchor/MRACE II), which extend further to encode the N-terminal part of PC, were obtained by RACE-PCR and RT-PCR. The PCR products were named after the primers used in the PCR amplification. Three functional domains within the coding region [biotin carboxylation (BCD), transcarboxylation (TCD) and biotin], the ATG start codon, the TGA stop codon and the relative length of each clone are shown.

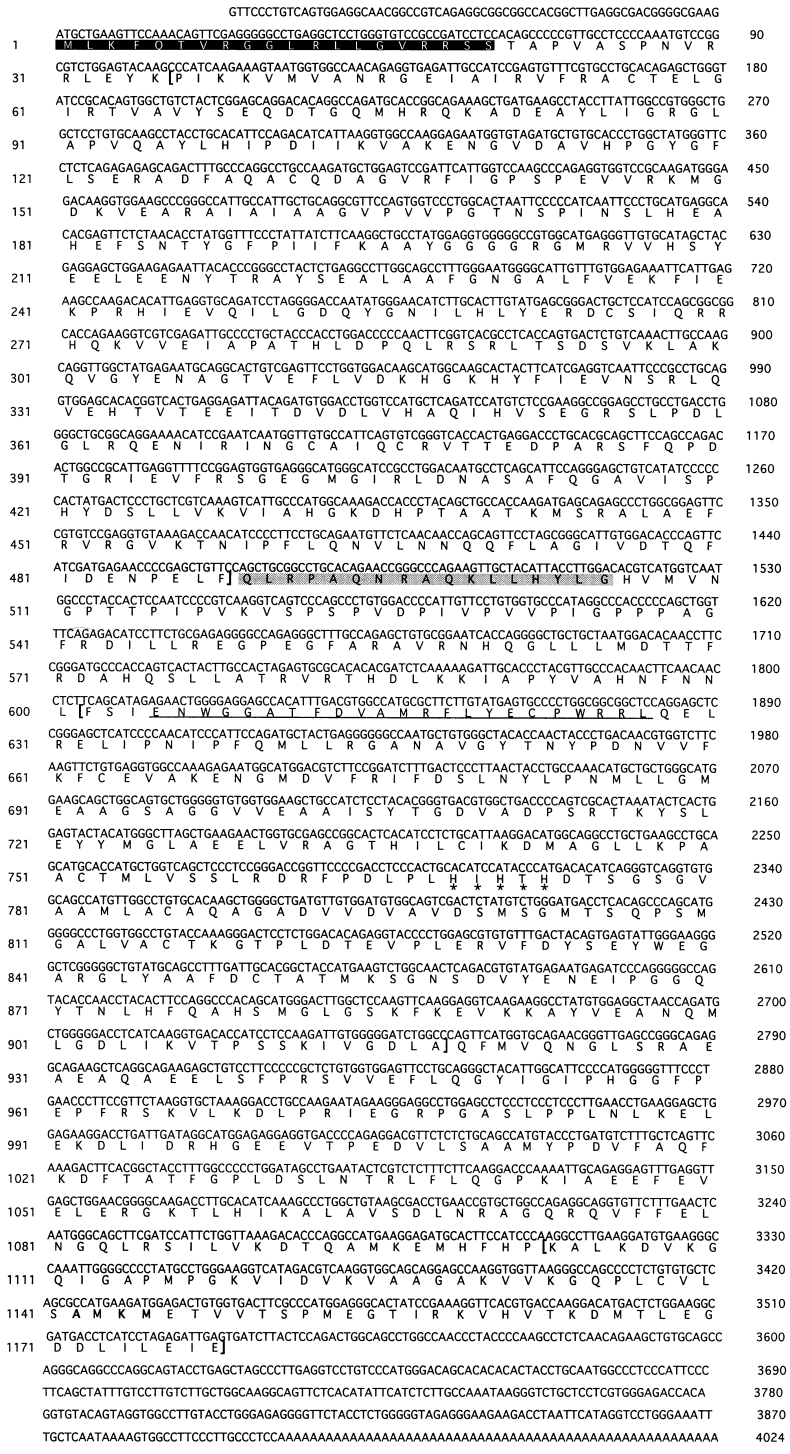

Figure 2 Nucleotide sequence and inferred amino acid sequence of rat PC

Amino acid residues are numbered at the left; nucleotide positions are numbered at the right. The predicted mitochondrial targeting sequence is shown in a black box. The domain boundaries, identified by highly significant degrees of similarity with other biotin enzymes, are indicated by square brackets. The conserved putative metal-binding motif (HXHXH) at residues 769–773 is identified by asterisks. The amino acid residues, which were identified by Edman degradation sequence analysis of the M_r approx. 75000 chymotryptic fragment (see Figure 4), are shown in a grey box. The putative carboxyl binding site (EXWGGATFDVAMRFLYECPWXRL) [22] within the pyruvate binding domain is underlined. The biotin attachment site (AMKM) is in bold type.

λRL2.53 (Figure 1). The inserts from these clones were subcloned into M13mp18 and M13mp19 and sequenced in both directions. The sequences showed more than 80% identity with the corresponding mouse and human PC nucleotide sequences. Attempts to isolate other clones from this library representing the remaining rat PC cDNA were unsuccessful. This is most probably due to the relatively large size (approx. 4 kbp) of the rat PC message and the inherent limitations of oligo-dT-primed cDNA libraries.

Isolation of cDNA clones by RACE-PCR

RACE-PCR was used to recover the remaining part of PC cDNA. Two antisense primers were used to synthesize two regions of PC cDNA. The first region was synthesized with a primer designed from a conserved sequence of mouse [8] and human [9,10] PC (nucleotide positions +358 to +380) followed by PCR with another internally conserved antisense primer corresponding to nucleotide positions +124 to +150 of mouse [8] and human [9,10] PC sequence and the 5′ adaptor primer. After the amplification, a 0.3 kbp product [RACE (Anchor/MRACE II)] was obtained. The second region was synthesized with the second primer designed from the 5′ end of λRL2.53 (nucleotide positions +1993 to +2014) followed by PCR with an internal antisense primer (nucleotide positions +1951 to +1974) and the 5′ end adaptor primer. The longest product obtained, [RACE (Anchor/PC 2)], was 1.6 kbp. DNA sequence analysis of these two fragments showed that approx. 200 nucleotides were missing from the junction between them owing to the inability of RT to synthesize the second region of cDNA beyond the 3′ end of the RACE (Anchor/MRACE II) fragment. The final fragment of PC cDNA was obtained by RT-PCR with a 3′ end sense primer of RACE (Anchor/MRACE II) fragment (nucleotide positions +61 to +75) and a 5′ end antisense primer of RACE (Anchor/PC 2) fragment. A 0.4 kbp product [RT-PCR (PC8/9)], which overlapped the above clones, was generated (Figure 1).

cDNA sequence and inferred amino acid sequence of rat PC

By screening a liver cDNA library and by PCRs, six overlapping clones spanning the entire coding region of rat PC were obtained. Figure 2 shows the complete sequence of the rat PC message encompassing 4024 nucleotides and including a 5′ untranslated region, coding region, stop codon and 3′ untranslated region. The largest open reading frame was 3537 nucleotides including the stop codon, encoding a polypeptide of 1178 amino acids with M_r 129 848. The 5′ untranslated region of rat PC is 66 nucleotides long, which is shorter than the 5′ untranslated region of mouse [8] and human PC [9]. The 3′ untranslated region of rat PC is 421 nucleotides long with a consensus polyadenylation signal (AATAAA) 27 nucleotides 5′ to a long poly(A) tail.

As a nuclear-encoded protein, PC is synthesized as a precursor protein in the cytosol and translocated to the mitochondria with a consequent reduction of molecular mass [18]. Many imported mitochondrial proteins are synthesized with an N-terminal leader sequence of 20–80 amino acid residues which typically contains several positively charged and several hydroxylated amino acids and no acidic amino acids [19], as is also the case with rat PC (see Figure 2). The mature N-terminal sequence of rat PC was previously reported to be Ser-Gly-Pro-Val-Ala-Pro-Leu-Asn-Val-Leu-Leu-Leu-Glu-Tyr-Pro by peptide sequencing [20]. However, this sequence was not found in the rat PC sequence that we inferred from the cDNA, nor does it exhibit any significant sequence identity with any biotin enzyme in the non-redundant databases listed in the experimental procedures. The proteolytic cleavage at the N-terminus of pre-mature PC during the translocation to mitochondria is predicted to occur between Ser[20] and Thr[21]. Cleavage at this site is a common feature of most mitochondrially imported proteins with their targeting sequence ending two residues after arginine [19]. The mitochondrial targeting sequence of rat PC showed high sequence similarity with mouse PC [8] and human PC [9,10], namely 95 % and 80 % respectively. When the amino acid sequence of rat PC was

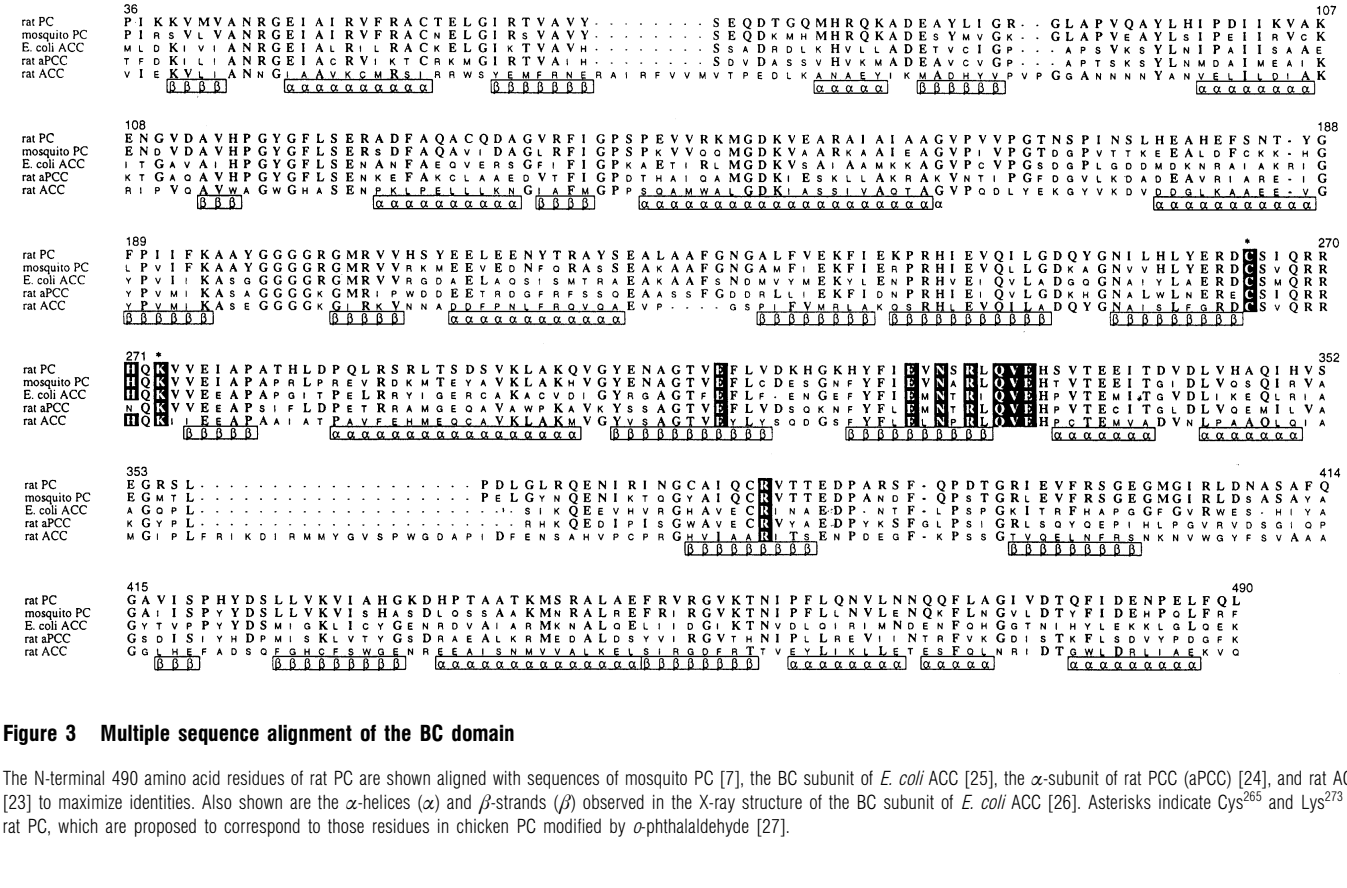

Figure 3 Multiple sequence alignment of the BC domain

The N-terminal 490 amino acid residues of rat PC are shown aligned with sequences of mosquito PC [7], the BC subunit of *E. coli* ACC [25], the α-subunit of rat PCC (aPCC) [24], and rat ACC [23] to maximize identities. Also shown are the α-helices (α) and β-strands (β) observed in the X-ray structure of the BC subunit of *E. coli* ACC [26]. Asterisks indicate Cys[265] and Lys[273] of rat PC, which are proposed to correspond to those residues in chicken PC modified by *o*-phthalaldehyde [27].

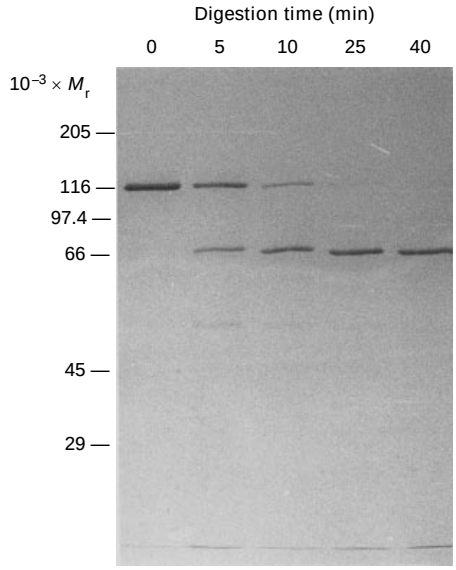

Figure 4 Time course of partial proteolysis of rat PC by chymotrypsin

Aliquots of the digests (see the Experimental section) were removed at the indicated times. Subsamples (5 μg) were then electrophoresed through SDS/PAGE (12.5% gel) and stained with Coomassie Brilliant Blue R250 as described previously [5].

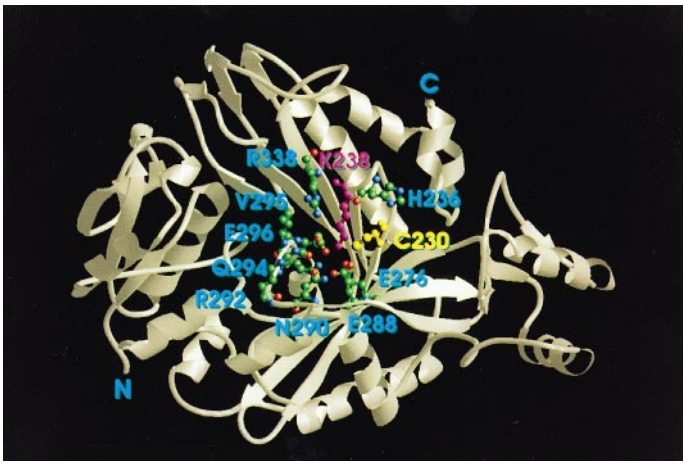

Figure 5 Structure of the BC subunit

A schematic representation of the three-dimensional structure of the BC subunit of *E. coli* ACC [26], showing the arrangement of β-strands and α-helices drawn with RIBBONS 2.0 [28]. The presumed active site is shown by the displayed side chains of highly conserved residues identified in Figure 3 (His^{236}, Glu^{276}, Glu^{288}, Asn^{290}, Arg^{292}, Gln^{294}, Val^{295}, Glu^{296}, Arg^{338}). For these residues indicated by their one-letter code and residue number, carbon atoms are shown in green, nitrogen in blue and oxygen in red. Lys^{238} and Cys^{230} are shown in magenta and gold respectively, and are proposed to correspond to the residues in chicken PC modified by *o*-phthalaldehyde [27]. Breaks in the polypeptide chain are the result of residues not present in the X-ray structure as well as *z*-clipping used to simplify the view.

aligned with sequences of PC from other organisms with the Gap program [21], it showed 96.6% overall identity with mouse PC, 96.3% identity with human PC, 68.4% identity with mosquito PC, and 53.5% identity with yeast PC1 and PC2.

Domain structure of rat PC

The biotin-dependent carboxylases are a family of enzymes with a diverse set of functions but with many common features and mechanisms [22]. Catalysis always involves two partial reactions: the carboxylation of the covalently attached biotin moiety and transcarboxylation from the carboxy-biotin to an acceptor molecule.

Biotin carboxylase (BC) domain

Figure 3 shows the multiple sequence alignment of the N-terminal region of rat and mosquito PC [7] with rat acetyl-CoA carboxylase (ACC) [23], the α-subunit of rat propionyl-CoA carboxylase (α-PCC) [24] and the BC subunit of *E. coli* ACC [25]. Also shown are the secondary structural elements identified from the X-ray crystal structure of the *E. coli* BC subunit [26]. It is clear from this alignment that these sequences are highly conserved and align well with the secondary structure of BC, implying that the overall folding of these proteins is likely to be conserved. This alignment extends the boundary of previously described homologous regions [5] to the N-terminus of the *E. coli* BC subunit, which corresponds to residue 36 of the inferred sequence of rat PC. The proposed boundary between the BC and transcarboxylase (TC) domains is consistent with the results of partial proteolysis experiments. Pure native rat PC, when treated with chymotrypsin, is cleaved into two main fragments in a time-dependent manner, as indicated by SDS/PAGE (Figure 4). The N-terminal sequence of the large fragment (M_r approx. 75000), which contains the biotin moiety (results not shown), was determined as Gln-Leu-Arg-Pro-Ala-Gln-Asn-Arg-Ala-Gln-Lys-Leu-Leu-His-Tyr-Leu-Gly. This matches exactly the cDNA-derived sequence for residues Gln^{489} to Gly^{505}, and places the chymotrypsin-labile peptide bond between Phe^{488} and Gln^{489}. This corresponds to the C-terminal extent of the sequence conservation and the structure of the BC subunit of *E. coli* ACC. We propose to call this conserved structural and functional unit within PC the 'BC domain' as it probably represents the first partial reaction site and incorporates the previously described ATP- and bicarbonate-binding domains [5].

The residues highlighted as white text in Figure 3 are those that are likely to play an important role in catalysis. The positions of these residues within the structure of the *E. coli* BC are shown in Figure 5. Waldrop et al. [26] identified His^{236} (His^{271} in rat PC), Lys^{238} (Lys^{273}), Glu^{276} (Glu^{311}), Glu^{288} (Glu^{324}), Asn^{290} (Asn^{326}), Arg^{292} (Arg^{328}), Gln^{294} (Gln^{330}), Val^{295} (Val^{331}), Glu^{296} (Glu^{332}) and Arg^{338} (Arg^{377}) as appropriately placed to interact with enzyme-bound biotin or phosphate molecules. Werneburg and Ash [27] showed that modification of a cysteine–lysine ion pair within the BC domain of chicken PC with *o*-phthalaldehyde resulted in inactivation of the first partial reaction. In *E. coli* BC the sulphur atom of Cys^{230} (Cys^{265} in rat PC) is 4.2 Å from the ϵ-amino atom of Lys^{238} (Lys^{273} in rat PC), which is sufficiently close to allow for cross-linking of these two residues by *o*-phthalaldehyde. From the sequence alignment it can be seen that these two residues are invariant within all the sequences shown in Figure 3 (also in these same enzymes from other species; results not shown) and reside within the presumed biotin carboxylation site (Figure 5). These results are consistent with the BC domain representing a conserved structural motif forming the first partial reaction subsite. Interestingly, the P-loop motif [GXXXXGK(TS)] commonly, though not universally, found in ATP- and GTP-binding proteins [29] is not present in rat or other PC sequences [5–10]. However, the sequence GGGRGMRVV between Gly^{199} and Val^{208} of rat PC is found without variation in all these PC sequences and in *E. coli* ACC, and with only conservative changes in rat ACC and PCC (see

Figure 3). In the *E. coli* BC structure [26], the corresponding residues Gly^{162} to Val^{171} are sufficiently mobile not to be observed in the X-ray structure.

Transcarboxylation domain

As was noted previously by Lim et al. [5] for residues 559–913 of yeast PC1, there is also a highly significant similarity between residues 602 and 918 of rat PC, the N-terminal region of the 5 S subunit of TC [EC 2.1.3.1] from *Propionibacterium shermanii* [22] and the α-subunit of oxaloacetate decarboxylase (ODC) [EC 4.1.1.3] from *Klebsiella pneumoniae* [30]. In particular, residues 605–627 [namely -ENWGGATFDVAMRFLYECPWRRL-] of rat PC conform to a consensus motif EXWGGATXDXXX-RFLXECPWXRL which is present in PCs from human [9,10], mouse [8], mosquito [7] and yeast [5,6] as well as the two oxaloacetate-metabolizing biotin-dependent enzymes of bacterial origin mentioned above. Kumar et al. [31] have shown that, in the presence of pyruvate, Trp^{73} (corresponding to the underlined Trp above) of the TC 5 S subunit is protected from modification by the tryptophan-specific reagent 2,4-dinitrophenyl sulphenyl chloride, suggesting that this residue is at or near the pyruvate binding site in that enzyme.

The motifs HXH [32] and HXXEH or HEXXH [33] are recognized as part of the metal binding sites of carbonic anhydrases and metalloproteinases respectively. Mozier et al. [34] have shown that a 24-residue tryptic peptide of protein kinase C inhibitor-1 (PKCI-1) containing the sequence HVHLH, as well as a number of synthetic peptides containing various length segments of the PKCI-1 sequence spanning the HVHLH site, would bind a single $^{65}Zn^{2+}$ ion. Because PC is known to be a metalloprotein [1] and *K. pneumoniae* ODC binds a Zn^{2+} ion [35], we have examined the known sequences for these enzymes to identify a putative metal binding site in the pyruvate domain of all PCs and ODC. The consensus sequence HXHXH, representing rat PC residues 769–773 (Figure 2) is consistent with the motifs for these other metal-binding enzymes mentioned above.

The putative transcarboxylation domain in rat PC is linked to the N-terminal biotin carboxylation domain and to the C-terminal biotinyl domain (see below) by proline-rich sequences. Of the 65 proline residues in rat PC, 13 are located between residues 492 and 550, just C-terminal of the biotin carboxylation domain, whereas 11 are located between residues 941 and 1042, just C-terminal of the transcarboxylation domain. There are also 9 alanine and 7 glycine residues between the BC domain and the transcarboxylation domain. A further 12 alanine and 11 glycine residues occur between the transcarboxylation domain and the biotinyl domain. However, in general they are not distributed in as close proximity to the prolines as in the α-subunit of ODC from *K. pneumoniae* [30] and in the biotin carrier protein of *P. shermanii* TC [22]. These unusual sequences have been proposed to provide flexibility for movement of the biotin prosthetic group between catalytic centres in a manner analogous to the highly mobile Pro-Ala sequences in lipoated proteins [36]. Because PC is composed of three domains that must fold together to form a single active site, it is reasonable to expect the presence of two proline-rich regions capable of forming hinge-like structures.

Biotinyl domain

The biotin-attachment site of PC has previously been identified from sheep and avian species by peptide sequencing [37]. The conserved motif seems to be AMKM with biotin covalently attached to the ϵ-amino group of the lysine via an amide bond. With the advance of recombinant DNA technology, more sequence information around biotin attachment sites of many biotin-containing enzymes have been obtained by sequencing cDNA and genomic DNA. Samols et al. [22] showed that the amino acid sequences within the biotinyl domains from a number of organisms have the conserved motif AMKM. The conserved proline that is located 27 residues N-terminal of the biocytin in rat PC has been found in other PCs and many other biotin-containing enzymes studied so far [22].

Thampy et al. [20] reported the sequence of 24 amino acid residues around the biotin attachment site of rat PC by peptide sequencing. However, seven amino acid differences were noted between the previously reported peptide sequence [20] and the sequence inferred here from cDNA sequencing and from the 3′ end of a genomic clone (results not shown). This, in addition to the difference in N-terminal peptide sequence, raises the possibility of there being two isoforms of rat PC as in *Saccharomyces cerevisiae* [5,6,38]. However, our sequence in this region is also identical with mouse PC [8] and human PC [9,10, 39], which were shown to be highly conserved. Only one gene for PC has been identified in humans [40]. The biocytin of rat PC is located 35 residues upstream from the C-terminus, in common with many biotin-containing enzymes [22].

Tissue-specific expression of rat PC

Northern analysis demonstrated the presence in a variety of adult rat tissues of a single PC mRNA that was approx. 4.2 kbp in length (see Figure 6), in agreement with the size of the PC cDNA (4024 bp). The expression of PC mRNA exhibits a tissue-specific pattern, being most abundant in liver and kidney, which are the gluconeogenic organs. The levels of PC message in these organs are also consistent with those in human liver and kidney [10]. The levels of mRNA are also substantial in adipose tissue and brain, in keeping with the relatively high levels of PC activity in these tissues [4]. This is not surprising because adipocyte PC plays a role in lipogenesis in the export of mitochondrial acetyl-CoA to the cytosol as citrate for the biosynthesis of fatty acids *de novo* [3]. Recently Zhang et al. [41] showed that an increase in the rate of synthesis of PC in mouse adipocytes is due to an increase in the concentration of this enzyme's translatable mRNA. In

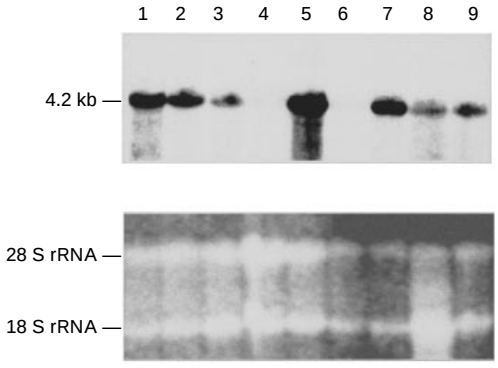

Figure 6 Northern analysis of PC mRNA in different tissues

Samples of total RNA (40 μg) from different rat tissues were electrophoresed and blotted on nylon membrane as described in the Experimental section, and then hybridized with the 1.6 kbp RACE (Anchor/PC2) clone (upper panel). RNA loading was assessed by comparison with the 18 S and 28 S RNA bands revealed by staining with ethidium bromide (lower panel). Lane 1, adipose tissue; lane 2, brain; lane 3, heart; lane 4, spleen; lane 5, liver; lane 6, skeletal muscle; lane 7, kidney; lane 8, lactating mammary gland; lane 9, adrenal gland.

brain, PC plays an anaplerotic role in the transfer of acetyl groups from mitochondria to the cytosol for the biosyntheses of several neurotransmitter substances [3]. Moderate expression of PC mRNA was detected in adrenal gland, heart and lactating mammary gland. A very low level of expression was found in skeletal muscle and spleen, suggesting the anaplerotic function of this enzyme may be less important for the normal functions of these tissues.

We thank Dr. F. Tomas and Dr. P. Rogers, CSIRO Division of Human Nutrition, Adelaide for rat tissue samples; Dr. A. Chapman-Smith, Dr. D. Val and Dr. M. Walker for valuable discussions; and Ms. D. Turner for N-terminal protein sequencing. This work was supported by Grant AO9600585 to J.C.W. from the Australian Research Council. S.J. is a recipient of a Royal Thai Government Scholarship.

REFERENCES

1 Wallace, J. C. and Easterbrook-Smith, S. B. (1985) in Pyruvate Carboxylase (Keech, D. B. and Wallace, J. C. eds.), pp. 65–108, CRC Series in Enzyme Biology, CRC Press, Boca Raton, FL
2 Attwood, P. V. (1995) Int. J. Biochem. Cell Biol. **27**, 231–249
3 Barritt, G. J. (1985) in Pyruvate Carboxylase (Keech, D. B. and Wallace, J. C. eds.), pp. 141–177, CRC Series in Enzyme Biology, CRC Press, Boca Raton, FL
4 Wallace, J. C. in Pyruvate Carboxylase (Keech, D. B. and Wallace, J. C. eds.), pp. 5–63, CRC Series in Enzyme Biology, CRC Press, Boca Raton, FL
5 Lim, F., Morris, C. P., Occhiodoro, F. and Wallace, J. C. (1988) J. Biol. Chem. **263**, 11493–11497
6 Stucka, R., Dequin, S., Salmon, J. M. and Gancedo, C. (1991) Mol. Gen. Genet. **229**, 307–315
7 Tu, Z. and Hagedorn, H. H. (1994) GenBank accession number L 36530
8 Zhang, J., Xia, W., Brew, K. and Ahmad, F. F. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 1766–1770
9 Mackay, N., Rigat, B., Douglas, C., Chen, H.-S. and Robinson, B. H. (1994) Biochem. Biophys. Res. Commun. **202**, 1009–1014
10 Wexler, I. D., Du, Y., Lisgaris, M. V., Mandal, S. K., Freytag, S. O., Yang, B., Liu, T., Hwon, M., Patel, M. S. and Kerr, D. S. (1994) Biochim. Biophys. Acta **1227**, 46–52
11 Lamhonwah, A.-M., Quan, F. and Gravel, R. A. (1987) Arch. Biochem. Biophys. **254**, 631–636
12 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 10.7
13 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) J. Mol. Biol. **215**, 403–410

Received 14 November 1995/24 January 1996; accepted 5 February 1996

14 Goss, N. H., Dyer, P. Y., Keech, D. B. and Wallace, J. C. (1979) J. Biol. Chem. **254**, 1734–1739
15 Aebersold, R. H., Teplow, D. B., Hood, L. E. and Kent, S. B. H. (1986) J. Biol. Chem. **261**, 4229–4238
16 Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. **162**, 156–159
17 Cathala, G., Savouret, J.-F., Mendez, B., West, B. L., Karin, M., Martial, J. A. and Baxter, J. D. (1983) DNA **2**, 329–335
18 Srivastava, G., Borthwick, I. A., Brooker, J. D., Wallace, J. C., May, B. K. and Elliott, W. H. (1983) Biochem. Biophys. Res. Commun. **117**, 344–349
19 Hendrick, J. P., Hodges, P. E. and Rosenberg, L. E. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 4056–4060
20 Thampy, K. G., Huang, W.-Y. and Wakil, S. J. (1988) Arch. Biochem. Biophys. **266**, 270–276
21 Program Manual for the Wisconsin Package, Ver. 8.1, Sept. 1995, GCG, 575 Science Drive, Madison, WI 53711, USA
22 Samols, D., Thornton, C. G., Murtif, V. L., Kumar, G. K., Haase, F. C. and Wood, H. G. (1988) J. Biol. Chem. **263**, 6461–6464
23 Lopez-Casillas, F., Bai, D.-H., Luo, X., Kong, I.-S., Hermodson, M. A. and Kim, K.-H. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 5784–5788
24 Browner, M. F., Taroni, F., Sztul, E. and Rosenberg, L. E. (1989) J. Biol. Chem. **264**, 12680–12685
25 Li, S. and Cronan, J. E., Jr. (1992) J. Biol. Chem. **267**, 855–863
26 Waldrop, G. L., Rayment, I. and Holden, H. M. (1994) Biochemistry **33**, 10249–10256
27 Werneberg, B. G. and Ash, D. E. (1993) Arch. Biochem. Biophys. **303**, 214–221
28 Carson, M. (1991) J. Appl. Crystallogr. **24**, 958–961
29 Saraste, M., Sibbald, P. R. and Wittinhofer, A. (1990) Trends Biochem. Sci. **15**, 430–434
30 Schwarz, E., Oesterhelt, D., Reinke, H., Beyreuther, K. and Dimroth, P. (1988) J. Biol. Chem. **263**, 9640–9645
31 Kumar, G. K., Haase, F. C., Phillips, N. F. B. and Wood, H. G. (1988) J. Biol. Chem. **27**, 5978–5983
32 Vallee, B. L. and Auld, D. S. (1990) Biochemistry **29**, 5647–5659
33 Jiang, W. and Bond, J. S. (1992) FEBS Lett. **312**, 110–114
34 Mozier, N. M., Walsh, M. P. and Pearson, J. D. (1991) FEBS Lett. **279**, 14–18
35 Dimroth, P. and Thorner, A. (1992) FEBS Lett. **300**, 67–70
36 Radford, J. E., Lane, E. D., Perham, R. N., Martin, S. R. and Appella, E. (1989) J. Biol. Chem. **264**, 767–775
37 Rylatt, D. B., Keech, D. B. and Wallace, J. C. (1977) Arch. Biochem. Biophys. **183**, 113–122
38 Walker, M. E., Val, D., Rhode, M., Devenish, R. J. and Wallace, J. C. (1991) Biochem. Biophys. Res. Commun. **176**, 1210–1217
39 Freytag, S. O. and Collier, K. J. (1984) J. Biol. Chem. **259**, 12831–12837
40 Walker, M. E., Baker, E., Wallace, J. C. and Sutherland, G. R. (1995) Cytogen. Cell Gen. **69**, 187–189
41 Zhang, J., Xia, W.-L. and Ahmad, F. (1995) Biochem. J. **306**, 205–210