Pyruvate Carboxylase: Relating Structure to Function

by
Grant William Booker, B.Sc.(Hons)

A thesis submitted to the
University of Adelaide, South Australia
For the degree of
Doctor of Philosophy.

awarded 1990

July, 1990
Department of Biochemistry,
University of Adelaide,
South Australia.

Table of Contents

SUMMARY

Pyruvate carboxylase (PC) [EC. 6.4.1.1] is a large tetrameric enzyme, found in a wide variety of organisms where it catalyses the conversion of pyruvate to oxaloacetate. PC is a member of the enzyme family known as the biotin-dependent carboxylases, where the covalently attached prosthetic group biotin is an absolute requirement for activity. The reaction mechanism involves the carboxylation of biotin, followed by transfer of this carboxylate moiety to pyruvate. These two reactions occur at spatially distinct subsites and are known as the 1st and 2nd partial reactions respectively. Although the catalytic mechanism of the above reactions has been extensively studied [see Knowles (1989): *Annu. Rev. Biochem.* **58**, p195-221], far less is known of the 3-dimensional structure of the tetrameric molecule and how this structure is involved in the various functions of the enzyme. This is due in part, to the absence of PC crystals suitable for X-ray diffraction studies and the size of the tetramer (approximately Mr 520,000) effectively precluding it from analysis by NMR.

Due to the lack of a detailed 3-D structural model, the study reported in this thesis was aimed at providing structural information on the functional aspects of PC, by the combination of a number of techniques including partial proteolysis, chemical modification, immunochemistry, electron microscopy, and molecular biology.

Partial cDNA and genomic clones encoding rat PC were sequenced, providing 753 amino acids of sequence data from the C-terminus of the protein. These are encoded by ten exons over approximately 5 kb of DNA. Significant sequence homologies were observed between PC from rat and yeast both at the protein and nucleotide levels.

Limited treatment of sheep, rat, yeast and chicken PC with chymotrypsin demonstrated the existence of a large C-terminal fragment which is stable to further proteolysis, and sequence analysis by Edman degradation, showed the N-termini of the corresponding fragments to be identical in all four species. Similar treatment of sheep PC with proteinase K resulted in cleavage at a site identical to chymotrypsin, but in addition a second fragment was produced which lacked approximately Mr 6,000 from the C-terminus. These results together with the sequence similarities between rat and yeast PC confirm the proposed ligand binding regions reported by Lim *et al.* [(1988): *J. Biol Chem.* **263**, p11493-11497].

The effects of chymotrypsin and proteinase K treatment on the tetrameric structure

truncated by 10 - 13% of the length of the molecule, although 40% of the total mass of the protein had been removed. A structural model has been proposed to explain this difference. Tetrameric particles could also be observed in samples where treatment with proteinase K had truncated the protein at positions Gly-17 and Arg-978. This indicates that the residues important for intersubunit binding are not contained in regions of the protein assigned to binding ATP or biotin.

Western analysis of partially proteolysed sheep PC, allowed the localisation of 5 monoclonal antibody epitopes within the length of the polypeptide. Monoclonal antibodies #12 and #42 bind to a region within approximately Mr 6,000 of the C-terminus, #6 and #18 bind to a region within approximately Mr 75,000 of the C-terminus while monoclonal antibody #60 binds to a region between Mr 75,000 and Mr 6,000 from the C-terminus of PC.

Limited treatment of PC with trinitrobenzene sulphonate (TNBS) results in modification of two lysine residues, one of which (Lys-A) is important for mediating the allosteric effect of acetyl-CoA. (Ashman *et al.* (1973) *Biochem. Biophys. Res Commun.* **51**, p924-931).

Acetyl-CoA and 3´5´ADP were shown to protect yeast PC from inhibition by TNBS as is the case for PC from sheep . A [^{14}C]acetyl-CoA binding assay was developed, and used to demonstrate that modification of Lys-A inhibits the binding of acetyl-CoA to yeast PC, indicating that Lys-A is at or near the acetyl-CoA binding site. Localisation of Lys-A within the protein sequence was attempted using peptide mapping techniques. Fab fragments of anti-TNP IgG were produced and observed as complexes with TNBS-treated yeast PC, using electron microscopy. Lys-A and hence the acetyl-CoA binding site was localised to the central region of the tetramer, near the inter-subunit junction using this technique.

Statement

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University. To the best of my knowledge and belief it contains no material that has been previously published by any other person except where due reference is made in the text. The author consents to the thesis being made available for photocopying and loan.

Grant W. Booker

Erratum

The following reference should be included in the bibliography

Taylor, B.L., Routman, S. and Utter, M.F. (1975) *J. Biol. Chem.* **250**, 2376-2382, The control of the synthesis of pyruvate carboxylase in *Pseudomonas citronellolis*.

Acknowledgements

I wish to thank Professors W.H. Elliott and G.E. Rogers for permission to undertake this research project in the Department of Biochemistry, University of Adelaide. I am eternally grateful to my supervisor, Dr. John C. Wallace for his support, careful criticism and friendly advice throughout. In addition, I acknowledge the financial support of a Commonwealth Postgraduate Research Award.

I would like to thank Professor Frank Mayer for his collaboration in the electron microscopy reported as part of Chapters 4 & 5, Dr. A. Ian Cassady for providing the cloned DNA analysed in Chapter 3 and Dr. Michelle Walker for her comments and criticisms during the preparation of this thesis.

In addition, I would like to thank the past and present members of the Biotin-Biocentre and Growth Faction, particularly Drs. John Carver, Phil Morris, Manfred Rohde, Ian Cassady, Filip Lim, Sue Gale, Karen Carey, David (there's only one) Mottershead, Chris Bagley, Michelle Walker and Briony Forbes, Zee Upton, Laszlo Szabo, Sharon Gargosky, Nick Kallincos, Neil Brewster, Kathy Clark and Filomena Occhiodoro for being the brightest, weirdest, most friendly bunch of people with whom to work. Special thanks must go to Barb Magee and Paddy, for allowing me to use their Macintosh computer.

I am grateful to Jenny Brazier and Yvonne Riese for expert technical assistance, Mark Snoswell and Denise Turner for their expertise in protein sequencing and Ian Dodd for helpful discussions on computing matters.

I especially thank Tricia and my expanded family, for their interest, constant support and patience.

Finally, I thank Athos, Porthos and D'Artagnan for the good times and bad wine.

Abbreviations

In addition to those accepted for use in the *Journal of Biological Chemistry* , the following abbreviations are used in this thesis:

A_λ	absorbance at wavelength λ
ARG	endoproteinase arg-C
BIO	biotin-carrier domain
CoA	coenzyme A
CHY	chymotrypsin
DTE	dithioerythritol
DTNB	5,5´-dithio-bis-(2-nitrobenzoic acid)
EPR	electron paramagnetic resonance
FDNB	1-fluoro-2,4-dinitrobenzene
Mr	relative molecular mass
Nem	N-ethyl morpholine
NMR	nuclear magnetic resonance
PK	proteinase K
Pi	inorganic phosphate
PMSF	phenylmethane sulphonyl fluoride
PTH-	phenylthio-hydantoin-
PYR	pyruvate/oxaloacetate-binding domain
rp-HPLC	reverse phase high performance liquid chromatography
RPC	rat pyruvate carboxylase
s.a.	specific activity
SDS	sodium dodecyl sulphate
SPC	sheep pyruvate carboxylase
TNBS	2,4,6 trinitrobenzene sulphonic acid
TNP-	trinitrophenyl-
TPCK	L-(tosylamido 2-phenyl)ethyl chloromethylketone
YPC	yeast pyruvate carboxylase

CHAPTER 1

INTRODUCTION

1.1 HISTORICAL BACKGROUND

Glycolysis is a central metabolic pathway which catalyses the conversion of glucose to pyruvate, and is found in a wide variety of aerobic and anaerobic organisms. This conversion results in the net production of energy, in the form of ATP and NADH, with the resultant pyruvate then being available to the cell to form acetyl-CoA, lactate or ethanol, depending on the cell type, species or metabolic circumstances.

The reversal of this pathway (gluconeogenesis) was shown to occur via a symmetrical dicarboxylic acid intermediate, by isotopic labelling experiments (Solomon *et al.*, 1940; Wood *et al.*, 1945; Topper & Hastings, 1949; Lorber *et al.*, 1950). These experiments demonstrated that the conversion of lactate or pyruvate to glycogen involved the fixation of CO_2. Although many of the reactions involved in glycolysis are readily reversible and are therefore shared by glycolysis and gluconeogenesis, three steps are considered to be unidirectional and the enzymes catalysing these reactions differ between the pathways. In glycolysis, these reactions are catalysed by hexokinase [EC 2.7.1.1], 6-phosphofructokinase [EC 2.7.1.11] and pyruvate kinase [EC 2.7.1.40].

The final step in glycolysis, catalysed by pyruvate kinase converts phosphoenolpyruvate to pyruvate and although under certain conditions this reaction can be reversed *in vitro* (Utter, 1959), the reversal *in vivo* is thought to be thermodynamically unfavoured (Krebs, 1954). The cell must, therefore, utilise some other means to produce phosphoenolpyruvate from pyruvate.

Utter and Kurahashi (1954a) showed that phosphoenolpyruvate could be formed from oxaloacetate, a reaction catalysed by phosphoenolpyruvate carboxykinase [EC 4.1.1.32], providing the first link between the dicarboxylic acids and phosphoenolpyruvate production. The requirement then was for a reaction scheme which produced oxaloacetate from pyruvate. A two step process was proposed (Utter and Kurahashi 1954b), involving the intermediate malate, but this too was considered thermodynamically unfavourable (Krebs, 1954).

In 1960, a mitochondrial matrix protein from chicken liver was discovered which provided the explanation for how the cell circumvents this thermodynamic barrier. The enzyme pyruvate carboxylase [EC.6.4.1.1] was isolated by Utter and Keech (1960) and

found to catalyse the conversion of pyruvate to oxaloacetate. Phosphoenolpyruvate carboxy kinase could then convert oxaloacetate into phosphoenolpyruvate. The two enzymes working in concert, are able thereby to side-step the essentially irreversible reaction of pyruvate kinase.

Pyruvate carboxylase (PC) was shown to catalyse the following overall reaction

$$\text{pyruvate} + HCO_3^- + \text{ATP} \xrightleftharpoons{Me^{2+},\ \text{acetyl-CoA}} \text{oxaloacetate} + \text{ADP} + P_i \quad \text{(equ 1.1)}$$

where divalent metal ions (Me^{2+}) and acetyl-Coenzyme-A (Acetyl-CoA) are activators of the reaction (Utter and Keech, 1960).

Apart from participating in gluconeogenesis, PC has an anaplerotic role in maintaining the levels of tricarboxylic acid cycle intermediates for use in biosynthesis (eg. amino acid, fatty acid and acetylcholine synthesis). An important component of the enzyme is the vitamin biotin, which is a post-translational modification of a particular lysine residue, approximately 35 residues from the C-terminus.

The significance of PC's role in metabolism can be seen in the few well characterised reports of human pyruvate carboxylase deficiency (for a review, see Barritt, 1985). These individuals usually die during infancy, with symptoms such as abnormally high serum pyruvate and lactate levels, neurological lesions and proximal renal acidosis. In instances where pyruvate carboxylase deficient individuals survive, they are severely mentally handicapped, presumably due to the occurrence of neurological lesions (De Vivo *et al.*, 1977; Oizumi *et al.*, 1983).

1.2 THE BIOTIN CARBOXYLASES

The biotin carboxylases are a family of enzymes which catalyse a diverse range of metabolic reactions, in each case utilising the vitamin biotin as a covalently attached cofactor. Ten members of this family have been identified, as shown in table 1.1. These enzymes can be classified into three groups, those involved with carboxylation, decarboxylation and transcarboxylation reactions (Wood and Barden, 1977).

(1) Carboxylation

$$R\text{-}H + HCO_3^- + ATP \rightleftharpoons R\text{-}CO_2^- + ADP + Pi \quad \text{(equ 1.2)}$$

(2) Decarboxylation

$$R\text{-}CO_2^- + H^+ + 2Na^+_{(in)} \rightleftharpoons R\text{-}H + HCO_3^- + 2Na^+_{(out)} \quad \text{(equ 1.3)}$$

(3) Transcarboxylation

$$^-O_2C\text{-}R\text{-}CO\text{-}CoA + R'\text{-}CO_2^- \rightleftharpoons R\text{-}CO\text{-}CoA + {}^-O_2C\text{-}R'\text{-}CO_2^- \quad \text{(equ 1.4)}$$

Based on these similarities and the quaternary structure of the various enzymes, Lynen (1975) proposed an evolutionary relationship between the members of the biotin carboxylases, suggesting that successive gene fusions had led to evolution of enzymes composed of multifunctional polypeptides, from precursors composed of a number of monofunctional peptides. This is illustrated in figure 1.1, where BC represents the biotin-carboxylation reaction centre, C represents the biotin-carrier protein, CT represents the trans-carboxylation reaction centre and R represents the regulatory region of the protein. It is only now, some 12-15 years later, that clear evidence for this proposal in the form of extensive sequence homologies, is accumulating through the application of recombinant DNA technology to this family of enzymes.

1.3 MECHANISM OF PYRUVATE CARBOXYLATION

The overall reaction catalysed by PC, shown in equation 1.1, occurs as a multi step process (Knowles, 1989) which can be simplified to the carboxylation and decarboxylation of biotin, as shown in equations 1.5 and 1.6. These two reactions have been shown to occur at separate subsites (Utter *et al.*, 1975), and the biotin moiety is thought to oscillate between the two subsites in a manner analogous to the lipoyl-arm serving the reaction centres of pyruvate dehydrogenase (Ambrose and Pernham, 1976; Grande *et al.*, 1976). The two reactions of PC can be assayed independently by isotope exchange (Scrutton *et al.*, 1965) and are called the 1st partial and 2nd partial reactions respectively.

1st Partial Reaction

$$\text{ENZ-biotin} + HCO_3^- + \text{ATP} \xrightleftharpoons{Me^{2+},\ \text{acetyl-CoA}} \text{ENZ-biotin-}CO_2^- + \text{ADP} + P_i \quad \text{(equ 1.5)}$$

2nd Partial Reaction

$$\text{ENZ-biotin-}CO_2^- + \text{pyruvate} \rightleftharpoons \text{ENZ-biotin} + \text{oxaloacetate} \quad \text{(equ 1.6)}$$

The detailed chemistry behind the two partial reactions has been the subject of much debate in the literature, leading to a variety of proposed mechanisms. Drawing all the available evidence together, Knowles (1989) proposed a "best fit" mechanism for the 1st partial reaction, shown in figure 1.2. This model involves the formation of a carboxy-phosphate intermediate and is based largely on the analogous reaction for carbamyl phosphate synthetase [EC 6.3.5.5]. Although the carboxy-phosphate intermediate proposed in this scheme has yet to be isolated, (or trapped as its trimethylester equivalent) evidence for its existence is based on the ability of PC and the biotin-carrier protein of *E. coli* acetyl-CoA carboxylase to catalyse the formation of ATP from ADP and carbamyl phosphate (Ashman and Keech, 1975; Polakis *et al.*, 1974; Ogita and Knowles, 1988) and radiochemical-labelling studies of Wallace *et al.* (1985). These latter authors, using gel filtration, were able to show that in the absence of acetyl-CoA, enzyme bound radioactivity could be detected after incubation with either $H^{14}CO_3^-$ or [γ-^{32}P]-ATP. When acetyl-CoA was added, however, $^{14}CO_2$ proceeded to be incorporated into oxaloacetate, and ^{32}Pi was released, suggesting that step II or III as shown in figure 1.2, is the likely point of influence of acetyl-CoA.

Binding of pyruvate to the enzyme induces the carboxy-biotin moiety to translocate to the second subsite (Goodall *et al.*, 1981), where carboxylation of pyruvate subsequently occurs. Using carbon-13 and deuterium isotope effects, Attwood *et al.* (1986b) were able to rule out concerted mechanisms for the decarboxylation of biotin, put forward by Goodall *et al.* (1983) and Retey and Lynen (1965), and instead proposed the scheme shown in figure 1.3.

The requirements set, by these authors, for this to be a viable mechanism are the presence of an amino acid such as cysteine (with a pKa $\leq$ 4.5) and another base such as lysine or histidine to act as an ion pair with the ureido-oxygen of biotin. Evidence for such residues comes from the chemical modification studies of Hudson *et al.* (1975) and Bagley

et al. (1983), suggesting the presence of at least one catalytically important cysteine in the pyruvate binding site. Although Attwood *et al.* (1986b) indicate direct coordination between Mn^{2+} and oxaloacetate in their proposed mechanism, it has been omitted from the scheme shown in figure 1.3, since studies by Carver *et al.*, (1988) and Fung *et al.*, (1973) suggest Mn^{2+} plays more of a structural role, influencing the stability of the tetramer (see section 1.4.4) .

1.4 PYRUVATE CARBOXYLASE STRUCTURE

1.4.1 Primary Structure

The first amino acid sequence data published for the biotin carboxylases was derived from proteolytically cleaved peptides, incorporating the site of biotinylation of PC (Rylatt *et al.*, 1977) and the biotin carrier proteins of transcarboxylase (Maloy *et al.*, 1979) and *E. coli* acetyl-CoA carboxylase (Sutton *et al.*, 1977). These sequences displayed strong inter-enzyme and inter-species similarities, and although consistent with Lynen's hypothesis, these similarities need not have been the result of gene fusion, but rather a consequence of functional similarities between the enzymes. More sequence data, spread over a wider enzyme and species range is required to support such an hypothesis of evolutionary relationship.

Clearly, members of the family of biotin carboxylases are too large to determine the primary structure efficiently by conventional protein sequencing, and require the application of recombinant DNA technology, with the amino acid sequence being inferred from the DNA sequence of either cDNA or genomic clones. Using such techniques, Freytag and Collier (1984) determined the C-terminal 44 amino acids of human PC, which represents approximately $1/30^{th}$ of the total expected length. Lim *et al.* (1988) have subsequently published the entire protein sequence of yeast PC, derived from the nucleotide sequence of genomic clones of *Saccharomyces cerevisiae*.

When this sequence is compared to data published for other biotin carboxylases: transcarboxylase (Maloy *et al.*, 1979; Wood & Kumar, 1985); chicken (Takai *et al.*, 1988) and rat (Lopez-Casillas *et al.*, 1988) acetyl-CoA carboxylase (ACC); human (Lamhonwah *et al.*, 1986;1989) and rat (Kraus *et al.*, 1986, Browner *et al.* , 1989) propionyl-CoA carboxylase (PCC); and *Klebsiella pneumoniae* oxaloacetate decarboxylase (Schwarz *et al.*, 1988), striking localised similarities are observed. Conservation of sequence has also

been observed between yeast PC and proteins from two other enzyme families, carbamyl phosphate synthetase (Lusty *et al.*, 1983; Nyunoya *et al.*, 1985) and pyruvate dehydrogenase (Guest *et al.*, 1985). From these comparisons Lim *et al.* (1988) proposed roles for three regions of the PC polypeptide, namely ATP/HCO_3^- binding, pyruvate binding and biotin attachment. The relative positions of these sequence similarities are indicated schematically in figure 1.4.

1.4.2 Quaternary Structure

Pyruvate carboxylases from all but one species so far investigated as homogeneous preparations, exist as active tetramers of approximately Mr 520,000. Each monomer (Mr 130,000) contains one bound biotin and one Mn^{2+} (or Zn^{2+}, as in yeast). The only exception to this is PC from *Pseudomonas citronellolis*, which is comprised of four pairs of non-identical subunits ($\alpha 4$, $\beta 4$). The alpha subunit is approximately Mr 65,000 and is biotinylated, where as the beta subunit is approximately Mr 54,000.

Although X-ray diffraction of protein crystals and more recently 2-dimensional NMR are the preferred methods of precisely determining the tertiary and quaternary structures of proteins, PC has so far not been amenable to study by these techniques as discussed further in section 4.1. Other methods, therefore, of defining the structure of PC have been utilised, including chemical modification, immunochemistry and electron microscopy.

Tetrameric PC is sufficiently large to enable visualisation under the electron microscope after negative staining with uranyl acetate. This technique has been used extensively to study the structure of PC. Initial experiments by Valentine *et al.* (1966) described the enzyme as square planar in shape. It was later shown (Goss, *et al.*, 1979, Cohen *et al.*, 1979) that this molecule was in fact a highly visible contaminant of these early preparations, and in 1980 Mayer *et al.* provided data which was consistent with a structure representing a tetrahedron-like molecule.

The tetramer was made up of 2 dimers, arranged orthogonally, joined at their convex surfaces as shown in figure 1.5. A cleft was clearly visible along the longitudinal axis of each monomer, and samples from liver mitochondria of chicken, sheep and rat were seen to be indistinguishable. Acetyl-CoA was observed to enhance the preservation of the tetrameric form of the enzyme, producing a "tight" conformation in which the cleft was less visible.

The location of the biotin moiety and hence the active site of PC was subsequently determined by observing complexes between PC and avidin using the electron microscope. In an elegant experiment, Johannssen *et al.* (1983) found that at a molar ratio of 1:1 PC could complex with avidin to form long unbranched chains. As the biotin binding sites of the tetrameric avidin molecule (Mr 68,000) have been characterised (Green *et al.*, 1971; Safer *et al.*, 1982; Gitlin *et al.*, 1988), it was clear that such chains could only form if the site of biotin attachment to PC was along the longitudinal axis, no further than 2-3 nm from the intersubunit junction (as shown in figure 1.6).

Electron microscopy conducted on PC samples from a number of invertebrate species has shown that the enzymes from *Aspergillus nidulans* (Osmani *et al.*, 1984), *S. cerevisiae* (Rohde *et al.*, 1986), *Rhizopus arrhizus* (Mayer *et al.*, 1985), *Bacillus thermophilus* (Libor *et al.*, 1979) and *Ps. citronellolis* (Fuchs *et al.*, 1988) all possess the same quaternary structure as the liver enzymes. This last finding is of particular interest (as discussed in section 6.3) given that PC from *Ps. citronellolis* is composed of 2 different subunits, and previous studies by Cohen *et al.* (1979) and particularly Goss *et al.* (1981), provided evidence to suggest that the structure of this PC was significantly different to that from the mammalian species.

1.4.3 1st Partial Reaction Subsite

As can be seen from equation 1.5, the first partial reaction involves the carboxylation of biotin with the concomitant cleavage of ATP. The structure of this reaction centre must therefore be such that ATP and HCO_3^- can bind and interact with the biotin prosthetic group.

Comparing the binding affinities of the magnesium salts of various nucleoside triphosphates, Scrutton and Utter (1965) found that inosine triphosphate (ITP) bound to chicken PC with a 6-7 fold lower affinity than ATP. These two ligands differ only in the substitution at the C-6 position of the purine ring (ITP: [C=O] for ATP: [C-NH_2]) suggesting that the amino group present in ATP undergoes hydrogen bonding with an appropriate basic residue in the active site.

The presence of a lysine residue in the 1st partial reaction subsite has been implicated by affinity labelling studies with a 2´3´ dialdehyde derivative of ATP (oATP; Easterbrook-Smith *et al.*, 1976). These authors were able to show that reduction with $NaBH_4$ of the MgoATP-PC complex resulted in modification of a single lysine residue. In

a similar study of yeast PC, Lim (1988) was able to show that oATP was a competitive inhibitor with respect to ATP prior to reduction and subsequently bound irreversibly after reduction with $NaBH_4$. Additionally, the stoichiometry of labelling was shown to extrapolate to 1 mole of label incorporated per mole of enzyme at zero residual activity, which is consistent with the modification results for sheep PC described above (Easterbrook-Smith *et al.*, 1976). As shown in figure 1.2 the role of such a lysine residue could be to stablise the enol form of the biotin moiety (Attwood and Keech , 1984).

While the identity of the catalytic base shown in figure 1.2 is far from certain, studies of the analogous reaction in the biotin carboxylase of *E. coli* (Tipton and Cleland, 1988) implicate a cysteinyl residue (pKa > 9.5) and an additional base with a pKa of 6.6. The finding of Palacian and Neet (1972) that ATP is able to partially protect chicken PC from inactivation by N-ethylmaleimide is consistent with the presence of a catalytic cysteine residue in this reaction subsite.

1.4.4 2nd Partial Reaction Subsite

At the 2nd partial reaction subsite the carboxy-biotin intermediate is broken down enabling the carboxylation of pyruvate as indicated in equation 1.6. This site therefore must be able to bind the carboxy-biotin moiety, pyruvate and oxaloacetate. In addition, the presence of a tightly bound divalent metal cation has been demonstrated in this region (for a review see Wallace and Easterbrook-Smith, 1985). In PCs from chicken, sheep and rat, the metal ion is found to be manganese [Mn(II)] while PC from yeast has a tightly bound zinc ion [Zn(II)].

Much of the information regarding the structure of the 2nd partial reaction subsite has come from electron spin resonance (ESR) and nuclear magnetic resonance (NMR) spectroscopy, and chemical modification studies. As indicated in figure 1.4 (and discussed further in section 3.3.3) this region of PC has a high level of sequence conservation with the transcarboxylation subunit (TC5S) of *Propionibacterium shermanii* transcarboxylase and the α-subunit of *Klebsiella pneumoniae* oxaloacetate decarboxylase consistent with their binding pyruvate and oxaloacetate. The considerable amount of study into the structure of the TC5S subunit is relevant therefore to consideration of the structure of PC.

Read and Scrutton (1974) reinterpreted the proton spin relaxation time data of Mildvan *et al.* (1968) to estimate the distance between the Mn(II) and pyruvate methyl-protons in PC to be 0.6 - 0.7 nm. This is consistent with similar studies by Fung *et al.*

(1974) who measured the distances between the bound paramagnetic ion cobalt (II) and pyruvate in TC 5S and concluded that pyruvate was not sufficiently close to be an inner sphere ligand of the metal ion. This has been confirmed for PC by the metal chelation studies of Carver *et al.* (1988) who proposed that the Mn(II) plays more of a structural role, enhancing the stability of the tetrameric enzyme.

The chemical modification studies of Hudson *et al.* (1975) showed treatment of sheep PC with bromopyruvate resulted in the rapid loss of catalytic activity, with protection afforded by the presence of oxaloacetate. Pyruvate was less effective than oxaloacetate in protecting against modification, although the combination of pyruvate and acetyl-CoA enhanced this protection. These authors also demonstrated that modification occurred at one cysteinyl residue, most likely in the 2nd partial reaction subsite, and to a lesser extent an additional residue in the 1st partial reaction subsite. These findings are consistent with those of Attwood *et al.* (1986b) who proposed that the mechanism of the second partial reaction was base catalysed, and that a cysteinyl sulphydryl represented the most likely candidate for such a base. Attwood *et al.* (1986b) further suggested that for ionisation to occur in such a group at neutral pH, it must have a pKa below 5, which in turn requires that it be in an ion pair with another base such as lysine, arginine or histidine.

Kumar *et al.* (1988b) using 2,4-dinitrophenylsulphenyl chloride (DNPS-Cl) were able to specifically label a tryptophan residue of TC5S (trp^{73}) in a pyruvate-protectable manner. This residue was identified after sequence analysis of a tryptic peptide isolated from labelled enzyme by reverse phase high performance liquid chromatography. Tryptophan had previously been implicated at or near the binding site of pyruvate by fluorescence quenching experiments using N-bromosuccinimide (NBS) labelled TC (Kumar *et al.*, 1988a).

1.5 INTERACTION WITH ACETYL-COA

The interaction between PC and acetyl-CoA is an especially interesting facet of this enzyme in that *in vitro* , acetyl-CoA has a variety of significant effects on PC including enhancing stability, enhancing enzyme activity and altering enzyme conformation.

1.5.1 Effect of Acetyl-CoA on Enzyme Stability and Conformation

Irias *et al.* (1969) showed that saturating levels of acetyl-CoA could prevent chicken PC from undergoing inactivation due to low temperature, and Ashman *et al.* (1972)

showed that similar levels of acetyl-CoA could protect sheep PC against inactivation upon dilution below 4 U/ml. Since both these phenomena have been shown to involve dissociation of tetramers to monomers which then associate irreversibly into inactive aggregates, it is likely that the stabilising influence of acetyl-CoA is effected by reducing the rate at which these tetramers dissociate to form monomers (Khew-Goodall, 1985). This is confirmed by a number of lines of evidence including the observations of Mayer *et al.*, (1980) that acetyl-CoA enhances the preservation of the tetrameric state, during mounting for electron microscopy.

Frey and Utter (1977) were able to observe a change in the difference spectrum of chicken PC when mixed with acetyl-CoA, with a small maximum at 280nm, indicating a subtle conformational change occurs upon binding. A change in the shape of tetramers was also noted by Attwood *et al.* (1986a) who showed that PC formed long chains when complexed with avidin only in the presence of acetyl-CoA (or high concentrations of pyruvate).

1.5.2 Effect of Acetyl-CoA on Enzyme Activity

The original proposal by Keech and Utter (1963) that acetyl-CoA is a physiologically significant regulator of PC, led Kornberg (1966) to propose that this allosteric regulation provided the mechanism by which the levels of oxaloacetate entering the citric acid cycle were coordinated with the levels of acetyl-CoA.

In vitro, PC is allosterically activated by acetyl-CoA although this effect shows variation with species. PC isolated from chicken liver is absolutely dependent on acetyl-CoA for activity (Utter and Keech, 1960). Under standard assay conditions and in the absence of acetyl-CoA, rat liver, (McClure *et al.*, 1971) sheep liver, (Ashman *et al.*, 1972) and thermophilic Bacillus PC, (Libor *et al.*, 1978) all exhibit ~2% of the acetyl-CoA stimulated activity . However, if the concentrations of pyruvate, ATP and HCO_3^- are substantially increased, this activity can approach 25% of the acetyl-CoA activated level.

The activity of PC from yeast is less dependent on acetyl-CoA, in that in the absence of the ligand, significant catalysis can be observed, although stimulation of between 4 and 10-fold can be achieved in buffers containing low concentrations of HCO_3^- and K^+. The acyl-CoA binding site of the yeast enzyme is also less selective than that of the liver enzyme, in that activation can be effected by acyl-CoA compounds with a wide range of acyl chain lengths (2-20 carbons; Myers *et al.*, 1983).

Although Osmani *et al.* (1981) concluded that the activity of PC from *Aspergillus nidulans* was enhanced by acetyl-CoA, it was only observed as a reversal of the inhibitory effects of L-aspartate and 2-oxoglutarate. As aspartate is known to promote the dissociation of tetramers of yeast (Rohde *et al.*, 1986) and *A. nidulans* PC (Osmani *et al.*, 1985) it is likely that the influence of acetyl-CoA on *A. nidulans* PC is at the level of tetramer stability alone and not directly involved in catalysis.

The enzymes from *Pseudomonas citronellolis* and *Aspergillus niger* are totally independent of acetyl-CoA (Utter *et al.*, 1975).

In vivo, acetyl-CoA has long been thought to regulate the levels of PC activity in the mitochondria of gluconeogenic tissues, particularly the liver, although the extent of acetyl-CoA's role in lipogenic tissues is less well defined (for a review see, Barritt, 1985). A more recent study (Griffith, *et al.*, 1989), aimed at characterising the role of benzoate administration to combat hyperammonemia, has confirmed this hypothesis in that addition of sodium benzoate inhibits PC activity in isolated rat mitochondria.

When sodium benzoate was added to intact mitochondria, intra-mitochondrial concentrations of benzoyl-CoA were shown to increase with a concomitant decrease in acetyl-CoA concentrations and pyruvate-dependent incorporation of $^{14}CO_2$. Under these conditions, no direct influence of sodium benzoate on PC activity was observed *in vitro*. These results are consistent with a study by Gatley (1977), who measured the effect of benzoate on ureogenesis and gluconeogenesis in isolated rat hepatocytes.

1.5.3 Studies on the Acetyl-CoA Binding Site

During attempts to study the mechanism of acetyl-CoA activation of pyruvate carboxylase, Keech and Farrant (1968) observed that amino-selective reagents fluorodinitro benzene (FDNB), trinitrobenzene sulphonate (TNBS) and potassium cyanate irreversibly inhibited sheep PC by modification of a single catalytically important residue, identified to be a lysine in the case of FDNB.

Ashman *et al.* (1973) showed that modification of a single lysine residue by TNBS inhibited the acetyl-CoA dependent activity, but enhanced the rate of the independent activity. That is, the second partial reaction (equation 1.6) was unaffected, but the first partial reaction (equation 1.5) had been desensitised to acetyl-CoA. It was shown (Ashman *et al.*, 1973; Chapman-Smith, 1981) that limited treatment of sheep PC in fact modified 2

lysine residues, one catalytically important residue (Lys-A) as well as another residue (Lys-B) which has no detectable effect on activity.

A number of species of PC have been treated similarly with TNBS, including rat and yeast (Scrutton and White, 1973), *A. nidulans* (Osmani *et al.*, 1981) and *Bacillus thermophilus* (Libor *et al.*, 1978), with the kinetics being very similar in all cases.

Protection studies with acetyl-CoA (Ashman *et al.* 1973) showed that saturating levels of acetyl-CoA could protect the sheep enzyme from loss of activity when treated with TNBS. Two models were proposed to explain this, either Lys-A is involved with acetyl-CoA binding (hence the bulky chemical modification would sterically hinder acetyl-CoA binding), or Lys-A is distinct from the acetyl-CoA binding site, but is involved in mediating the allosteric effect of acetyl-CoA. No clear evidence supporting either model has been provided from analogue studies performed by Scrutton *et al.* (1977) and Chapman-Smith (1981), who came to different conclusions regarding the role of Lys-A in the activation of the enzyme. Ashman *et al.* (1973) argued that acetyl-CoA could bind to TNBS-inactivated sheep PC, because the deacylation activity associated with these preparations was unaffected by the treatment. Although the PC samples used in these studies represented highly purified material, this argument was hampered by the absence of definitive evidence that this deacylation activity was an integral part of the enzyme itself.

The identity of Lys-A could not be established from these studies, due to the insensitive nature of the amino acid sequencing technology available at that time (Chapman-Smith, 1981).

1.6 PROJECT RATIONALE

PC is a large, complex and metabolically important protein that has been the subject of study by many groups since its initial description in 1960. Much of the work has revolved around powerful kinetic studies, which have provided a significant understanding of the chemical mechanism of the reaction. However, these studies have been unable to explain how the structure and amino acid composition of the enzyme enables this critical chemistry to proceed.

Clearly required to complete the characterisation of this enzyme (as with any other), are details of the amino acid sequence, the conformation taken by the polypeptide chain in three dimensions, and the functional groups on the amino acid side chains

particularly important for activity. Due to the absence of X-ray crystallographic data, the elucidation of the tertiary structure of PC must be derived from other methodologies such as sequence analysis, chemical modification, immunochemistry and electron microscopy.

At the outset of this study, little amino acid sequence information was available for any of the members of the biotin carboxylase family, and for pyruvate carboxylase in particular, as only the sequence surrounding the biotin attachment site (Rylatt *et al.*, 1977) and the carboxy-terminal 44 amino acids of human PC (Freytag & Collier, 1984) had been published. Work initiated in this laboratory subsequently provided the first complete amino acid sequence of pyruvate carboxylase (from yeast, Lim *et al.*, 1988) and partial cDNA and genomic clones for pyruvate carboxylase from rat (Cassady, 1987). Although the isolation and sequencing of the yeast PC gene provided information regarding the functional significance of some regions of the molecule, the identity of the acetyl-CoA binding site remained unsolved.

Lopez-Casillas *et al.* (1988) speculated that a region of 30 amino acids in rat ACC is involved in binding acetyl-CoA. This proposal was based, however, on a suspiciously low level of sequence similarity to a region of the acetyl-CoA binding site of citrate synthase [EC 4.1.3.7], the structure of which is known (Remington *et al.*, 1982). No counterpart for this sequence, can be found in yeast PC, although this may be due to the different roles acetyl-CoA takes in these two enzymes, that of substrate (ACC) and activator (PC).

Another consequence of the isolation and characterisation of the yeast PC gene within this laboratory, was the possibility of site-directed mutagenesis and expression of mutant PC proteins in a homologous host. It was clear therefore, that regions and residues implicated by chemical studies, as important for binding acetyl-CoA, or maintaining enzyme structure could then be tested through mutagenesis.

1.7 AIMS

(1) To provide further amino acid sequence data for pyruvate carboxylase from rat PC, so as to enable more complete inter-species comparisons, in addition to inter-enzyme comparisons. In this way it is hoped to gain an insight into which residues, or regions of PC are important for activity. Such information,

combined with the relative positions of exons, provided by sequencing of the rat PC gene, should also enable a more rigorous testing of Lynen's hypothesis.

(2) In view of the absence of X-ray diffraction, refine the structural model for PC using electron microscopy, immuno-chemistry and partial proteolysis.

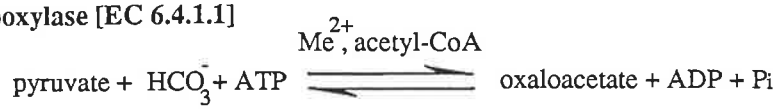
(3) Localise the acetyl-CoA binding site, within this refined model, by affinity labelling or other chemical modification techniques.

Table 1.1 The Reactions Catalysed by the Biotin-Dependent Carboxylases

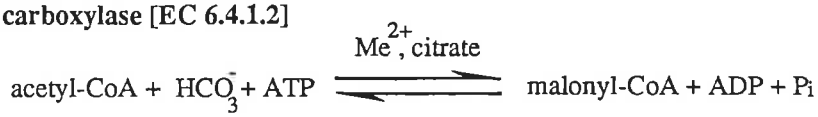
The reactions catalysed by the biotin dependent carboxylases are shown (Wood and Barden, 1977; Dimroth, 1980; Hilpert and Dimroth, 1982; Buckel and Semmler, 1982) along with known cofactors. Me^{2+} represents a divalent metal cation, acetyl-CoA is an allosteric activator of pyruvate carboxylase and citrate is an allosteric activator of acetyl-CoA carboxylase (Lane *et al.*, 1974).

Carboxylases

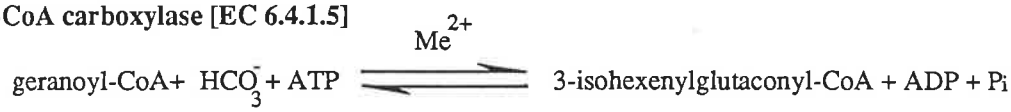
pyruvate carboxylase [EC 6.4.1.1]



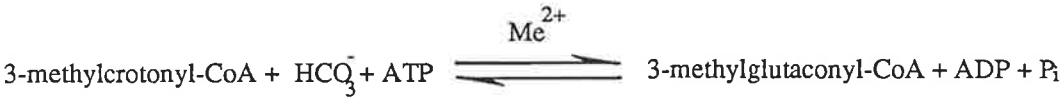
acetyl-CoA carboxylase [EC 6.4.1.2]



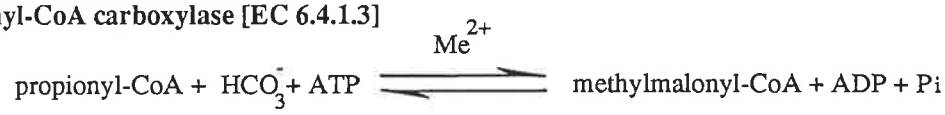
geranoyl-CoA carboxylase [EC 6.4.1.5]



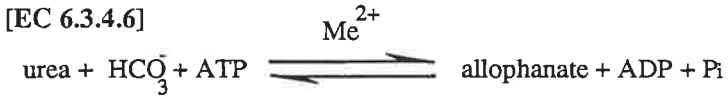
3-methylcrotonyl-CoA carboxylase [EC 6.4.1.4]



propionyl-CoA carboxylase [EC 6.4.1.3]



urea carboxylase [EC 6.3.4.6]



Decarboxylases

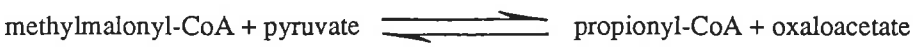
glutaconyl-CoA decarboxylase

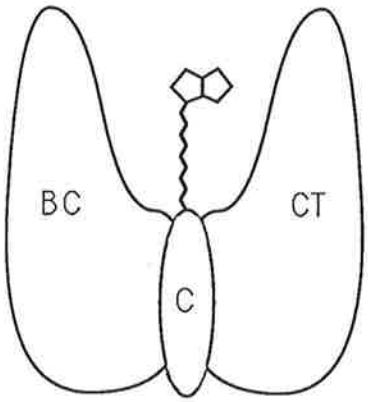
methylmalonyl-CoA decarboxylase[EC 4.1.1.41]

oxaloacetate decarboxylase [EC 4.1.1.3]

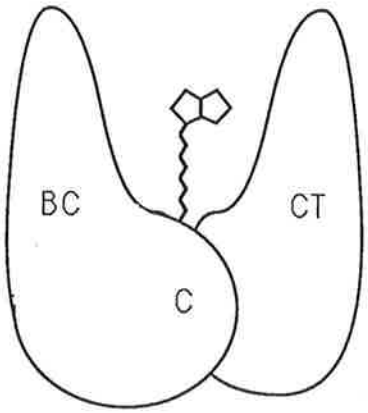
Transcarboxylase

transcarboxylase [EC 2.1.3.1]

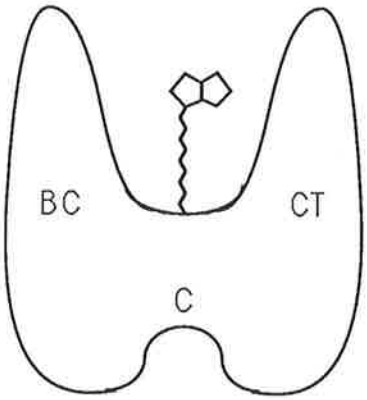




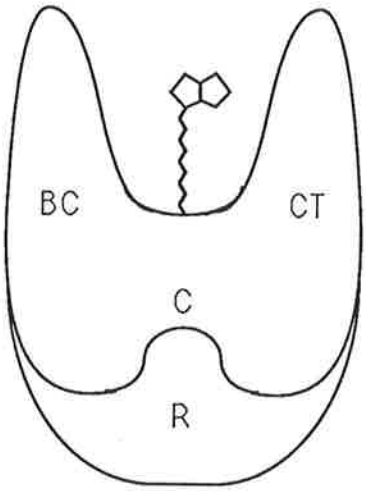
acetyl-CoA carboxylase (*E. coli*)
trans-carboxylase (*P. shermanii*)



β-methyl crotonyl-CoA carboxylase (*Achromobacter*)
pyruvate carboxylase (*P. citronellolis*)



pyruvate carboxylase (*A.. nidulans*)
acetyl-CoA carboxylase (yeast)



pyruvate carboxylase (avian, mammalian)
acetyl-CoA carboxylase (avian, mammalian)

Figure 1.1 Subunit structure of biotin carboxylases
Stylised diagram adapted from Lynen (1975), showing the four categories of subunit composition of biotin carboxylases.

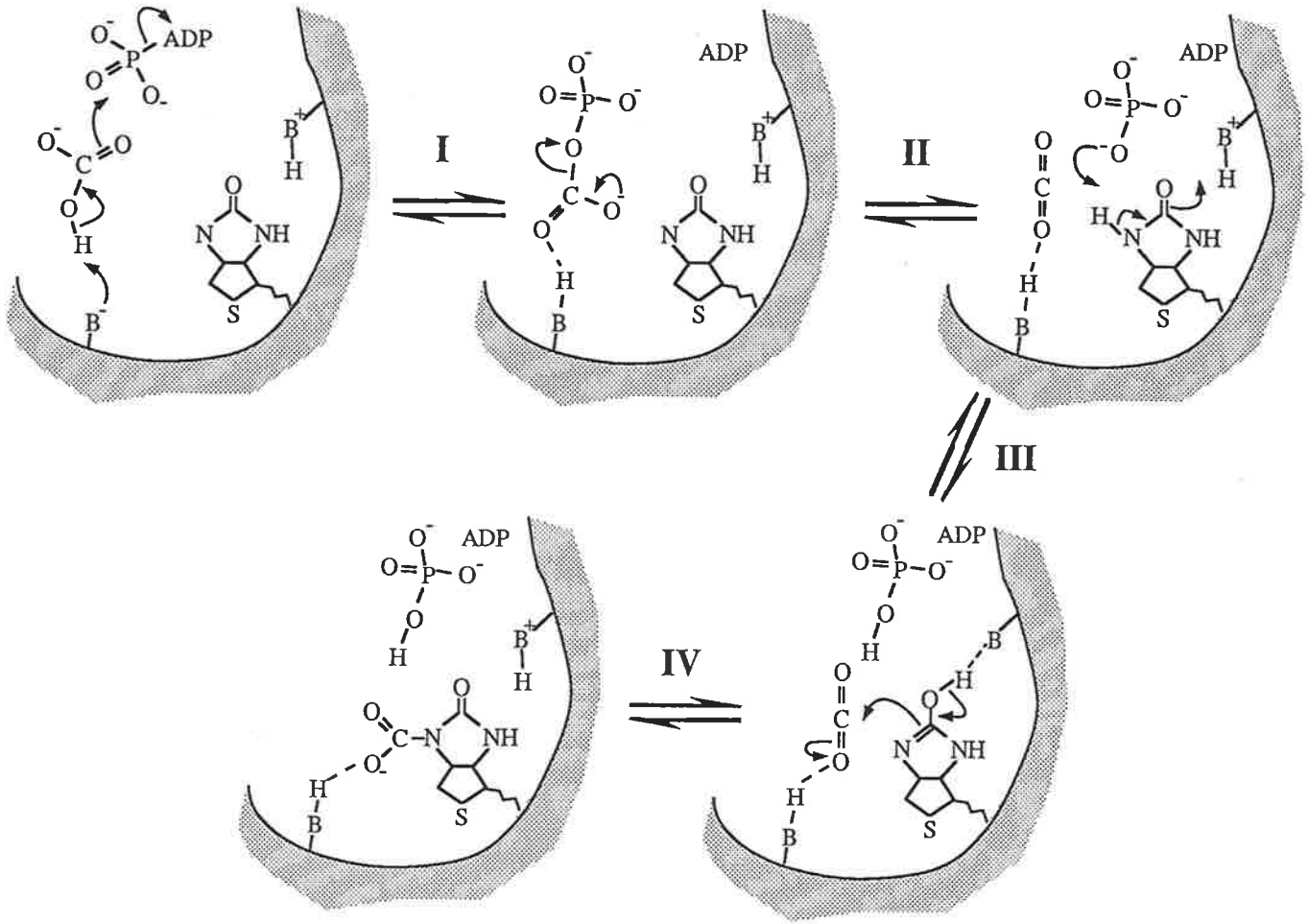


Figure 1.2 Mechanism of the First Partial Reaction. Adapted from the reaction scheme proposed by Knowles (1989), showing the carboxylation of the biotin moiety at a stylised first partial reaction subsite.

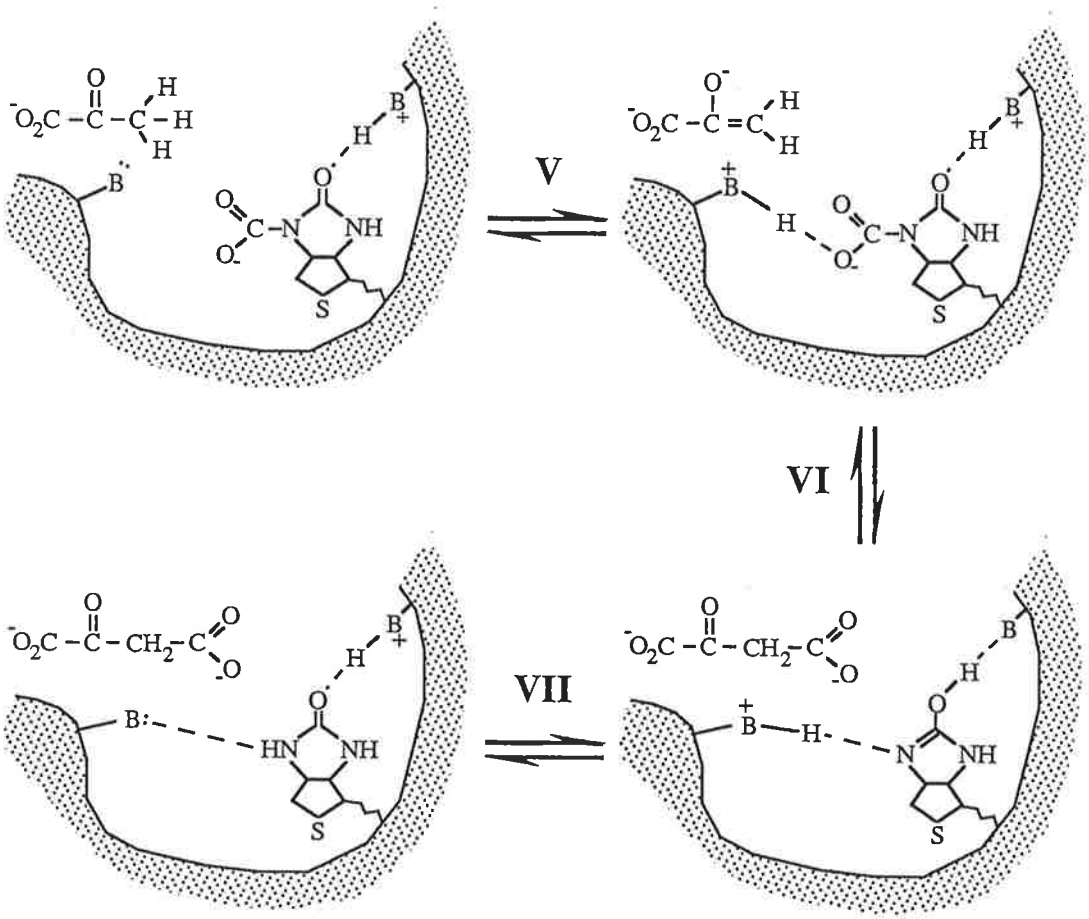
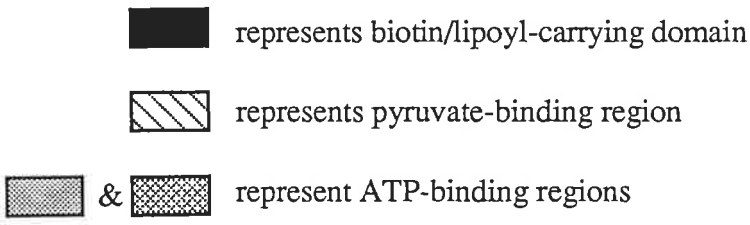


Figure 1.3 Mechanism of the Second Partial Reaction. Adapted from the reaction scheme proposed by Attwood *et al.* (1986b), showing the decarboxylation of the biotin moiety at a stylised second partial reaction subsite.

Figure 1.4 Schematic Representation of Sequence Conservations Observed between Biotin Carboxylases and Other Enzyme Groups

Regions of sequence homology between members of the biotin-dependent carboxylases and other enzyme classes, updated and adapted from Lim *et al.* (1988). Sequences are from:
PC (yeast) - Lim *et al.* (1988)
rat acetyl-CoA carboxylase - Lopez-Casillas *et al.* (1988)
chicken acetyl-CoA carboxylase - Takai *et al.* (1988)
rat propionyl-CoA carboxylase (α) Browner *et al.* (1989)
& (β) Kraus *et al.* (1986)
human propionyl-CoA carboxylase - (α) Lamhonwah *et al.* (1989)
& (β) Lamhonwah*et al.* (1986)
oxaloacetate decarboxylase - (α) Schwarz *et al.* (1988)
& (β,γ) Laußermair *et al.* (1989)
Transcarboxylase - (12S) Thornton *et al.* (1987), (5S) Samols *et al.* (1988) & (1.3S) Maloy *et al.* (1979)
yeast carbamyl phosphate synthetase - (large) Lusty *et al.* (1983) & (small) Nyunoya & Lusty (1984)
rat carbamyl phosphate synthetase - Nyunoya *et al.* (1985)
E. coli pyruvate dehydrogenase - Guest *et al.* (1985)



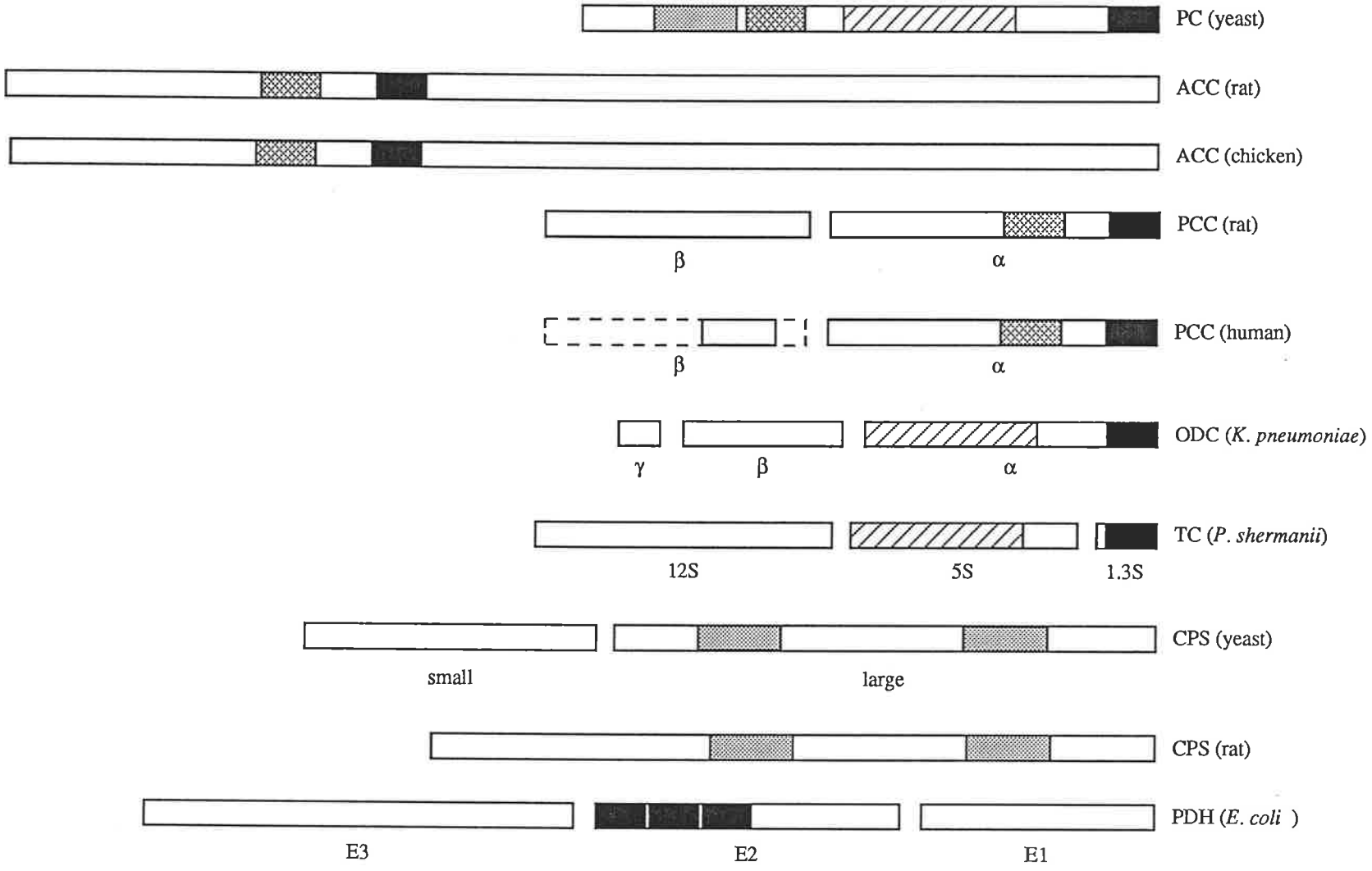

PC (yeast)
ACC (rat)
ACC (chicken)
PCC (rat)
β
α
PCC (human)
β
α
ODC (K. pneumoniae)
γ
β
α
TC (P. shermanii)
12S
5S
1.3S
CPS (yeast)
small
large
CPS (rat)
PDH (E. coli)
E3
E2
E1

Figure 1.5 A Model for the Quaternary Structure of Pyruvate Carboxylase from Vertebrates

The clay models shown represent the tetrameric structure of pyruvate carboxylase as interpreted from electron micrographs of chicken, sheep and rat pyruvate carboxylase by Mayer *et al.* (1980).

Series 13 represent the low-resolution model. Plates 1-4 show views from various directions. Plate (5) shows a partially flattened model.

Series 14 represent a high-resolution model, indicating the presence of a longitudinal cleft. Plates 1-5 show views from various directions. Plates 6 & 7 show some degree of flattening and Plate 8 represents completely flattened and disintegrated particles.[reprinted from Mayer *et al.* (1980) with permission from the authors]

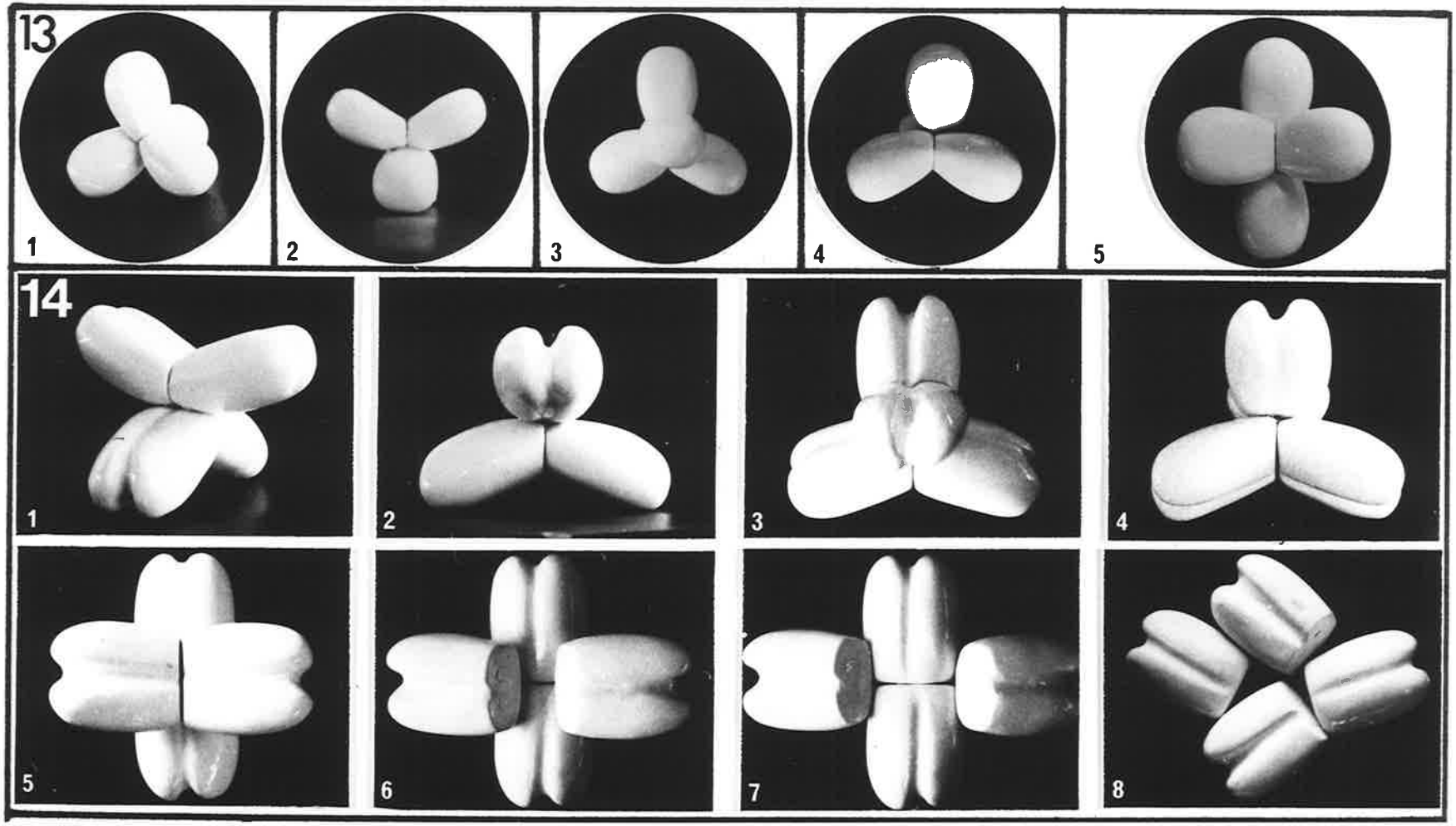
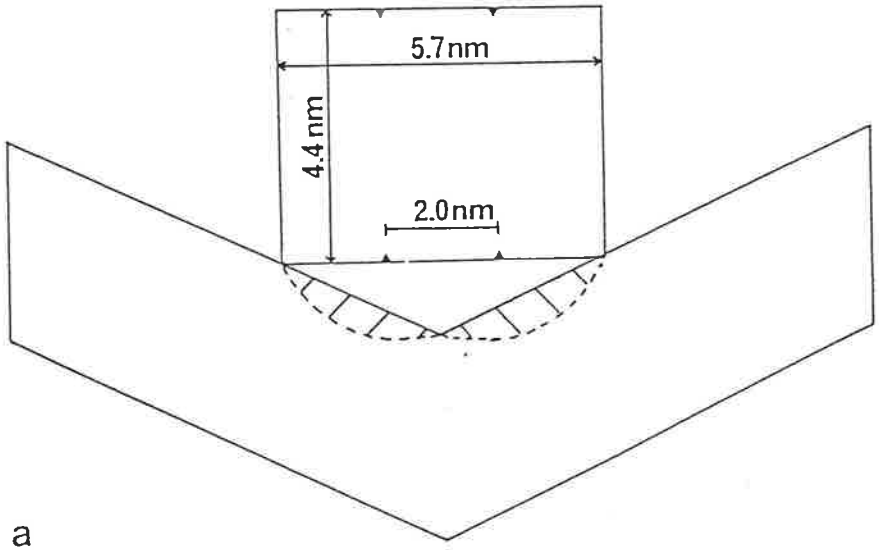

13
1
2
3
4
5
14
1
2
3
4
5
6
7
8

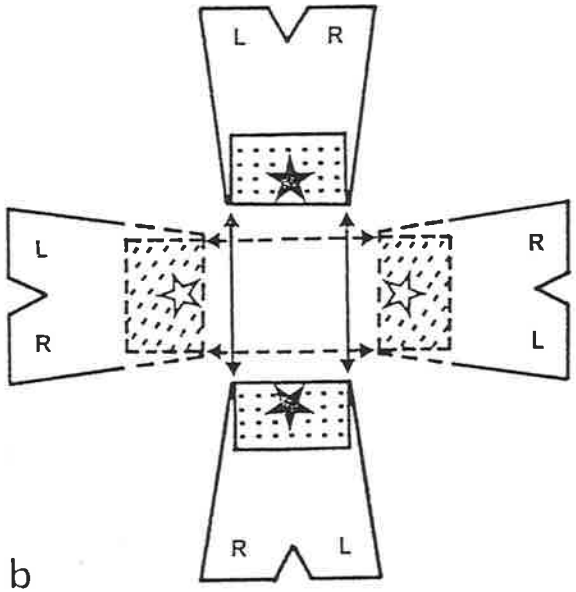
Figure 1.6 Diagramatic Views of a Complex between Avidin and Pyruvate Carboxylase

(a) A side view of a complex between an avidin tetramer and one half of a pyruvate carboxylase tetramer indicating the dimensions of avidin. The biotin-binding sites on avidin are indicated (▲) along with the region expected to contain the biotin moiety of pyruvate carboxylase (shaded).

(b) An exploded face-view of a tetramer of pyruvate carboxylase, indicating the region of contact with avidin (shaded). The location of biotin within each subunit is indicated. (★) represents the biotin moiety on the upper pair of subunits and ☆ represents the biotin moiety on the lower pair of subunits. Reproduced from Johannssen *et al.* (1983) with permission.



a



b

CHAPTER 2

MATERIALS & METHODS

2.1 MATERIALS

2.1.1 Enzymes and Proteins

Bovine serum albumin, trypsin (TPCK treated)[EC 3.4.21.4], chymotrypsin [EC 3.4.21.1], proteinase K, papain [EC 3.4.22.2], malate dehydrogenase [EC 1.1.1.37], avidin (egg white), anti-mouse and anti-rabbit IgG coupled to alkaline-phosphatase, and molecular weight markers (SDS-6H) were purchased from Sigma Chemical Co., St. Louis, MO., USA. T_4 DNA ligase [EC 6.5.1.1] , *E. coli* DNA polymerase I (Klenow fragment), avidin-alkaline phosphatase conjugate were obtained from Bresatec Ltd., Adelaide, South Australia. Calf intestinal phosphatase [EC 3.1.3.1] and Endoproteinase-Arg-C were purchased from Boehringer Mannheim Australia, Sydney, Australia. Restriction endonucleases were principally from New England Biolabs Inc., MA, USA. and Pharmacia LKB Biotechnology, Uppsala, Sweden. Monoclonal antibodies raised against native sheep PC were a kind gift of Dr. K. Carey and polyclonal antiserum raised against TNP-BSA was supplied by Ms. J. Brazier.

2.1.2 Radio-chemicals

[α-^{32}P] dATP, [γ-^{32}P] ATP and [α-^{32}P] dCTP (3000 Ci/mmol) were purchased from Bresatec Ltd., Adelaide, South Australia. $NaH^{14}CO_3$ (50-60 mCi/mmol) and D[carbonyl-^{14}C]biotin (40-60 mCi/mmol) were from Amersham, England.

2.1.3 General chemicals

ATP (disodium salt, grade I), acrylamide, agarose (type I), ampicillin, chloramphenicol, coenzyme A, Coomassie Brilliant Blue (R_{250}, G_{250}), D-biotin, 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (BCIG), 5-bromo-4-chloro-3-indolyl phosphate (BCIP), dithioerythritol (DTE), ethylenediaminetetraacetic acid (EDTA), ethidium bromide, kanamycin, L-lysine, $MgCl_2$, 2-mercaptoethanol, NADH, N-ethyl morpholine, N,N´´-methylene-bis-acrylamide, nitro blue tetrazolium (NBT), phenylmethylsulphonyl fluoride (PMSF), polyethylene glycol (PEG) 6000 & 8000, sodium pyruvate (type II, dimer free), sodium dodecyl sulphate (SDS), tetracycline, trinitrobenzene sulphonate (TNBS), were supplied by Sigma Chemical Co., St. Louis, MO., USA.

2.1.4 Kits for Molecular Biology

Oligonucleotide labelling kits (Cat. No. OLK-A) and dideoxynucleotide sequencing kits (Cat. No. DSK-A) were purchased from Bresatec Ltd. Gene-ATAQ™ taq polymerase sequencing kits (Cat. No. 27-1688-01) were purchased from Pharmacia LKB Biotechnology.

2.1.5 Chromatographic Media

Sephadex G-25, DEAE-Sephacel, protein-A-Sepharose and cyanogen-bromide-activated Sepharose CL-4B were purchased from Pharmacia LKB Biotechnology. Reverse-phase columns, C4 (4.6 x 30mm - Cat. No. B03-GU) and C4 (2.1 x 30mm - Cat. No. B03-032) were from Brownlee Labs. The TSK 3000 SW gel filtration column (0.5cm I.D. x 60cm) was obtained from Toya Soda, Japan.

2.1.6 Clones and Bacterial Strains

Rat genomic clone λRG1.2 and rat cDNA clone λRL2.53 were kind gifts of Dr. A. I. Cassady.

E.coli XL1-B	[*sup E44 hsdR17 recA1 endA1 gyrA46 thi relA1 lac* $^-$ (F′*proAB*$^+$ *lac* I^q *lac* ZΔM15 Tn10 (*tet* ^{r}))]
E.coli JM101	[Δ(*lac, pro*) *supE, thi.* F′ *traD36 proAB, lac* I^q *lac* ZΔM15]
E.coli AA102F′lac/low	[*recA pro thi supE endA hsdR tet* ^{R}Δ1(*gal-chl -pgl-att*) (F′*traD36, proAB, lac* I^q *lac* ZΔM15)]

2.1.7 Culture Media

Media were prepared using distilled water, and then autoclaved prior to the addition of any antibiotics. Media plates were prepared by the inclusion of 12 g of agar per litre.

L-broth	10 g of bacto-tryptone, 5 g of bacto-yeast extract, 10 g of NaCl in a final volume of 1 litre.

2 x YT- medium	16 g of bacto-tryptone, 10 g of bacto-yeast extract, 5 g of NaCl in a final volume of 1 litre.
MacConkey-galactose	40 g of MacConkey agar base and 10 g galactose in a volume of 1 litre.

2.1.8 Oligonucleotides

RPC02 5′-CCAGAGGAGTCCCTTTGGTAC-3′ (Bresatec Ltd., Adelaide.)
Probe directed against the coding sequence of the rat PC cDNA (see figures 3.5 & 3.6).

TSP 5′-AAAACACCATCATACACTAA-3′ (Bresatec Ltd., Adelaide.)
Sequencing primer for use with the Sequenest II™ transposon mediated deletion kit.

2.2 METHODS

2.2.1 Purification of pyruvate carboxylase

Pyruvate carboxylase was purified from freeze-dried rat or sheep mitochondria by the method of Goss *et al.* (1979), except that DEAE-Sephacel replaced DEAE-Sephadex as the chromatographic matrix. Avidin-sepharose affinity chromatography (Henrickson *et al.*, 1979) was then used to produce homogeneous enzyme. PC from vertebrate sources was stored at -80°C in sucrose storage buffer (0.1M N-ethylmorpholine-HCl, pH 7.2, 1.6 M sucrose).

Yeast PC was purified from baker's yeast (*Saccharomyces cerevisiae*) as described by Lim *et al.* (1988). Yeast PC was stored at -80°C in 100 mM Tris-Cl pH 7.2 or 100 mM N-ethylmorpholine-HCl pH 7.2.

2.2.2 Activity assay

PC activity was routinely measured using the enzyme-linked spectophotometric assay described by Utter and Keech (1963), which measures the change in absorbance at 340nm as NADH is used to convert oxaloacetate to malate, in the presence of saturating levels of malate dehydrogenase.

Whenever a large number of samples were to be processed, such as in kinetic experiments, the radio-chemical assay of (Keech and Farrant, 1968) was used. This

procedure measures the rate of incorporation of radioactivity from $H^{14}CO_3^-$ into an acid stable form. All assay reactions contained 250 μM acetyl-CoA unless otherwise stated.

One unit of PC is defined as catalysing the carboxylation of 1 μmole of pyruvate per minute at 30°C. Rat and sheep PC were assayed at pH 8.4 and yeast PC was assayed at pH 8.0.

2.2.3 Acetyl-CoA-independent activity assay

Under conditions of elevated salt and HCO_3^- concentrations sheep and rat PC are able to catalyse the carboxylation of pyruvate in the absence of acetyl-CoA at a rate approaching 25% of that observed in the presence of acetyl-CoA. This activity is termed the acetyl-CoA-independent activity and was measured using the radiochemical assay developed by Ashman *et al.* (1972).

2.2.4 Protein determination

Protein concentrations were routinely measured by the dye-binding method of Bradford (1976), as adapted for microtitre plates by Rylatt and Parish (1982). A standard curve was determined for each assay using bovine serum albumin, except when measuring concentrations of IgG or Fab solutions, where Gamma-globulin standards were used. Absorbance measurements of microtitre plates were made using a Titertek™ multiscan MC (Flow Laboratories).

During the purification of PC, protein concentration was approximated by measuring the u.v. absorbance of the samples at 260nm and 280nm. These values were then substituted into the equation of Layne (1957), giving a value in mg/ml.

$$\text{protein concentration (mg/ml)} = 1.55 \times A_{280\,nm} - 0.76 \times A_{260\,nm}$$

2.2.5 Microcentrifuge buffer exchange

Buffer exchanges on small volumes of enzyme (20 - 100 μl) were effected using the microcentrifuge gel filtration technique of Helmerhorst and Stokes (1980). The enzyme was centrifuged for 2 minutes at 1500 x g through a 2 ml bed volume of Sephadex G-25 which had been pre-equilibrated in the buffer of choice. This method was particularly useful for buffer exchange of enzyme from sucrose storage buffer, which is used to maintain vertebrate PC stability at low temperatures.

2.2.6 Synthesis of acetyl-CoA

Acetyl-CoA was synthesised from reduced CoA and acetic anhydride, using a procedure modified from Simon and Shemin (1953). Reduced CoA (130 µmoles) was dissolved in 5 ml of H_20 and the pH adjusted to 7.2 with potassium bicarbonate. Acetic anhydride (500 µmoles) was then added and the solution left at room temperature for 10 minutes. The mixture was then acidified by the addition of 2-3 drops of 6M H_2SO_4. Any remaining reactants were removed by 3 ether extractions, after which the aqueous phase was returned to neutrality with NaOH. The purity of the acetyl-CoA solution was measured by comparison of the absorbances at 232 nm and 260 nm. Samples used for routine enzyme assays had a 232nm/260nm ratio between 0.5 and 0.55 (Dawson *et al*, 1969). Aliquots of 1ml were stored frozen at a concentration of 5mM. When samples of more highly purified acetyl-CoA were required, such as for binding assays, further purification through a high performance liquid chromatography strong anion exchange column (HPLC-SAX) was undertaken. This involved injecting the sample onto a HPLC-SAX column (Brownlee Labs) pre-equilibrated in 5 mM triethylamine acetate/10% (v/v) acetonitrile, followed by elution with a linear gradient to 450 mM triethylamine acetate/10% (v/v) acetonitrile over 20 minutes. Acetyl-CoA, coenzyme-A, ATP and 3´dephospho-coenzyme-A can be resolved using this technique (Booker *et al.*, 1988).

2.2.7 Synthesis of [^{14}C] acetyl-CoA

To synthesise [^{14}C] acetyl-CoA, [^{14}C] acetic anhydride (9 nmoles) was added to aqueous reduced CoA (28 ug, 36 nmoles) at pH 7.2. After 20 min at room temperature, an excess of non-radioactive acetic anhydride (40 nmole) was added to ensure that the reaction had gone to completion. After a further 20 minutes at room temperature, 1ml of purified 5mM acetyl-CoA (section 2.2.5) was added to the reaction as a carrier, and the mixture was then extracted 3 times with ether as normal. This resulted in 4.5mM [^{14}C]-acetyl-CoA at a specific activity of 10^7 cpm / µmole.

2.2.8 Limited proteolysis

Fragmentation of PC was achieved by treatment with a number of proteases, under native conditions, as shown in Table 2.1 The incubation time varied with each experiment, to a maximum of 80 minutes, as described in the figure legends.

SPECIES	PROTEASE	PC:PROTEASE	TEMP	BUFFER
sheep & rat	chymotrypsin	100:1	30°C	NemCl pH 8.4
	proteinase K	50:1	30°C	NemCl pH 8.4
	endo Arg-C	50:1	30°C	NemCl pH 8.4
yeast	chymotrypsin	12.5:1	30°C	TrisCl pH 7.2
	proteinase K	12.5:1	30°C	TrisCl pH 7.2

Table 2.1 Conditions of partial proteolysis

Proteolysis reactions involving chymotrypsin and proteinase K were terminated by the addition of phenylmethylsulphonyl fluoride (PMSF) to a final concentration of 10 mM. Endoproteinase arg-C digestion was inhibited by the addition of EDTA to a final concentration of 10 mM.

2.2.9 Polyacrylamide gel electrophoresis

Polyacrylamide gel electrophoresis was carried out in the presence of SDS and 2-mercaptoethanol, as described by Laemmli (1970). Between 3 µg and 5µg of protein per track were routinely electrophoresed through acrylamide gels of dimensions; 16cm x 16cm x 0.05cm, at 30 mA for 1.5 hours. Gels were stained after electrophoresis in a 0.1% (w/w) solution of Coomassie Brilliant Blue R_{250} (in methanol (50% v/v) / acetic acid (10% (v/v)) for 1 hour. Destaining was effected in a solution of methanol (5% v/v)/acetic acid (10% v/v). Molecular weight markers (Sigma SDS-6H) consisting of: myosin (Mr 205,000), β galactosidase (Mr 116,000), phosphorylase b (Mr 97,400), bovine serum albumin (Mr 66,300), ovalbumin (Mr 45,000) and carbonic anhydrase (Mr 29,000) were co-electrophoresed to allow estimation of sample molecular weight.

2.2.10 Sequencing after electrophoresis

Fragments of sheep and rat PC, produced by partial proteolysis, were isolated for sequencing by the method of Aebersold *et al.* (1986). Subsamples of proteolysis reactions (40 µg) were electrophoresed through polyacrylamide gels (16 x 16 x 0.15 cm) and electrotransferred to 3-aminopropyltriethoxysilane-derivatised glass fibre (AP-glass) in 100mM NemTFA pH 8.4, as described by Svoboda *et al.* (1985). Protein was then visualised by staining in 3,3´ dipentyloxacarbocyanine iodide (1mg/ml) in 50mM $NaHCO_3$

pH8.4 /10% (v/v) methanol. Regions of the filter fluorescing upon u.v. irradiation were then excised and applied directly to the sample chamber of an Applied Biosystems 470 A Sequenator and subjected to Edman degradation, as described by Hunkapiller *et al.* (1983).

2.2.11 Western analysis

Proteolysed samples of PC were transferred to nitrocellulose after polyacrylamide gel electrophoresis as described by Svoboda *et al.* (1985). After electrotransfer, filters were washed in PBS-Tween (10 mM potassium phosphate, 150mM NaCl pH 7.2, 0.05% v/v Tween 20), and placed in a blotting apparatus (Miniblotter, Immunetics), allowing a single filter to be probed with up to 45 different antibody solutions without cross contamination. 120 µl of monoclonal antibody (0.1 mg/ml) or avidin-alkaline phosphatase conjugate (1.0 µg/ml) in PBS-Tween was pipetted into the appropriate slot of the blotting apparatus and incubated with shaking for 2.5 hours at 37°C.

After removing the probing solutions, each slot was washed thoroughly with PBS-Tween. The nitrocellulose filter could then be removed for further washing, after which it was incubated in a second antibody solution (anti-mouse IgG - alkaline phosphatase conjugate, 0.6 µg/ml) in PBS-Tween for 1 hour at 37°C. The filter was again thoroughly washed in PBS-Tween. A positive signal (binding of antibody or avidin) was detected by incubating the filter with substrates (0.3 mg/ml nitro blue tetrazolium and 0.2 mg/ml 5-bromo-4-chlor-3-indolyl phosphate) in 100 mM Tris-Cl pH 9.5, 100mM NaCl, 5mM MgCl for up to 15 minutes in subdued lighting. The reaction was quenched by washing in PBS-Tween containing 10 mM EDTA.

Sigma SDS-6H molecular weight markers were biotinylated by treatment with 1 mM N-hydroxysuccinimide-biotin for 10 minutes at pH 8.0. The reaction was quenched by the addition of L-lysine to 10 mM. This treatment did not significantly alter the electrophoretic mobility of the proteins.

2.2.12 TNBS modification

Yeast PC at 1mg/ml was chemically modified with the lysine specific reagent, trinitrobenzene sulphonate (TNBS) for up to 80 minutes at 30°C at a TNBS concentration of 0.5 mM. Although the reaction buffer (100mM Tris-Cl pH7.2) was chosen for historical reasons (Ashman *et al.*, 1973) and contains a primary amine, the nitrogen moiety of Tris is clearly not sufficiently nucleophilic to react significantly with TNBS. The modification of the protein was quenched by the addition of L-lysine to 10 mM. In some

experiments acetyl-CoA (0.5 mM) was included in the reaction mixture, as described in the figure legends.

2.2.13 Acetyl-CoA binding assay

The ability of PC to retard the progress of [^{14}C]acetyl-CoA through a Mr 30,000 cut off ultrafiltration membrane was measured as an indicator of acetyl-CoA binding. PC at concentrations of 0.5 mg/ml, 1.0 mg/ml and 2.0 mg/ml was combined with [^{14}C]acetyl-CoA (specific activity 10^7 c.p.m./ μmole) at a final concentration of 70 μM, in Amicon MPS-1™ ultrafiltration tubes and immediately centrifuged at 1500 x g for 10 minutes at 4°C. The solution was buffered by 100 mM Tris-Cl, at pH 7.2. Subsamples of the filtrate (50 μl) were taken into Triton-toluene scintillation fluid (2ml) and counted for 10 minutes in a Beckman scintillation counter. The number of counts in the filtrate was then compared to samples where no PC was included in the assay. In some samples, PC was pretreated with TNBS until 40% initial activity remained (as described in section 2.2.12), allowing a comparison to be made between TNBS-treated and untreated PC, in their ability to bind acetyl-CoA.

2.2.14 Biotin assay

The biotin content of a solution was estimated using the method of Rylatt *et al.* (1977). Samples were pre-treated with *Streptomyces griseus* protease for 1 hour at 30°C at a protease to substrate mass ratio of 1:100, so that steric hinderence by protein secondary structure did not influence the measurements. The protease was inactivated by incubation at 90°C for 5 minutes.

2.2.15 Fab isolation

Fab fragments of IgG were isolated from serum raised against trinitrophenyl-labelled bovine serum albumin (TNP-BSA) by papain fragmentation and protein A-Sepharose chromatography.

Serum (3ml) from a rabbit immunised with TNP-BSA, was loaded onto a protein A-Sepharose (5ml bed volume) column equilibrated in phosphate buffered saline (PBS). After washing the column with 5 volumes of PBS, the IgG fraction was eluted from the column with 100 mM glycine-HCl pH 3.0. The protein was precipitated with an equal volume of $(NH_4)_2SO_4$ at pH 7.0 and resuspended in PBS (24mg at 13mg/ml). Fab fragments were then produced by treating 13mg of IgG at 3mg/ml with papain (1% w/w) in 100mM Na acetate pH 5.8, 1mM DTE, 2mM EDTA at 37°C for 90 minutes. The proteolysis reaction was terminated by the addition of PMSF to 10 mM, after which the

mixture was chromatographed on protein A-Sepharose as before. The Fab molecules passed through the column unimpeded, and after precipitation with an equal volume of saturated $(NH_4)_2SO_4$, were resuspended to a concentration of 300 µg/ml in PBS-azide (0.1% w/v azide)

2.2.16 Electron Microscopy of Protease-Treated YPC

Samples of yeast PC at 1mg/ml were treated with chymotrypsin or proteinase K at a PC:protease ratio of 12.5:1, in 0.1 mM Tris-Cl, pH 7.2 for 80 minutes at 30°C. The reactions were terminated by the addition of PMSF to a final concentration of 10 mM. In order to preserve any resulting tetramers against effects of dilution during the mounting procedure, samples were treated with 1 mM dithio-bis-(succinimidyl propionate) (DTSP) in 50 mM potassium phosphate, pH 7.2 for 10 minutes at 30°C. This reaction was terminated by the addition of L-lysine to 10 mM. Sample mounting and microscopy, negatively staining with uranyl acetate, were performed as described by Mayer *et al.* (1980). Particle dimensions were measured to an accuracy of 0.1mm (equivalent to a length of 0.4nm in the sample), from prints of micrographs at a final magnification of 250,000x, using a calibrated magnifying glass. As it is known that deep stain and shallow stain conditions may have an influence on the measurable dimensions of PC (Mayer *et al.*, 1980) and that the staining conditions are usually not fully reproducible, the data used for figure 4.12, were taken equally from deep stain, shallow stain and intermediate depth of stain samples, and the obtained values were averaged.

2.2.17 Electron Microscopy of TNBS-Treated YPC

Samples of yeast PC at 1mg/ml (specific activity 21U/mg) were treated with 0.5 mM TNBS, in Tris-Cl, pH 7.2 at 30°C. After 80 minutes, the crosslinking reagent dithio-bis-(succinimidyl propionate) (DTSP) was added to a final concentration of 1 mM. After an additional 10 minutes at 30°C, both of these reactions were terminated by the addition of L-lysine to a final concentration of 10 mM, in order to preserve any resulting tetramers against effects of dilution during the mounting procedure. Residual cross linking reagent and TNP-lysine were removed from the samples by high performance gel filtration (TSK 3000 SW, Toya Soda) in 10 mM potassium phosphate pH 7.2, 100 mM NaCl.

The PC samples were concentrated five-fold under reduced pressure using a Speedvac concentrator (Savant) and complexed with anti-TNP Fab molecules (prepared as described in section 2.2.15) overnight at 4°C. Complexes and residual PC tetramers were separated from unbound Fab fragments by high performance gel filtration (TSK 3000 SW,

Toya Soda) in 10 mM potassium posphate pH 7.2, 100 mM NaCl. Fractions across the peak of high molecular weight material were mounted and negatively stained (uranyl acetate) for electron microscopy as described by Mayer *et al.* (1980).

Interpretations of the position of complex formation were made from prints of micrographs with a final calibrated magnification of 250,000x.

2.2.18 Reverse-phase High Performance Liquid Chromatography

Separation of complex peptide mixtures and preparation of peptides for N-terminal sequence analysis, was performed on silica-based reverse-phase columns at low flow rates using chromatographic equipment purchased from Waters™ (Millipore Corporation). The columns and elution conditions used in individual experiments are specified in the figure legends. Generally, butyl- or octyl-silica columns with internal dimensions of 2.1mm x 30mm or 4.6mm x 30mm were used with flow rates of between 0.2ml/minute and 1.0ml/minute. In most instances, chromatography was undertaken at room temperature, with initial aqueous buffer conditions, using either trifluoroacetic acid (TFA) or heptafluorobutyric acid (HFBA) as counter ions (0.1% v/v). The elution buffers consisted of organic solvent, either acetonitrile or propan-2-ol, containing the same counter ion as the initial buffer (0.1% v/v). Materials eluting from the column were monitored using a flow-through ultraviolet detector (Waters, model 490). This detector allowed simultaneous detection at up to four different wavelengths, and was routinely set to monitor at 215nm, 280nm, 345nm, and 420nm. Data and pump management were controlled by Waters Expert™ Chromatography software, on a Digital Electronic Corporation P350 personal computer.

2.2.19 Manipulation of DNA and Recombinant Organisms

Unless otherwise stated, DNA was cloned, isolated and analysed using standard procedures, as described by Maniatis *et al.* (1982). All manipulations involving recombinant DNA or recombinant organisms were performed in accordance with the regulations, and with the approval of the Recombinant DNA Monitoring Committee of the Australian Academy of Science and the Council of the University of Adelaide.

2.2.20 Preparation of DNA for Subcloning

Double stranded fragment and vector DNA were prepared by restriction endonuclease digestion and electrophoresis through 1% (w/v) agarose/TBE (90mM TrisCl pH 8.3, 40mM borate, 2.5 mM EDTA) gels. The DNA was electroeluted from gel slices

into TE (25mM TrisCl pH8.0, 10mM EDTA) in dialysis tubing and precipitated with ethanol after phenol extraction.

2.2.21 Ligation

Fragment DNA and vector, at a molar ratio of approximately 5:1, were incubated in a ligation buffer (50mM TrisCl pH7.5, 10mM $MgCl_2$, 1mM DTT, 1mM ATP) containing 1 unit of T4 DNA ligase, for 2 hours at 15°C.

2.2.22 Bacterial Transformation

E. coli XL1-B were cultured in 2YT + tetracycline (10 µg/ml) to an OD_{600} of 0.6. The cells were pelleted and resuspended in 50 mM $CaCl_2$, 10mM $MgCl_2$. After 40 min at 4°C, the cells were competent to take up DNA. Ligated DNA (~5-10 µg) was mixed with competent cells (300 µl) and after 40 min at 4°C the mixture was overlayed onto L-agar (+/- antibiotics as appropriate) in 3ml of molten 2YT + 0.7% (w/v) agarose. When pUC and M13 were used as vectors, 10 µl of 100mM IPTG and 30 µl of 0.2% (w/v) BCIG were included in the mixture just prior to overlaying. The plates were then incubated overnight at 37°C.

2.2.23 Rapid Preparation of Double Stranded Plasmid DNA

Double stranded plasmid DNA was prepared from overnight broth cultures by the alkaline lysis method based on the procedure of Birnboim and Doly (1979). Pelleted cells were resuspended in TEG buffer (25mM TrisCl pH8, 10mM EDTA, 50mM glucose) and treated with 2 volumes of 200mM NaOH, 1% (w/v) SDS. After 5min at 4°C a half volume of 5 M potassium acetate was added, followed by centrifugation at 12,000 x g for 10 min. The supernatant was extracted with phenol and chloroform, and the DNA was precipitated with 2 volumes of ethanol.

2.2.24 Rapid Preparation of Single Stranded Phage or Phag-mid DNA

Single stranded DNA was prepared from M13 clones and deletion clones for sequence analysis as described by Maniatis *et al.* (1982). Recombinant M13 phage particles were produced from infected XL1-B cells after incubation in 2YT + tetracycline (10µg/ml) for 6 hours at 37°C. Phage particles carrying single stranded copies of plasmid DNA were produced by incubation of plasmid bearing XL1-B cells with M13KO7 helper phage for 2 hours at 37°C, followed by incubation overnight at 37°C in the presence of kanamycin (70µg/ml).

Phage particles were precipitated from culture supernatants by the addition of 700mM NaCl / 5.5% (w/v) PEG 6000, followed by centrifugation at 12,000 x g for 10

min. The pellet was resuspended in TE and extracted with phenol. The DNA was precipitated with ethanol and 300mM Na-acetate, pH 5.0. In some cases, remaining protein contaminants were removed by precipitation with ammonium acetate (2.5 M final concentration). The DNA was then precipitated with 3 volumes of ethanol.

2.2.25 DNA Sequencing

Single stranded DNA was sequenced by the dideoxy chain termination method of Sanger *et al.* (1977). Second strand synthesis was carried out using a commercially optimised reagent kit (Bresatec Ltd.), and involved the incorporation of either α-[^{35}S]-dATP or α-[^{32}P]-dATP.

The reactions were then electrophoresed through 6% polyacrylamide gels (acrylamide:bis, 19:1) containing 8M urea in TBE buffer. The resultant gels were either fixed with 20% (v/v) ethanol /10% (v/v) acetic acid and baked onto a glass plate at 100°C for 1 hour or dried onto Whatman 3MM paper. The dried gel was then exposed to X-ray film (Konica) at room temperature until bands were adequately visible (usually 16 hours for ^{32}P, 48 hours for ^{35}S).

2.2.26 Preparation of Overlapping Deletion Clones

A series of nested deletions of the 4kb EcoRI-BamHI fragment of λRG1.2 were generated using the Sequenest II™ (Gold BioTechnology) transposon deletion system of Ahmed (1985). The strategy of this method, described below, is illustrated in figure 2.1.

This involved subcloning the fragment into vectors pAA-PZ718 and pAA-PZ719 in *E. coli* XL1-B. The construct was then transformed into the permissive host strain *E. coli* AA102F'lac/low, stabilising the transposable element with chloramphenicol resistance. Relieving this selection and growing in the presence of galactose, the transposable element of Tn 9 is able to excise, occasionally removing adjacent vector DNA in the process. The size of these deletions can vary, but always occur from the left hand (IS1-L) element of the transposon. This enables a series of deletions ranging in size to be produced in a uni-directional manner by simply allowing the transposable element to excise. As this is a rare event, even in a permissive strain of *E. coli* , a powerful selection pressure is applied to enhance the rate of formation of deletions. Between the transposable element and the cloning site, the vector contains a region of the *E. coli* lac operon. When grown in the presence of galactose, the proteins of this operon are induced, but due to the lack of a functional epimerase enzyme in this strain, the resulting build up of UDP-galactose in the cell kills the host. Thus, Deletions through the lac operon are selected for, as removal of

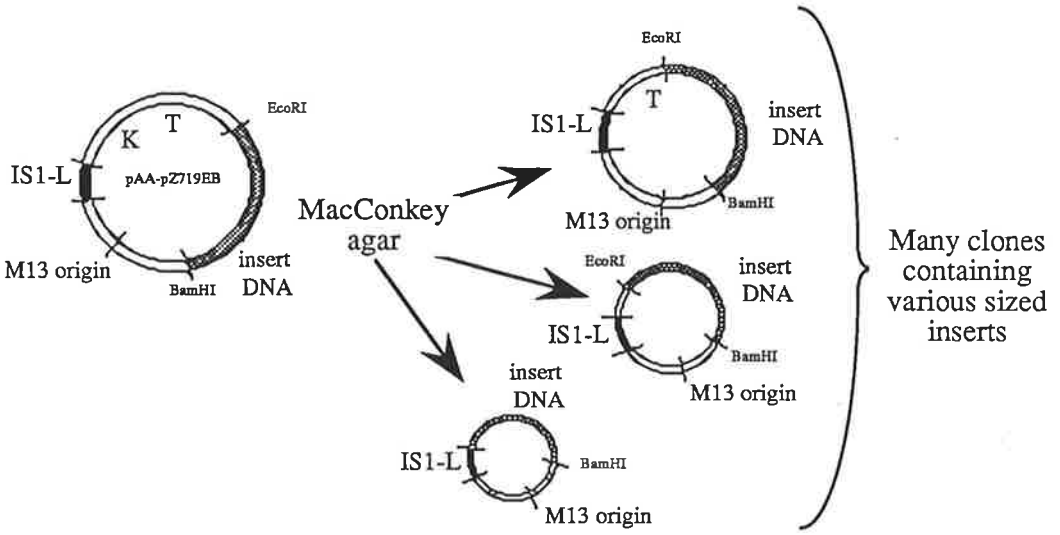


Figure 2.1 Sequenest II A schematic representation of the transposon deletion system of Ahmed(1985) (GOLD BioTechnology Inc.). When galactose is used as the carbon source in the growth media, transposition events lead to deletions random length from one end of the IS1-L element (in a clockwise manner with respect to this diagram) resulting in a series of clones of varying length. Second strand synthesis, priming from an oligonucleotide directed to an invarient sequence within the IS1-L element allows sequencing of the deletions in this same direction (clockwise). K and T represent the gal K and gal T genes of the *E. coli* lac operon respectively.

gal K or gal T activity are required for cell survival. To ensure that deletions progress into the DNA fragment of interest, a second selection is made by making a mixed DNA

preparation from many transposon excision events, and treating this DNA with a unique restriction endonuclease (preferably the enzyme corresponding to the cloning site at the end of the clone nearer the transposable element). Constructs where deletion has proceeded into the insert will no longer possess this restriction site and will remain circular. When this mixed DNA is transformed into *E. coli* XL1-B, the linearised DNA will not transform efficiently, thereby selecting for colonies which contain deletions into the insert . These clones can then be isolated, DNA prepared and sized by gel electrophoresis. As the vector contains the M13 origin of replication, single stranded DNA can be produced from clones of interest by superinfection with a helper phage M13K07. Sequencing is then undertaken, priming the second strand synthesis with an oligonucleotide complementary to the invariant sequence at the 3´end of the Is1-L element (oligonucleotide TSP - section 2.1.8).

Constructs pAA-PZ718-EB and pAA-PZ719-EB were transformed into *E. coli* AA102F´lac/low on L + ampicillin (100µg/ml) + tetracycline (12.5µg/ml) + chloramphenicol (12.5 µg/ml) plates. Recombinants were then grown overnight on MacConkey agar + ampicillin (100µg/ml) + tetracycline (12.5µg/ml) + galactose (10

μg/ml). Double stranded DNA was prepared from mixed cultures of a large number of colonies grown in 2YT + ampicillin + tetracycline. Deletions into the insert DNA were ensured by treating the mixed-plasmid DNA preparation with either EcoRI (pAA-PZ719-EB) or BamHI (pAA-PZ718-EB) as appropriate. The restricted DNA was then used to transform XL1-B, as described in section 2.2.22.

Double stranded preparations from a large number of the resulting colonies were performed in 96 well mictrotitre plates, as described in Ausubel *et al.* (1989). Clones suitable for sequencing were selected by sizing the linearised plasmids on 0.8% (w/v) agarose gels. Single stranded DNA was produced by infection of XL1-B clones with M13K07, as described in section 2.2.24. A sequencing primer (TSP), complementary to the -23 region of IS1-L was used to sequence putative deletion clones, as described in section 2.2.25).

2.2.27 Computer Programs

DNA sequence was read from gel autoradiographs using a digitiser (Science Acessories Corporation) with a program written in PASCAL by Mr. M. Snoswell (Adelaide). DNA sequence overlaps were found using SPCOMP (Dr. A. Sivaprasad, Adelaide), a modified version of DBCOMP (Staden, 1982). ANALYSEQ, a suite of DNA analysis programmes (Staden, 1982) was used to identify restriction enzyme sites, translate coding regions into amino acid sequence and produce sequence data of the complementary strand.

Localised protein sequence similarities were identified using SEQHP (Dayhoff *et al.*, 1978, Goad and Kanehisa, 1982). Sequence databases GenBank, EMBL, NBRF and VECBase were screened using the suite of programs from Genetics Computing Group of the University of Wisconsin (Devereux *et al.*, 1984) particularly WORDSEARCH, FASTA, TFASA. The codon usage of the 3´ end of rat PC gene shown in table 3.2 was tabulated by the program CODONUSAGE. The RNA secondary structure prediction shown in figure 3.3 was calculated using the program FOLD (using the method of Zucker and Stieger, 1981) and displayed with the program SQUIGGLES.

The protein secondary structure prediction shown in figure 4.13 was calculated using the program PEPTIDESTRUCTURE (using the method of Garnier *et al.*, 1978) and displayed with the program PLOTSTRUCTURE.

Chromatographs were produced from printed output using a digitiser (Houston Instruments) and AutoCAD (Autodesk Inc.). Word processing and typesetting were

performed on a Macintosh SE personal computer using Word version 4.00 A (Microsoft Corporation). Diagrams and charts were produced using MacDraw II (Claris), Excel (Microsoft Corporation) and Cricket Graph (Cricket).

CHAPTER 3

ANALYSIS OF THE PRIMARY STRUCTURE

3.1 INTRODUCTION

Although the folding pathway of large complex proteins may be a consequence of many external forces acting in concert, clearly the critical starting point in all cases is the amino acid sequence of the polypeptide (Anfinsen *et al.*, 1961). This has been well demonstrated in such examples as ribonuclease A (White, 1961), growth hormone (Holzman *et al.*, 1986) and bovine pancreatic trypsin inhibitor (Creighton, 1977), which can be re-folded into native structures after denaturation, indicating that the primary amino acid sequence alone is sufficient to dictate the tertiary structures of these proteins.

White (1960) was able to show that activity could be regenerated from fully reduced and denatured ribonuclease A upon reoxidation . Subsequent structural comparisons (White, 1961; Bello *et al.*, 1961) between reduced-reoxidised ribonuclease A (reox-RNase A) and the native enzyme, showed that the structures of the two proteins were essentially identical, indicating that the amino acid sequence alone is sufficient to dictate the tertiary structure. Further confirmation of this finding was given by Taniuchi (1970) who showed that truncation of the C-terminal 4 amino acids from ribonuclease A prior to reduction, inhibited the ability of the protein to regenerate the correctly matched disulphide bonds and hence the correct overall structure, upon reoxidation.

Although the structure of a protein is essentially dictated by its amino acid sequence, some flexibility can clearly be tolerated. Bovine pancreatic trypsin inhibitor (BPTI) is a 58 amino acid protein that contains 6 cysteine residues, which are paired into 3 disulphide bonds (cys^{14}/cys^{38}, cys^{30}/cys^{51} & cys^{5}/cys^{55}), and readily refolds after reduction and reoxidation to form biologically-active protein (Creighton, 1977). When Marks *et al.* (1987) replaced cysteines #14 and #38 of BPTI with alanines or threonines, the mutant proteins were found to fold into native conformations, albeit at a slower rate, suggesting that the presence of these 2 cysteines (and hence disulphide bond) was not critical for the formation of the tertiary structure.

The determination of the amino acid sequence of a protein therefore, is a necessarily desirable prerequisite in the analysis of its overall structure and function. This sequence data can then be utilised in comparisons with sequences from other protein groups or the same protein from other species in an attempt to provide information about regions or residues important for structure or function. In addition, knowledge of the amino acid sequence of a protein aids in the rational development of further experimentation. Present

N-terminal sequencing technology is too inefficient to allow rapid sequence analysis of proteins larger than 70 to 80 amino acids, thus for large proteins the sequence is more effectively determined from analysis of the gene(s) or cDNA(s) which encode the protein product.

Analysis of the gene in itself may provide useful information regarding the structure or ancestry of the protein, especially when compared with other biochemical and structural evidence. Do, for example, the exons found within some genes represent discrete functional or structural domains? Traut (1988) surveyed the sequence data base, examining the correlation between exons and domains. While numerous examples could be detailed for both sides of the argument, the conclusion drawn was that such a correlation could be observed too often to be due to chance alone.

Based on the evolutionary relationship between biotin carboxylases proposed by Lynen (1975), it may not be unreasonable to expect functional regions or domains of PC to be encoded by one or more exons. The proposed assembly of genes encoding a single multifunctional polypeptide may have been facilitated therefore, by duplication or rearrangement of discrete units of structure in the form of exons.

An additional benefit of characterisation of the gene or cDNA encoding a vertebrate PC would be the identification of the mitochondrial targetting and import sequences. As PC from vertebrate sources is a mitochondrial matrix protein (Utter and Keech, 1960) that is synthesised in the cytosol (Srivastava *et al.*, 1983), these signalling sequences are cleaved from the precusor during translocation into the mitochondria (for recent reviews on importation into mitochondria see Hartl *et al.*, 1989 & Pfanner *et al.*, 1988). Since isolation from vertebrate cells of sufficient quantities of the precursor to enable characterisation by N-terminal sequence analysis would not be feasible in the absence of an appropriate cDNA clone, sequencing of the cDNA is clearly the most efficient method for determining the signal sequence.

The long term regulation of the levels of PC in vertebrate tissues is at the level of transcription of the gene (Angus *et al.*, 1981). Therefore study of the mechanism of this control, which is influenced by the nutritional and endocrine status of the animal, requires isolation and characterisation of the 5′ region of the PC gene containing the transcriptional regulatory sequences.

It is clear therefore that characterisation of the rat PC gene and cDNA would be a particularly useful adjunct to the determination of the primary amino acid sequence of the protein.

Dr. A. I. Cassady, in this laboratory, identified 4 short overlapping cDNA clones from a λgt_{10} rat liver cDNA library (supplied by Dr. G.J. Howlett, Melbourne), using a fragment of the human PC cDNA clone (supplied by Dr. R. Gravel, Toronto) as a probe. The largest of these clones, λRL1.1, was sequenced, providing 939bp of sequence including the entire 3´non-coding region, and its identity confirmed by the presence of the characteristic biotin-attachment site in the inferred protein sequence. This clone was also shown by Northern analysis to hybridise to a rat RNA species of ~ 4.2 kb, which is consistent with the expected size of the PC mRNA (Freytag and Collier, 1984). This rat clone (λRL1.1) was then used to re-screen the same rat cDNA library in order to isolate longer clones. Of 68 initial positively hybridising clones, 3 were shown to contain sequences further 5´to the existing clone λRL1.1. The 5´ EcoRI-PstI fragment of the longest of these clones, λRL2.35, was sequenced and shown to contain 400 bp of sequence further 5´to the extent of λRL1.1. Together, these two cDNA clones represent the 3´ 1338 bp of the rat PC mRNA, which encodes the C-terminal 306 amino acids or ~ 26% of the length of the protein (Cassady, 1987; for reference see figures 3.1 & 3.2).

In addition to his study of the rat PC cDNAs, Dr Cassady also isolated 2 overlapping clones (λRG1.2 & λRG1.4) from a λCharon 4A rat genomic library (Phytogen Corp., CA, USA). Initial characterisation of these clones (Cassady, 1987) showed that the rat cDNA clone λRL1.1 hybridised to a 1.4 kb BamHI-HindIII fragment which was common to both clones. This fragment was subcloned and sequenced, identifying two exonic regions at the 3´ end of the rat PC gene (Cassady, 1987; for reference see figures 3.6 & 3.7).

At the conclusion of Dr. Cassady's studies it was clear that another of the rat cDNA clones that he had isolated (λRL2.53) contained sequences further 5´ to the two clones (λRL1.1 & λRL2.35) discussed above. In addition, a preliminary experiment indicated that the ~4 kb of the rat genomic clone λRG1.2 may contain further exonic regions.

Described in this chapter are

(i) sequencing of the EcoRI-KpnI fragment of cDNA clone λRL2.53

(ii) sequencing of the ~ 4 kb EcoRI-BamHI fragment of λRG1.2

(iii) comparison of the inferred amino acid sequence with other proteins.

3.2 RESULTS AND DISCUSSION

3.2.1 Characterisation of the cDNA

Although the mRNA for pyruvate carboxylase is relatively abundant in liver tissue (estimated at 0.042%; Cassady, 1987), cloned cDNA representing the entire coding region of PC has been elusive. Freytag and Collier (1984) isolated a short cDNA clone representing the 3´end of the coding region from a human liver cDNA library. Cassady (1987) isolated and sequenced 2 rat liver cDNA clones (λRL1.1 & λRL2.35) which together provided some 26% of the coding region.

An overlapping rat cDNA clone, λRL2.53 was subsequently purified and partially sequenced by Dr. A.I. Cassady in this laboratory. The remaining 500 bp of sequence were determined as part of this study, after subcloning restriction fragments of this clone into M13 vectors (mp18 & mp19). The sequencing strategy for this region, along with the restriction maps of clones λRL2.35 and λRL2.53 are shown in figure 3.1. Taken together with the nucleotide sequence of λRL2.35, this data represents the 3´ 2067 bp (~49%) of the rat PC mRNA, which has been estimated by northern analysis to be ~4.2 kb in length (Cassady, 1987).

As shown in figure 3.1, although λRL2.53 extends the cDNA sequence in the 5´direction with respect to λRL2.35, the two clones are in fact of a similar size. The 3´ truncation of λRL2.53 explains how this clone was missed in the initial screening for cDNA clones which were further 5´ to λRL1.1(Cassady, 1987), in that the oligonucleotide probe used for this screening (directed against the biotin-attachment site) would not be able to hybridise to this clone. As the library from which this clone was isolated was constructed from cDNAs synthesised using poly-deoxythymidine oligonucleotide (oligo-dT) priming, clones are expected to reflect the mRNA sequence from the 3´end. However in the case of λRL2.53 a region of up to 550 bp has been truncated from the 3´end of this clone, during the cloning process.

These two cDNA clones (λRL2.35 & λRL2.53) are the longest isolates, from a number of cDNA library screenings (Cassady, 1987 & unpublished results), suggesting that isolation of a cDNA which represents the full length of the rat PC mRNA is unlikely

from the available cDNA libraries. This conclusion is supported by the failure of attempts in this laboratory to synthesise rat PC specific cDNAs using an oligonucleotide primer (RPC02) directed against the known cDNA sequence at the 5′end of λRL2.35 (Cassady, *pers. comm.*). On this basis, it is possible that under the conditions used for cDNA synthesis, the rat PC mRNA is present in a highly structured form, thereby restricting the synthesis of long cDNA molecules.

An analysis of this region of the known rat PC mRNA sequence (residues 1-1200) using the secondary structure prediction program FOLD (Zuker, 1986) is shown in figure 3.3. The free-energy (ΔG) value calculated in this example (-1840 kJ/mole or -440 kcal/mole) is comparable to values predicted for the proposed intron-splicing structure of*Tetrahymena* ribosomal RNA (Cech *et al.*, 1983), suggesting that formation of such a structure is at least plausible. Interestingly, when adjacent sequences (5′- predicted from the gene sequence, section 3.2.2; 3′- residues 800 - 2000) are subjected to the prediction program, little or no significant secondary structures result, which is consistent with the observed difficulties in cDNA synthesis noted above. However, any interpretation of these predictions must be mindful that only a portion of the total mRNA molecule can been analysed at any one time (caused by the size limitation of the program - 1200 nucleotides) and the program does not take into account any influences tertiary conformations may have on the stability of the mRNA.

The search for full length cDNA clones has not been aided, however, by the use of nucleotide probes derived from primarily the 3′ end of the mRNA. Until recently, little protein sequence from the N-terminal half of rat PC, or indeed PC from any species, has been available for generation of nucleotide probes. The sequence of the N-termini of two chymotryptic fragments of rat PC are reported in chapter 4 of this thesis, and a report of the sequence of the N-terminus of the native rat PC has recently been published (Thampy *et al.*, 1988). These sequences however, contain amino acid residues such as leucine and serine which may be represented by any of 6 codons in the nucleotide sequence, thereby limiting their value as regions for oligonucleotide probes.

The report of the N-terminal sequence for mature rat PC is particularly interesting, in that the N-termini of PC from other species are considered to be N-terminally blocked (Rylatt *et al.*, 1977; Lim, 1988). In addition this sequence bears little or no similarity to the protein sequence for yeast PC inferred from the sequence of the corresponding gene. The overlapping protein sequences for 3 biotin-containing tryptic fragments were also

reported by Thampy *et al.*, 1988, and are shown in figure 3.4, along with the corresponding amino acid sequence inferred from sequencing of the rat PC cDNA and gene (Cassady, 1987), yeast PC gene (Lim *et al.*, 1988) and human PC cDNA (Freytag and Collier, 1984).

As can be seen from this figure (3.4), a number of residues differ between the rat PC sequence reported by Thampy *et al.* (1988) and that of Cassady (1987). The peptide data reports an alanine residue at position 2 and a threonine at residue position 24, whereas both the gene and cDNA clones encode glutamine and methionine residues at these positions respectively. In addition, a cysteine residue predicted from the gene and mRNA at position 5 has not been detected by protein sequencing. As the peptides subjected to sequence analysis were not treated with an alkylating reagent such as iodoacetic acid (which protects cysteine residues during sequencing reactions), it is perhaps not surprising that the cysteine residue predicted from the DNA sequence was not detected after Edman degradation, although a "blank cycle" should have been observed in this case.

Although these differences may reflect some form of post-transcriptional editing mechanism this is very unlikely. Exchange of an individual amino acid from within a polypeptide with another amino acid after synthesis would be an unprecedented finding, and require an extensive enzymic frame work in order to function reliably. The observed differences are unlikely to be a result of artifacts introduced during the cloning process in that the gene and cDNA clone have be constructed and isolated quite independently. The possibility that these differences are due to intra-species variation (Sprague-Dawley versus Wistar rats) seems most unlikely, in view of the identity of the rat and human PC in this region.

As shown in figure 3.4, in the case of the alanine residue at position 2, the assignment was based on a single Edman degradation experiment, where as the assignment of threonine (position 24) was apparently based on three separate determinations. The fact that the DNA sequences for the rat PC gene and cDNA concur and have been determined by sequencing of both strands of the DNA, strongly suggest that they represent the more "correct" sequence for the amino acids surrounding the biotin attachment site. This conclusion is supported by the additional comparison with inferred sequences for human PC and yeast PC shown in figure 3.4. The latter sequence has also been confirmed by Edman degradation of the corresponding biotin peptide (Lim *et al.*, 1988).

The additional sequence reported by these authors (Thampy *et al.*, 1988), as representing the N-terminus of rat PC should therefore be treated with some degree of caution, particularly for use in the design of oligonucleotide probes to the 5´end of the rat PC gene. Although some difficulty has been encountered in the isolation of long cDNAs for PC from any species, the rapid and recent development of polymerase chain reaction (PCR) technology provides a future avenue of investigation. Jansen *et al.* (1989) isolated full length cDNA clones for methylmalonyl-CoA mutase [MCM; EC 5.4.99.2] in a clear demonstration of the power of the PCR technique. These authors synthesised oligonucleotide primers from the sequence of the 5´ end of an existing partial cDNA clone of MCM and also to a region in the cloning vector λgt_{11}. Using PCR, they were then able to "screen" a very large number of recombinants from a λgt_{11} cDNA library by selectively amplifying DNA which consisted of MCM sequences further 5´ to the existing clone. These DNA fragments could then be sized, subcloned and sequenced. Such methodologies should be readily applicable to the construction of a complete cDNA sequence for rat PC, thereby complementing the present study into the structure of the rat PC gene.

3.2.2 Characterisation of the Gene

Given some of the difficulties, detailed in the previous section, of isolating and characterising cDNA clones representing the entire PC mRNA, determination of the rat PC protein sequence was undertaken, directly from the gene sequence.

A 12 kb clone of rat genomic DNA (λRG1.2) was isolated and shown to contain the 3´ end of the rat PC gene (Cassady, 1987). The identity of this clone was confirmed by hybridisation to rat cDNA clone λRL2.35 and subsequent sequencing of the 1.4 kb BamHI-HindIII fragment, which included the sequence of the biotin attachment site (Cassady, 1987). This sequence data also revealed two exonic regions separated by a small intron, of only 78 nucleotides. Southern analysis of λRG1.2, probing with the rat cDNA clone λRL1.1, provided evidence to suggest that the region of λRG1.2 further 5´ to that already sequenced contained additional exons (Cassady, 1987). The level of hybridisation observed in this experiment was quite low, however, suggesting that prior to sequencing of this region, more convincing evidence for further exons was required.

The ~ 4 kb EcoRI-BamHI fragment of λRG1.2 was subcloned into pUC19 and the resultant construct was designated pRGEB.1. In order to confirm that this region of the rat PC gene contained one or more exons, restriction fragments of this clone were subjected

to Southern analysis, probing with an oligonucleotide derived from a region of the cDNA towards the 5′end of λRL2.35. This oligonucleotide, RPC02 (the sequence of which is shown in section 2.1.8), hybridised to a number of restriction fragments including a PstI-KpnI fragment of ~400 bp, as shown in figure 3.5. The position of this fragment within the restriction map of pRGEB.1 is shown in figure 3.6, along with the parent clone λRG1.2, clearly indicating that pRGEB.1 encoded at least one exon of the rat PC gene. In addition, a comparison of the restriction pattern of pRGEB.1, with that of the cDNA clones (compare figure 3.1 with figure 3.6), suggested that a significant proportion of the DNA 3′ of the KpnI site in pRGEB.1 may be represented in the cDNA.

Sequence analysis of the 3′ ~1.2 kb (KpnI-BamHI) fragment of pRGEB.1 was undertaken by subcloning various KpnI-PstI and PstI-BamHI fragments into M13 Vectors (mp18 & mp19), and sequencing the resultant single stranded DNA by the chain termination method of Sanger *et al.* (1977). The sequencing strategy for this region is shown in figure 3.7. Three complete exons were identified within this region of DNA, based on the corresponding cDNA sequences. Three complete introns were also identified, the largest of which consisted of 205 nucleotides. The small size of the introns so far determined, together with evidence for an exon at the 5′end of this sequence, suggested that completely sequencing pRGEB.1 would provide a significant proportion of the coding sequence of the gene.

The ~4 kb EcoRI-BamHI fragment from pRGEB.1 was cloned into Sequenest II™ vectors pAA-PZ718 and pAA-PZ719, as described in section 2.2.21. A series of deletions into the insert DNA were made from these constructs, as described in section 2.2.26, allowing efficient sequencing of ~ 3 kb of DNA. This method of nested deletion generation has advantages over other methods in that the deletions occur *in vivo*, and are therefore not dependent on *in vitro* exonuclease activities or reaction time-courses. In addition, these vectors are able to readily maintain larger DNA inserts than phage vectors such as M13, allowing significantly larger sequencing projects to be handled with a minimum of subcloning.

Figure 3.7 shows the sequencing strategy for pRGEB.1. Regions of coding sequence where only one strand of the DNA have been sequenced were confirmed by comparison to the corresponding cDNA sequences. Also shown in figure 3.4 are the positions of the 10 identified exons, with respect to a limited restriction map of the gene and

the region previously sequenced (Cassady, 1987). The limit of the cDNA sequence is indicated in the figure.

The sequence of pRGEB.1 completes the sequence of the 3′ 5436 bp of the rat PC gene, which is shown in figure 3.8 as an EcoRI-HindIII fragment of the rat genomic clone λRG1.2. Ten exons have been identified within this sequence, encoding 750 amino acids, interspersed with eleven introns. Of the 10 exons, seven can be readily identified by direct comparison to the known cDNA, but the remaining three however, represent sequences further 5′ to the present limit of cDNA sequence. The positions of these three exons have been proposed on the basis of conformity to the splice acceptor and donor consensus sites of Cech (1983) shown in figure 3.9, and strong conservation of the resulting inferred sequence of rat PC with that of yeast PC and human PC (figure 3.10). As described in section 4.2.3, the amino acid sequence of an internal fragment of rat PC has been determined and found to match the inferred sequence derived from exon X, thereby confirming the coding nature of this region of the gene.

The statistical data shown in table 3.1 provides an indication of the high proportion of coding sequences in the 3′ end of the gene in that the average exon length is ~260 bp, including a notably small exon of 43 bp (exon IX) and the average intron length is a remarkably small ~220 bp. This provides a coding to non-coding ratio of 54% for the 3′end of the gene, which if consistent throughout would result in a gene extending over approximately 8 kb.

Characterisation of the rat PC gene not only identifies the amino acid sequence of the protein, it also provides information not accessible from any other source, such as the intron-exon boundaries and 5′ regulatory regions. In addition the codon usage within the coding region of the gene can be ascertained (Table 3.2), providing useful information regarding the preference for particular codons for the various amino acids. The identification of the 5′ regulatory regions, which requires isolation of genomic DNA further 5′ to pRGEB.1, would be particularly interesting given that the site of control of PC in rats is principally at the level of transcription (Angus *et al.*, 1981).

As described in section 1.2, Lynen (1975) proposed that the biotin carboxylases represent a group of enzymes which have evolved into complex multifunctional proteins from smaller monofunctional precursors through successive gene fusions. While evidence for this hypothesis is steadily accumulating with the elucidation of the protein sequences for the various members of the biotin carboxylases, the hypothesis is one of gene fusion and

therefore requires evidence that at least gene fusion may have occurred. Such evidence is most likely to result from studies of the genes for the various biotin carboxylases, which to date have mostly been limited to non–vertebrate species, and thus do not contain introns. As the data presented here is the first example of an extensive study into the gene structure of a vertebrate biotin carboxylase, few conclusions can be drawn regarding the role of gene fusion in assembly of the various biotin carboxylases.

Two recent reports in the literature, however, indicate that studies into the gene structures of rat acetyl-CoA carboxylase (ACC; Luo *et al.*, 1989) and human propionyl-CoA carboxylase (PCC; Tahara *et al.*, 1990) have been initiated. Luo and co-workers (1989) have investigated the 5′untranslated region of the ACC gene in an attempt to determine the mechanism by which the observed 5′ end heterogeneity of ACC mRNA is produced, whereas Tahara *et al.*, (1990) have studied the incidence of particular mutations found in the β subunit of PCC from patients with propionic acidemia. As both of these studies involve regions of their respective genes which are not represented in PC, no meaningful comparisons can be drawn between the structures of the genes encoding these proteins.

In the absence of suitable comparative data between intron-containing biotin carboxylase genes, it is however interesting to compare the positioning of the introns in the rat PC gene with the separate genes for bacterial enzymes such as transcarboxylase. This comparison will be dealt with in the following section.

3.3.3 Characterisation of the Inferred Protein Sequence

As outlined in the introduction to this chapter, the single most important influence on the structure of a protein molecule is its amino acid sequence. Partly as a result of this study, the sequence of the C-terminal 750 amino acids of rat PC has now been determined, inferred from the DNA sequence of the corresponding cDNA and gene. This sequence is shown in figure 3.10, along with the region of protein sequence determined by Edman degradation of a proteolytic fragment of rat PC (determined in section 4.2.3 of this thesis). While methods to reliably predict the tertiary structure of a protein solely from its amino acid sequence are not yet available, these results represent the first step in the long process of determining the overall structure of rat PC.

Comparing the partial amino acid sequence of rat PC with the corresponding sequences of yeast PC (complete - Lim *et al.*, 1988) and human PC (incomplete -

Lahmonwah *et al.*, 1987), significant similarities are observed (figure 3.10). When aligned over their respective lengths, PCs from rat and yeast are 46% identical while PCs from rat and human are 96% identical. The level of conservation between these sequences is consistent with the commonly held view that the rat and human species diverged more recently in evolutionary terms than the fungi and animals. From figure 3.10 it can also be seen that the conservation of sequence between rat PC and yeast PC occurs in clusters, separated by less well conserved regions. As the two proteins are known to catalyse the same reaction and have the same overall structure (as discussed in section 1.4.2) these regions of similarity highlight the portions of the protein most likely to be functionally important. These regions are further highlighted when sequence comparisons are made with proteins from different enzyme groups.

As summarised in figure 1.4, Lim *et al.* (1988), compared the amino acid sequence of yeast PC to published sequences from other members of the biotin carboxylase family and noted areas of strong conservation. Sequence similarities were also observed between yeast PC and proteins from other enzyme groups. When these conserved regions were compared to the known biochemical data for the various polypeptides, an interesting correlation between sequence conservation and protein function was observed. In particular, regions have been ascribed to the binding of ATP (1^{st} partial reaction subsite), pyruvate/oxaloacetate (2^{nd} partial reaction subsite) and the attachment of biotin.

Inclusion of the protein sequence of rat PC in these comparisons confirms these original findings, highlighting the presence of the biotin-carrier region (Cassady, 1987) and the pyruvate binding region. Figure 3.11 shows the comparison between the sequences of rat and yeast PC, transcarboxylase 5S subunit and the α–subunit of oxaloacetate decarboxylase in the region ascribed to pyruvate binding by Lim *et al.* (1988) and Samols *et al.* (1988), highlighting the instances when at least three of the four residues are identical. These comparisons become particularly interesting when compared with the results of biochemical studies.

As detailed in section 1.4.4, the structure of the pyruvate binding sites of PC and the 5S subunit of TC have been studied using chemical labelling, fluorescence quenching and NMR. In the case of PC, the site is thought to contain a tightly bound divalent metal ion (Mn(II) in chicken, sheep & rat ; Zn(II) in yeast), at least one catalytically competent cysteine residue and a base with a pKa ~4.5. TC 5S subunit contains tightly bound Zn(II) and Co(II) metal ions, and at least one critical tryptophan residue.

Kumar *et al.* (1988b) demonstrated that modification of a tryptophan residue (trp^{73}) of TC 5S with 2,4-dinitrophenylsulphonyl chloride (DNPS-Cl) was prevented by coincubation with pyruvate, implicating this residue with the pyruvate binding site of TC. Interestingly, this residue is absolutely conserved between the four sequences shown in figure 3.11 and is surrounded by residues which are also highly conserved. Another tryptophan, 17 residues further N-terminal, is also absolutely conserved in this comparison and either of these may be the tryptophan observed to interact with the divalent metal ion in TC (Kumar *et al.*, 1988a).

Of particular interest is the incidence of an asparagine residue at position 178 in the rat sequence. Previously the conservation of a cysteine at this position in the remaining proteins had led to speculation that this residue may represent one of two likely basic residues involved in the reaction mechanism, as explained in section 1.4.4. Although it is possible that the mammalian PCs have evolved a subtly different reaction mechanism for the decarboxylation of biotin, this seems somewhat unlikely, given the overall high conservation of sequence and structure between the PCs.

Interestingly, a number of basic side chain amino acids within this region are absolutely conserved between the four enzymes, which suggests that although there is evidence to show that cysteine residues are within the vicinity of the pyruvate binding site (Hudson *et al.*, 1975; Bagley, 1982; Bagley *et al.*, 1983), the catalytic base in the decarboxylation reaction may indeed be a more traditional basic residue such as arginine or histidine. Some evidence for this argument comes from inactivation studies using the arginine modifying reagent diacetyl ($CH_3COCOCH_3$; Penberthy, 1970). At pH 8.2, diacetyl was found to rapidly inactivate sheep PC following pseudo-first order kinetics, causing modification of a single residue. Analysis of these results is complicated however, in that at this pH, some reaction with lysine residues is likely, while at pH 6.5, where the reagent is more selective, the kinetics of the reaction are no longer pseudo-first order.

The identity of the catalytic base in the second partial reaction site may well be best answered by site directed mutagenesis replacing, in the first instance, the cysteine residue (cys^{600}) in PC from yeast and observing the effect on the decarboxylation reaction.

Another facit of the proteins compared in figure 3.11 is the coordination of a tightly bound divalent metal ion. The detailed structures of a number of metallo-enzymes have been determined, including alcohol dehydrogenase (Eklund *et al.*, 1976), carbonic anhydrase (Eriksson *et al.*, 1988), carboxypeptidase A (Reeke *et al.*, 1967) and

thermolysin (Kester and Matthews, 1977), and although other residues are able to interact with metal ions, the co-ordinating amino acids in these examples are histidines, glutamates and cysteines. Interestingly, as shown in figure 3.11, a number of these residues are absolutely conserved between these four biotin carboxylases, namely his^{146}, glu^{177}, and his^{345} (numbering relative to rat PC). Additionally a cysteine residue is conserved between yeast PC, ODC and TC5S which is represented by an asparagine residue (asn^{178}) in rat PC as mentioned above. This difference may reflect the fact that Zn(II) is the bound metal in these three proteins, while rat PC binds a Mn(II) ion.

Identification of the ion-coordinating amino acids in these proteins would provide valuable information regarding the conformation of the polypeptide chain, and hence the overall structure of these proteins. However, such experiments may be hampered by the apparent instability of PC in the absence of a bound metal ion (Carver *et al.*, 1988). One possible solution to such a problem is to introduce amino acids substitutions (by site directed mutagenesis) aimed at altering the metal ion binding preference of PC. That is, rather than attempting to create metal ion deficient molecules, such a regime would attempt to alter rat PC such that it now binds Zn(II). Alternately, attempts could be made to encourage yeast PC to bind Mn(II) or Mg(II) preferentially.

As detailed in section 1.4.1, the amino acid sequences of a number of biotin carboxylases have been published during the course of this study. The rapid pace of these advances, in addition to the constantly expanding sequence data bases (Genbank, EMBL etc), dictates the need to regularly update sequence comparisons. Lim *et al.* (1988) identified a number of regions of sequence similarity between various members of the biotin carboxylases, in addition to similarities with other proteins. With the advance of sequence information, other regions of sequence similarity have become evident, as shown schematically in figure 3.12.

The addition of the complete sequences of acetyl-CoA carboxylase (ACC) and propionyl-CoA carboxylase (PCC) to the original comparisons of Lim *et al.* (1988) has highlighted the dual nature of the ATP / HCO_3^- binding region of yeast PC. Significant sequence conservation between yeast PC (residues 147 - 467) and regions of rat ACC (residues 280 - 614, 32%) and rat PCC α (residues 184 - 499, 37%) were observed using the program FASTA, and are shown in figure 3.13. In addition, a portion of this conserved region is also significantly similar to the ATP-binding sequences of carbamyl phosphate synthetase (CPS), as shown in figure 3.14.

Of particular interest in these comparisons is the absolute conservation of a number of residues between the 7 sequences shown. Given the structural similarity between the product of CPS, namely carbamyl phosphate, and the proposed intermediate in the biotin carboxylation reaction (carboxy-phosphate, figure 3.15) it is tempting to suggest that these highly conserved residues are important for either maintenance of a particular peptide conformation or interaction with the ligands.

Also of interest is the cysteine residue (yeast PC - cys^{249}) which is absolutely conserved between the biotin carboxylases shown. This residue may therefore represent the catalytic cysteine detected in the first partial reaction site by Palacian and Neet (1972) and Tipton and Cleland (1988) as described in section 1.4.3. In addition, a number of lysine residues (yeast PC - lys^{178}, lys^{257} and lys^{284}) are absolutely conserved, one of which may represent the catalytically important lysine residue modified with the affinity label oATP by Easterbrook-Smith *et al.* (1976) and Lim (1988). Substitution of these residues in yeast PC by site directed mutagenesis, might be expected therefore to provide further evidence on the nature of the reaction occuring at the 1^{st} partial reaction site.

As indicated in figure 3.16 the region of the yeast PC sequence (residues 21 - 146) N-terminal to the ATP/HCO_3^- binding site (viz. residues 147-346) displays significant similarity (40%) to a region near the N-terminus of PCC α–subunit (residues 58 - 183). Interestingly, no corresponding sequences could be found in any other protein in the sequence database (Genbank release 63; EMBL release 20 and NBRF release 23, December 1989) and particularly, this conservation was not evident in the sequences of chicken or rat ACC. The significance of this region in the structures or functions of PC and PCC is not clear, although future comparison to sequences from as yet undetermined proteins may provide some clues.

Comparison of the sequence of the TC 12S subunit with the rat PCC β-subunit (figure 3.17) reveals a remarkably high level of sequence conservation, where amino acids 18 - 508 of TC 12S subunit are 46% identical to residues 39 - 541 of the PCC β–subunit . Although such a level of similarity may not be surprising, in that both of these subunits catalyse essentially the same reaction, converting between methylmalonyl-CoA and propionyl-CoA (see Table 1.1), it does have some implications for the structures of these two proteins which are difficult to interpret. As indicated in figure 3.12, a portion of this region is also conserved with a region of rat and chicken ACC (residues 1711-2019 and 1733-2041 respectively). As coenzyme-A binding is common to all three enzyme types,

this conserved region is likely to represent a coenzyme-A-binding site. The implications of the sequence homology between these proteins will be discussed further in section 5.2.4.

In order to investigate whether exons encode functional domains in PC, a comparison of the position of the intron-exon boundaries and the amino acid sequence is required. Shown schematically in figure 3.18 is the relationship between the position of introns in the coding sequence of rat PC and regions designated as functionally important by sequence comparisons (Lim *et al.*, 1988; figures 3.11, 3.13 & 3.17). As can been seen from this figure, the position of the most 3´ intron (intron A) coincides well with the proposed N-terminal boundary of the biotin-carrier domain (Lim *et al.*, 1988), and the adjacent intron (intron B) approximates the N-terminus of the 1.3S subunit of TC.

The pyruvate-binding domain is bounded by introns D and H, with the protease susceptible region (section 4.3.1) found between introns I and J (exon X). A more detailed inspection within the pyruvate-binding region, shows that the level of sequence similarity observed in figure 3.11 is not consistent along the length of the region. As shown in figure 3.15, the level of identity between the sequences differs markedly when the region is divided into the four corresponding exons of rat PC. The percent identity (39%) observed in exon VIII is significantly higher than that observed in the remaining exons, and hence is higher than the value averaged over the entire region shown (21%). This is consistent with the presence of discrete sub-domain regions within the pyruvate/metal ion -binding domain, each having individual functional significance such as pyruvate binding, metal ion binding, intersubunit binding and catalysis.

Although it is tempting to suggest that some degree of correlation exists between exons and functional domains in PC, the role or significance of such a relationship is presently unclear. A greater degree of knowledge regarding the functional and structural units within the enzyme is required before any meaningful conclusions may be drawn.

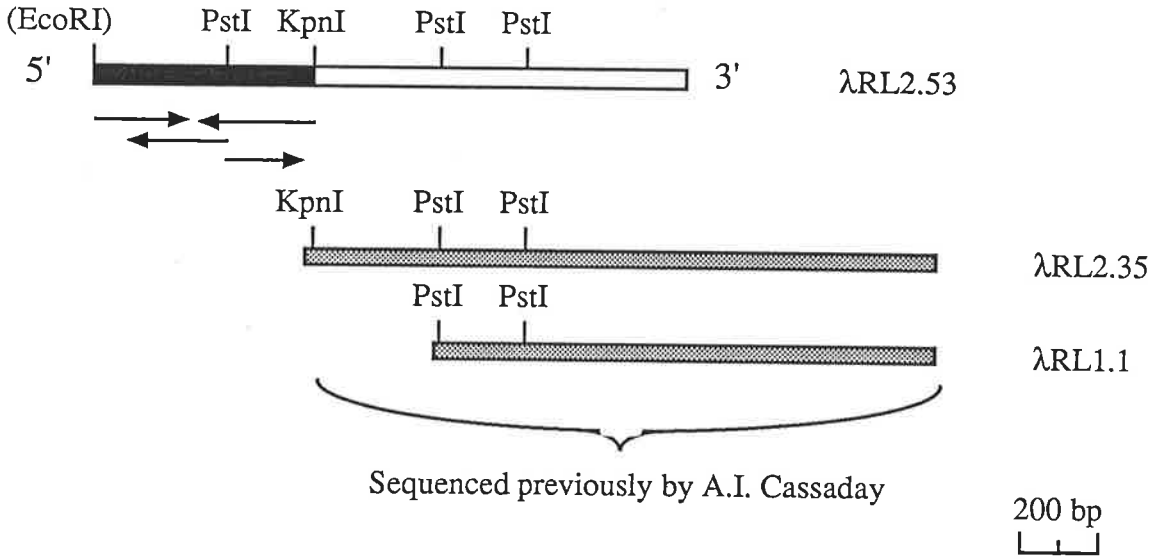


Figure 3.1 Partial restriction map and sequencing strategy of rat cDNA The λRL2.53 cDNA insert was digested with KpnI or KpnI/PstI and subcloned into M13 vectors mp18 and mp19 as described in section 2.2.21. The resultant subclones were sequenced by the chain termination method of Sanger *et al.* (1977) and the extent and orientation of each sequence is indicated by an arrow. The relative size and position of the cDNA inserts from cDNA clones λRL2.35 & λRL1.1 are shown for comparison. The regions of cDNA indicated by ▒▒ were sequenced by Cassady (1987).

Figure 3.2 Partial cDNA Sequence of Rat Pyruvate Carboxylase

The nucleotide sequence representing the 3´ 2067 bases of the rat pyruvate carboxylase mRNA are shown, along with the inferred amino acid sequence. This sequence represents a compilation of data derived from the sequences of cDNA clones λRL1.1, λRL2.35 & λRL2.53, which overlap as indicated in figure 3.1. The triangle represents the 5´extent of the previous sequence data as determined by Cassady (1987).

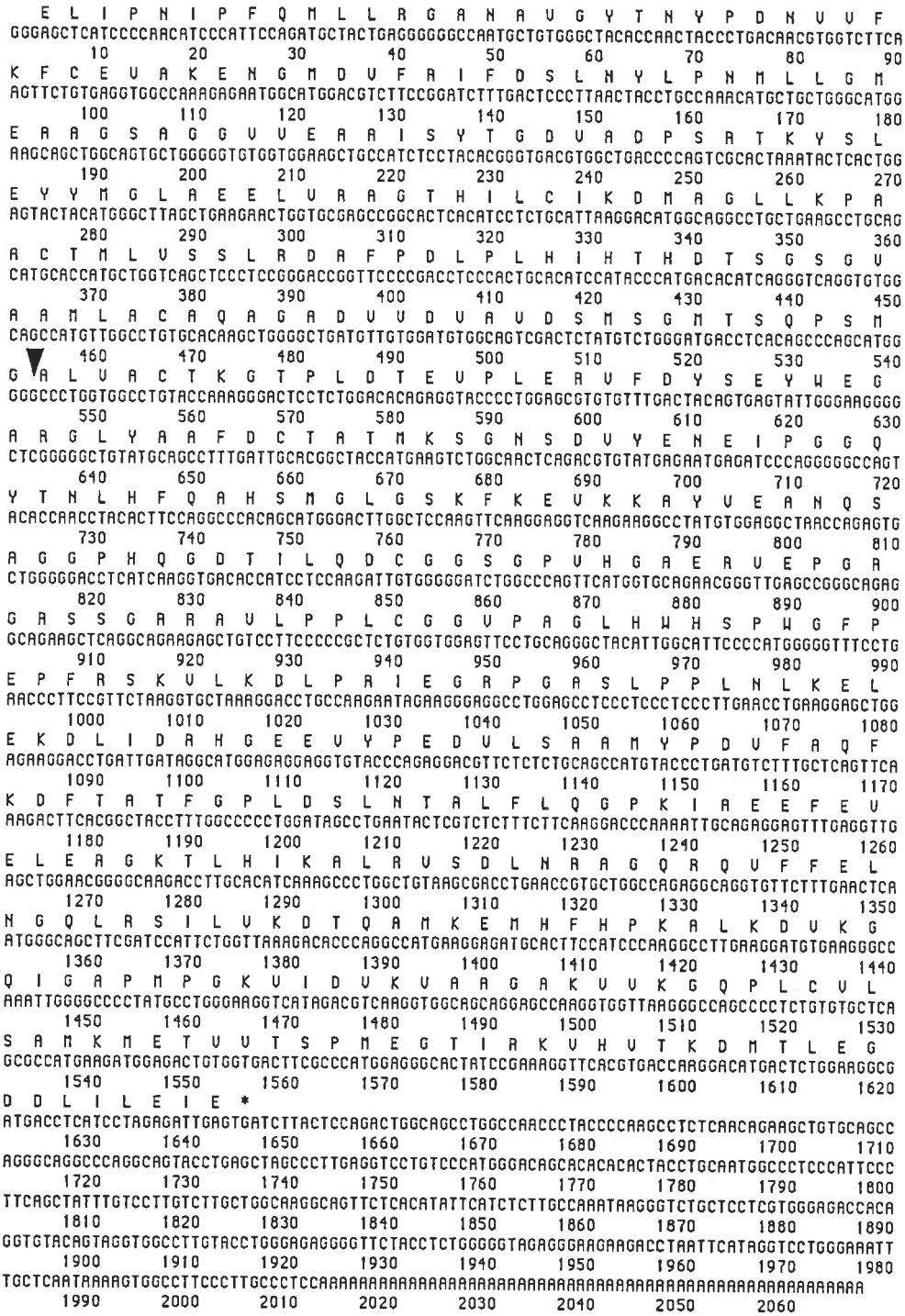


Figure 3.3 Computer Predicted Secondary Structure of a Portion of the Rat Pyruvate Carboxylase mRNA

The mRNA sequence of rat pyruvate carboxylase, represented by residues 1 - 1200 in the cDNA sequence (figure 3.2) was analysed for secondary structure propensity by the method of Zucker and Stiegler (1981).

The analysis was undertaken using the program FOLD and the results have been displayed in graphic form using the program SQUIGGLES. Due to the size limit of this program (1200 nucleotides) Residue 1 has been designated the 5′ end while residue 1200 has been designated the 3′ end. A Separate analysis of adjacent regions did not reveal significant alternative structures.

Regions of base pairing are indicated in the figure by a ladder, and the residue numbers are printed at intervals. The free energy (ΔG) calculated for this potential structure is -1840 kJ/mole (-440 kcal/mol).

1 (5') 1200 (3')

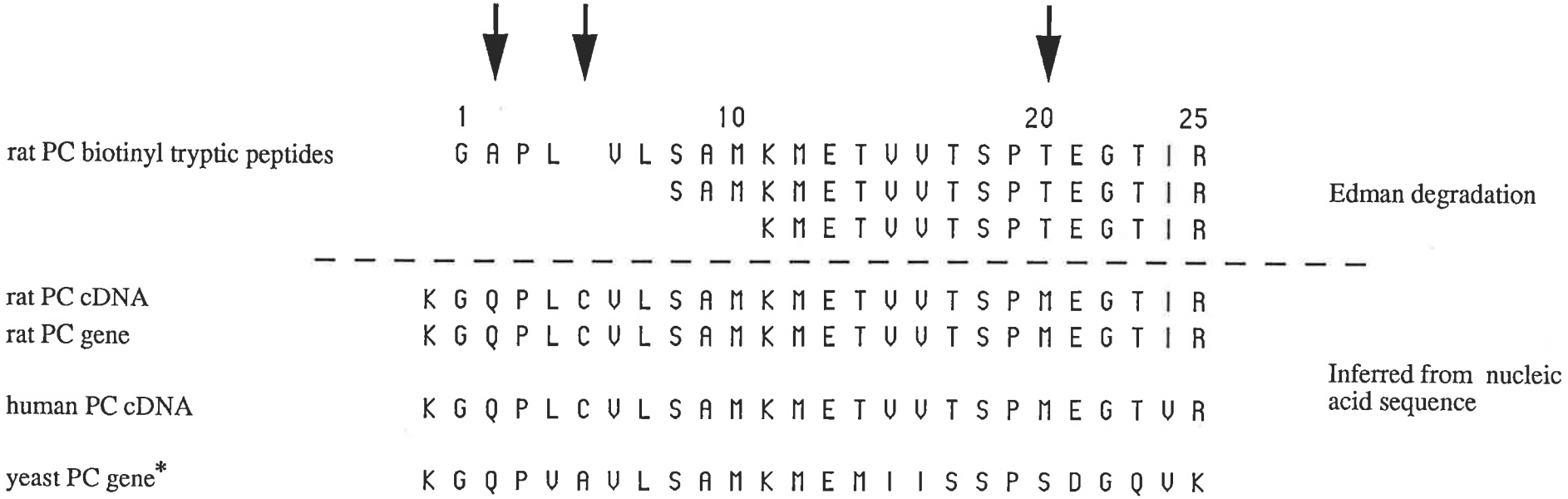


Figure 3.4 Comparison between protein seqences around the biotin attachment site of rat PC as predicted by Edman degradation and nucleic acid sequence. Rat PC biotinyl tryptic peptide sequences were published by Thampy *et al.* (1988); rat PC cDNA and genomic sequences were determined by Cassady (1987); human PC sequence was published by Freytag & Collier (1984) and yeast PC sequence was published by Lim et al. (1988). The arrows indicate the three positions where the Edman degradation results contradict the sequence inferred from nucleic acid sequences.

* - The sequence of the yeast PC biotin attachment site shown was also determined by Edman degradation of the biotinyl tryptic fragment by these authors.

Figure 3.5 Restriction Digests and Southern Analysis of pRGEB.1 insert DNA

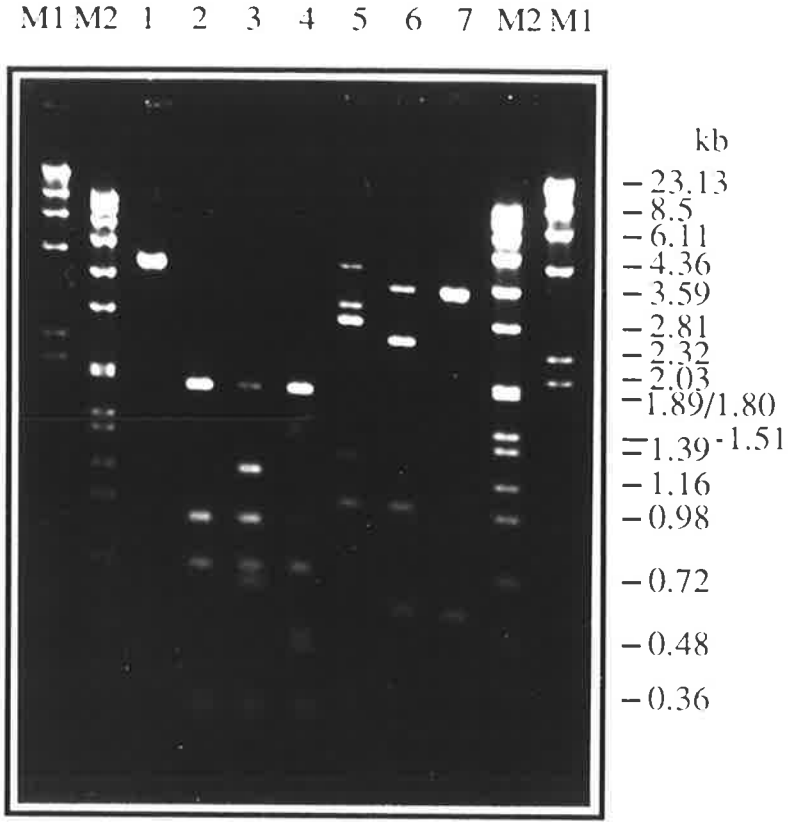
DNA representing the insert of rat pyruvate carboxylase genomic subclone pRGEB.1 was prepared by digestion of the plasmid with EcoRI and BamHI followed by electrophoresis on a 1% (w/v) agarose gel. The band representing the insert was excised from the gel and concentrated by ethanol precipitation.

A Restriction digests of pRGEB.1 insert DNA. 10 µg aliquots of this DNA were treated with the restriction endonucleases indicated and the products separated by electrophoresis through a 1% (w/v) agarose gel containing 1 x TBE, 0.3 µg/ml ethidium bromide. The molecular weight markers are shown in kilobases.

Track 1 - insert DNA (EcoRI-BamHI)
Track 2 - EcoRI-BamHI/PstI
Track 3 - EcoRI-BamHI/PstI/NarI
Track 4 - EcoRI-BamHI/PstI/KpnI
Track 5 - EcoRI-BamHI/Sal I/KpnI
Track 6 - EcoRI-BamHI/Bgl I/KpnI
Track 7 - EcoRI-BamHI/Bgl I
Track M1 - molecular weight marker - λ phage /HindIII
Track M2 - molecular weight marker - SPP1 phage /EcoRI

B Southern transfer of restriction digests of pRGEB.1 insert DNA. After separation of the restriction products shown above, the DNA was transferred to nitrocellulose by capillary action. The DNA was then probed with a 5′-[^{32}P]-labelled oligonucleotide (RPC02) derived from the sequence of the rat pyruvate carboxylase cDNA. The filter was washed in 1 x SSC/0.1% SDS at 42°C then autoradiographed for 20 hours at -80°C.

A


M1 M2 1 2 3 4 5 6 7 M2 M1
kb
23.13
8.5
6.11
4.36
3.59
2.81
2.32
2.03
1.89/1.80
1.39
1.51
1.16
0.98
0.72
0.48
0.36

B

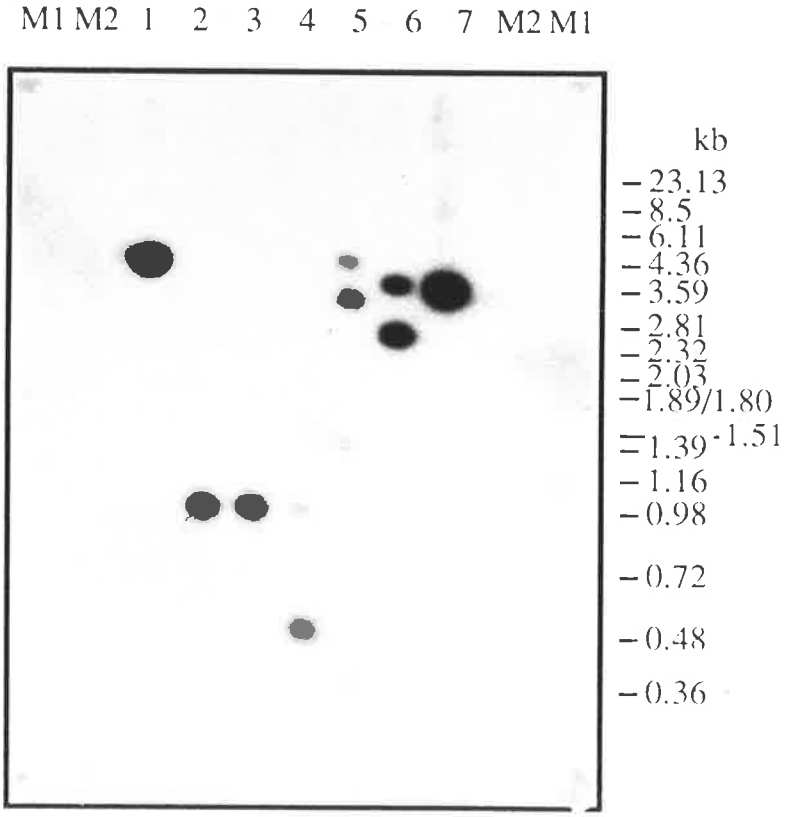
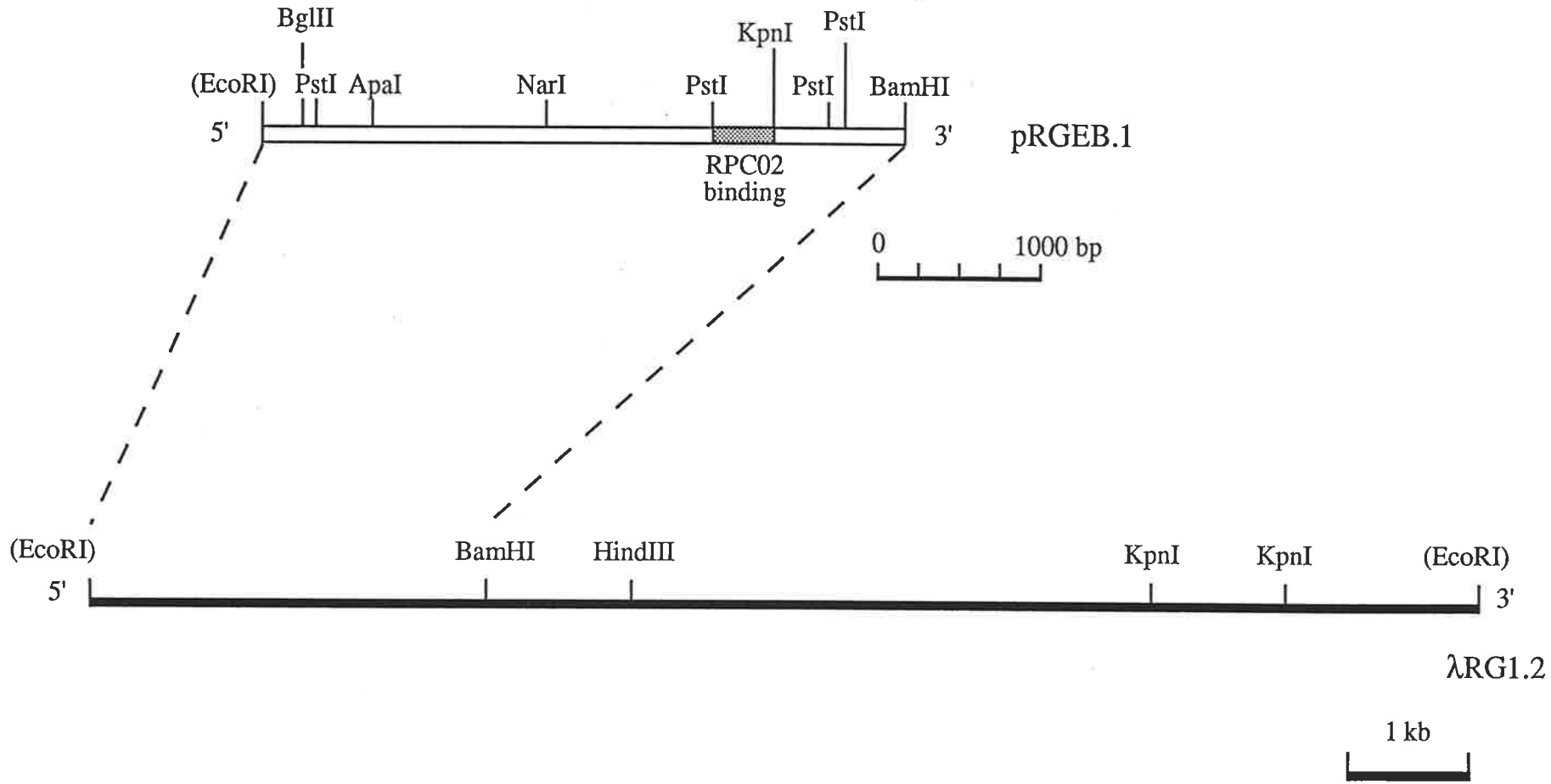

M1 M2 1 2 3 4 5 6 7 M2 M1
kb
23.13
8.5
6.11
4.36
3.59
2.81
2.32
2.03
1.89/1.80
1.39
1.51
1.16
0.98
0.72
0.48
0.36

Figure 3.6 Restriction Map of pRGEB.1 insert DNA

The restriction fragmentation pattern observed in figure 3.4 is summarised here. Also shown is the PstI-KpnI fragment to which the oligonucleotide RPC02 hybridises. The parent clone λRG1.2 (Cassady, 1987), representing the 3´ end of the rat pyruvate carboxylase gene, is shown for comparison. The restriction enzyme sites indicated within parentheses have been created during the cloning process.


BglII
KpnI
PstI
(EcoRI)
PstI
ApaI
NarI
PstI
PstI
BamHI
5'
3'
pRGEB.1
RPC02
binding
0
1000 bp
(EcoRI)
BamHI
HindIII
KpnI
KpnI
(EcoRI)
5'
3'
λRG1.2
1 kb

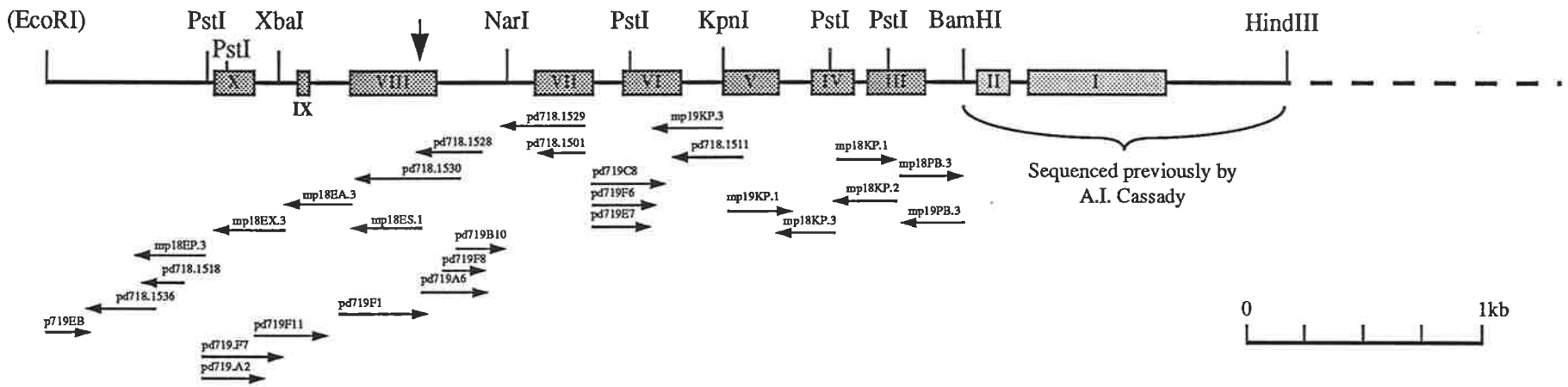
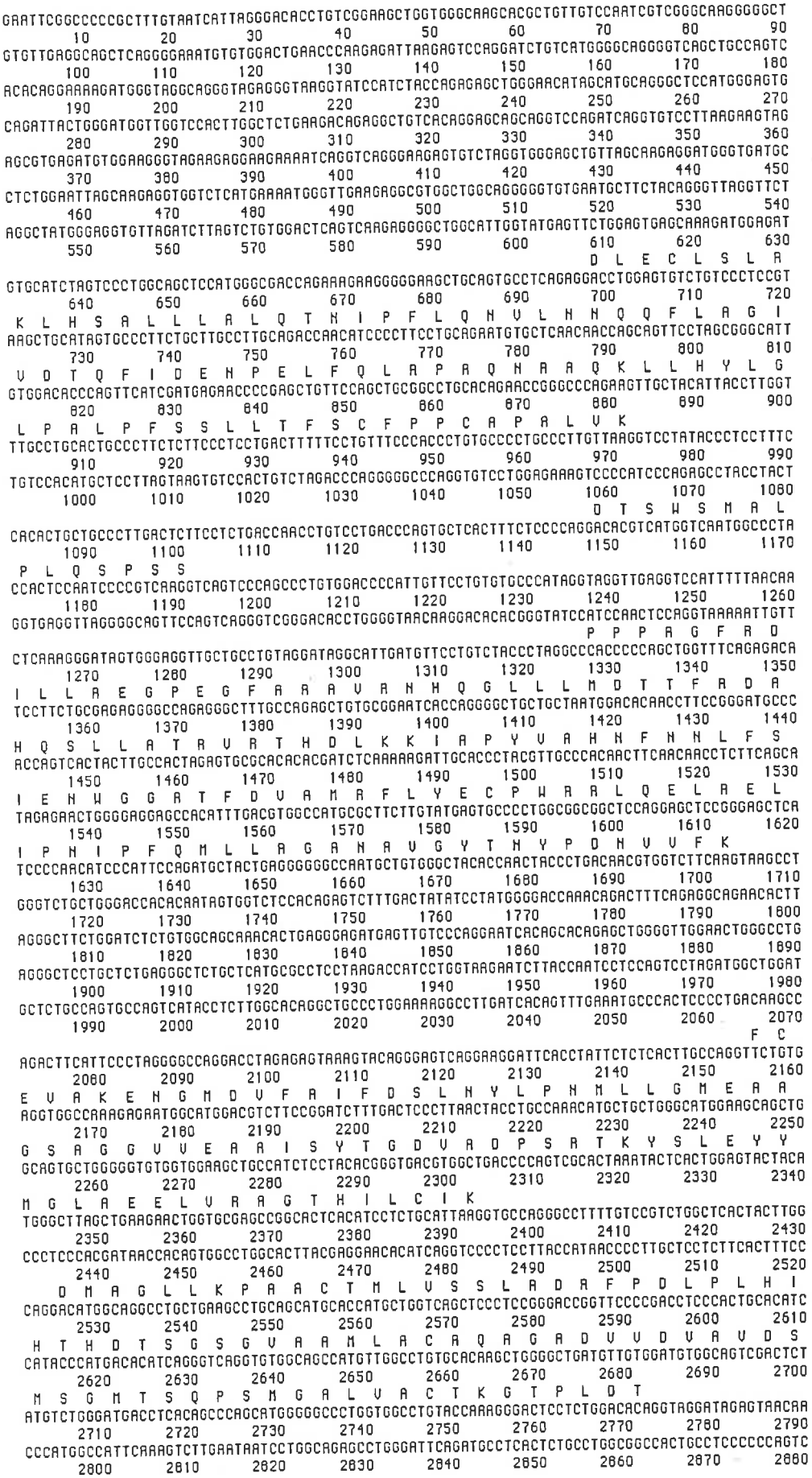


Figure 3.7 Sequencing strategy of EcoRI - BamHI fragment of rat PC gene The EcoRI - BamHI fragment of pRGEB.1 was subcloned and sequenced as described in section 2.2.26. The direction and extent of individual sequencing experiments are indicated by the horizontal arrows, along with the appropriate subclone title. The exonic regions are shown as boxed regions along the length of the 3' end of the rat PC gene and are denoted by roman numerals, with the most 3' exon as exon I. The 5' extent of the corresponding cDNA sequence is indicated by a vertical arrow. Exonic regions where only one strand of the DNA had been sequenced were confirmed by comparison to the cDNA sequence.



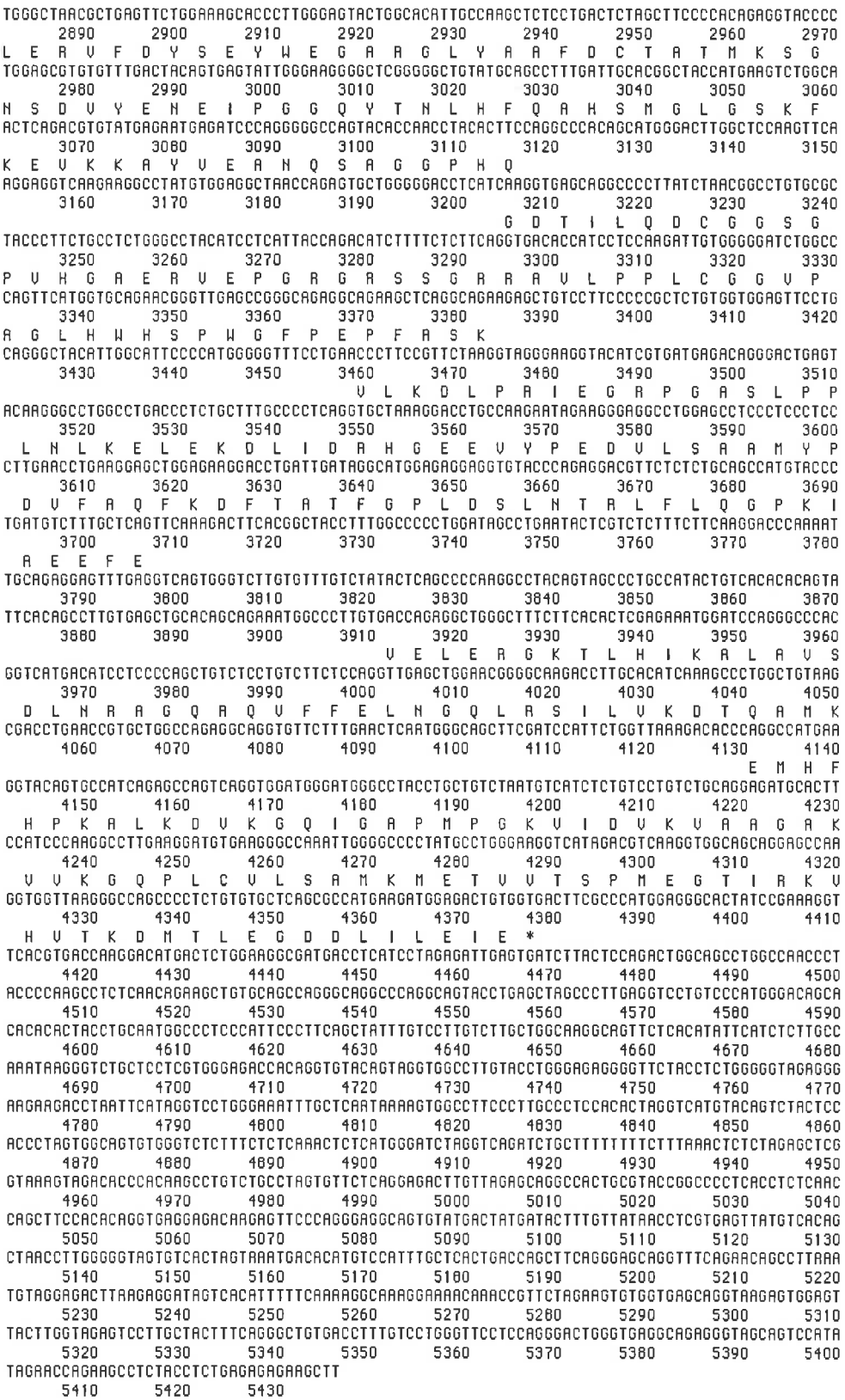


Figure 3.8 - PARTIAL GENE SEQUENCE OF RAT PYRUVATE CARBOXYLASE

The nucleotide sequence of the EcoRI - HindIII fragment from λRG1.2, along with the inferred protein sequence for rat PC.

Figure 3.9 **Comparison of the Nucleotide Sequences of the Acceptor and Donor Splice Sites of the Rat Pyruvate Carboxylase Gene**

Shown here are the nucleotide sequences for the excision sites of the 10 introns of the rat pyruvate carboxylase gene so far identified. Also shown are the splice consensus sites for messenger RNA as described by Cech (1983). Py represents positions that favour pyrimidine nucleotides (C & T). The vertical arrows indicate the site of cleavage of the precursor transcripts.

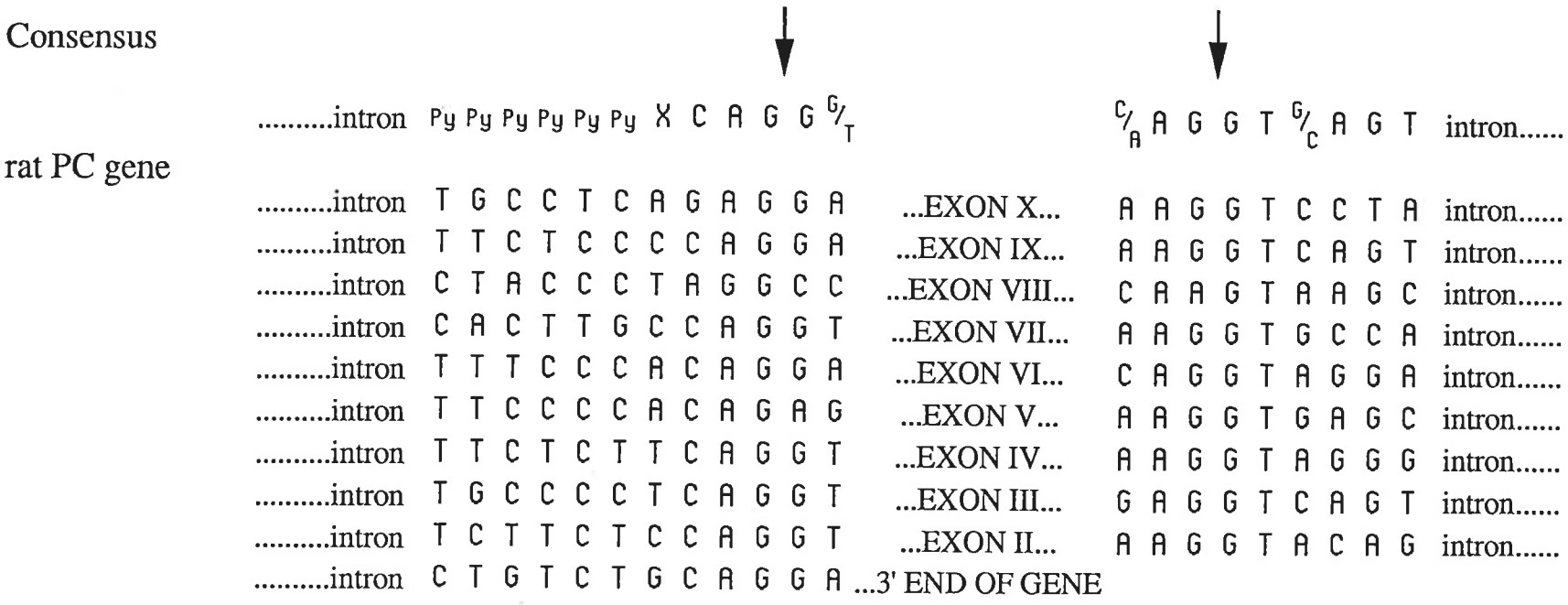

Consensus
.....intron Py Py Py Py Py Py X C A G G G/T C/A A G G T G/C A G T intron.....
rat PC gene
.....intron T G C C T C A G A G G A ...EXON X... A A G G T C C T A intron.....
.....intron T T C T C C C C A G G A ...EXON IX... A A G G T C A G T intron.....
.....intron C T A C C C T A G G C C ...EXON VIII... C A A G T A A G C intron.....
.....intron C A C T T G C C A G G T ...EXON VII... A A G G T G C C A intron.....
.....intron T T T C C C A C A G G A ...EXON VI... C A G G T A G G A intron.....
.....intron T T C C C C A C A G A G ...EXON V... A A G G T G A G C intron.....
.....intron T T C T C T T C A G G T ...EXON IV... A A G G T A G G G intron.....
.....intron T G C C C C T C A G G T ...EXON III... G A G G T C A G T intron.....
.....intron T C T T C T C C A G G T ...EXON II... A A G G T A C A G intron.....
.....intron C T G T C T G C A G G A ...3' END OF GENE

Table 3.1 Statistics of the Coding and Non-Coding Regions of the 3' end of the Rat PC Gene.

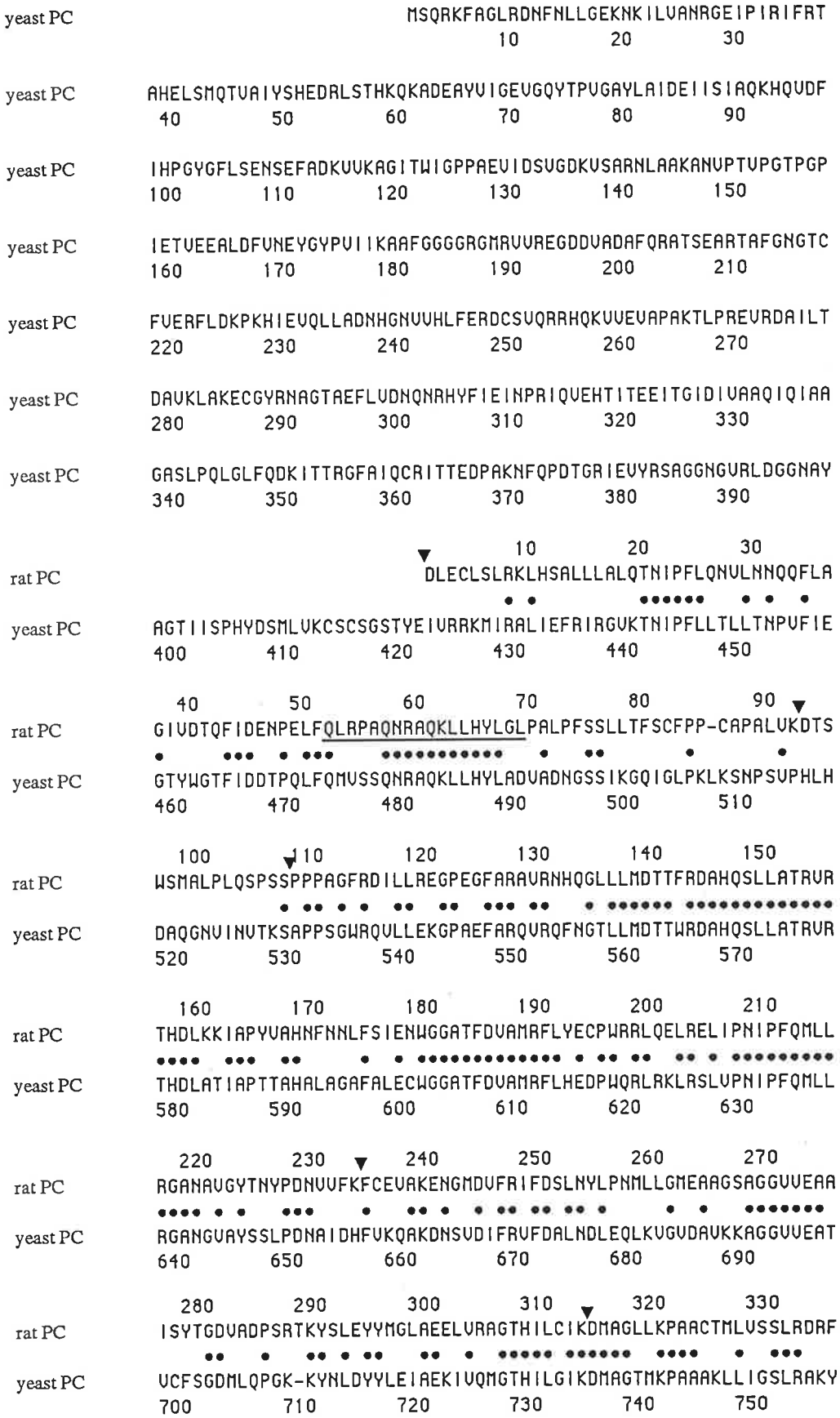
The number of nucleotides (bp) in each of the introns and exons so far determined for the rat PC gene are shown, along with the resulting average figure. The exons are designated by roman numerals starting with the 3' most exon. The introns are designated alphabetically, with the 3' most intron (found between exons I & II) as intron A. The 5' boundary for intron J has yet to be determined.

Table 3.2 Codon Usage.

The frequency at which the various codons are utilised within the partially characterised rat PC gene are shown. Tabulated using the program CODONUSAGE.

	Exons	(bp)	Introns	(bp)
3'	Exon I	614		
			Intron A	79
	Exon II	140		
			Intron B	205
	Exon III	248		
			Intron C	66
	Exon IV	187		
			Intron D	88
	Exon V	245		
			Intron E	189
	Exon VI	249		
			Intron F	131
	Exon VII	240		
			Intron G	451
	Exon VIII	377		
			Intron H	135
	Exon IX	43		
			Intron I	175
5'	Exon X	275		
			Intron J	>697
Average:		261		221
Ratio:		54%		46%

Gly	GGG	20	Arg	AGG	4	Trp	TGG	6	Arg	CGG	13
Gly	GGA	10	Arg	AGA	8	End	TGA	1	Arg	CGA	4
Gly	GGT	8	Ser	AGT	5	Cys	TGT	9	Arg	CGT	5
Gly	GGC	23	Ser	AGC	9	Cys	TGC	4	Arg	CGC	3
Glu	GAG	34	Lys	AAG	28	End	TAG	0	Gln	CAG	22
Glu	GAA	11	Lys	AAA	8	End	TAA	0	Gln	CAA	6
Asp	GAT	12	Asn	AAT	7	Tyr	TAT	5	His	CAT	10
Asp	GAC	29	Asn	AAC	19	Tyr	TAC	13	His	CAC	11
Val	GTG	29	Met	ATG	24	Leu	TTG	8	Leu	CTG	37
Val	GTA	2	Ile	ATA	3	Leu	TTA	1	Leu	CTA	10
Val	GTT	11	Ile	ATT	8	Phe	TTT	12	Leu	CTT	11
Val	GTC	9	Ile	ATC	14	Phe	TTC	23	Leu	CTC	21
Ala	GCG	1	Thr	ACG	4	Ser	TCG	1	Pro	CCG	3
Ala	GCA	17	Thr	ACA	5	Ser	TCA	10	Pro	CCA	13
Ala	GCT	19	Thr	ACT	10	Ser	TCT	7	Pro	CCT	14
Ala	GCC	31	Thr	ACC	15	Ser	TCC	11	Pro	CCC	20


yeast PC
MSQRKFAGLRDNFNLLGEKNKILVANRGEIPIRIFRT
10 20 30
yeast PC
AHELSMQTVAIYSHEDRLSTHKQKADEAYVIGEVGQYTPVGAYLAIDEIISIAQKHQVDF
40 50 60 70 80 90
yeast PC
IHPGYGFLSENSEFADKVVKAGITWIGPPAEVIDSVGDKVSARNLAAKANVPTVPGTPGP
100 110 120 130 140 150
yeast PC
IETVEEALDFVNEYGYPVIIKAAFGGGGRGMRVVREGDDVADAFQRATSEARTAFGNGTC
160 170 180 190 200 210
yeast PC
FVERFLDKPKHIEVQLLADNHGNVVHLFERDCSVQRRHQKVVEVAPAKTLPREVRDAILT
220 230 240 250 260 270
yeast PC
DAVKLAKECGYRNAGTAEFLVDNQNRHYFIEINPRIQVEHTITEEITGIDIVAAQIQIAA
280 290 300 310 320 330
yeast PC
GASLPQLGLFQDKITTRGFAIQCRITTEDPAKNFQPDTGRIEVYRSAGGNGVRLDGGNAY
340 350 360 370 380 390
10 20 30
rat PC
DLECLSLRKLHSALLLALQTNIPFLQNVLNNQQFLA
yeast PC
AGTIISPHYDSMLVKCSCSGSTYEIVRRKMIRALIEFRIRGVKTNIPFLLTLLTNPVFIE
400 410 420 430 440 450
40 50 60 70 80 90
rat PC
GIVDTQFIDENPELFQLRPAQNRAQKLLHYLGLPALPFSSLLTFSCFPP-CAPALVKDTS
yeast PC
GTYWGTFIDDTPQLFQMVSSQNRAQKLLHYLADVADNGSSIKGQIGLPKLKSNPSVPHLH
460 470 480 490 500 510
100 110 120 130 140 150
rat PC
WSMALPLQSPSSPPPAGFRDILLREGPEGFARAVRNHQGLLLMDTTFRDAHQSLLATRVR
yeast PC
DAQGNVINVTKSAPPSGWRQVLLEKGPAEFARQVRQFNGTLLMDTTWRDAHQSLLATRVR
520 530 540 550 560 570
160 170 180 190 200 210
rat PC
THDLKKIAPYVAHNFNNLFSIENWGGATFDVAMRFLYECPWRRLQELRELIPNIPFQMLL
yeast PC
THDLATIAPTTAHALAGAFALECWGGATFDVAMRFLHEDPWQRLRKLRSLVPNIPFQMLL
580 590 600 610 620 630
220 230 240 250 260 270
rat PC
RGANAVGYTNYPDNVVFKFCEVAKENGMDVFRIFDSLNYLPNMLLGMEAAGSAGGVVEAA
yeast PC
RGANGVAYSSLPDNAIDHFVKQAKDNSVDIFRVFDALNDLEQLKVGVDAVKKAGGVVEAT
640 650 660 670 680 690
280 290 300 310 320 330
rat PC
ISYTGDVADPSRTKYSLEYYMGLAEELVRAGTHILCIKDMAGLLKPAACTMLVSSLRDRF
yeast PC
VCFSGDMLQPGK-KYNLDYYLEIAEKIVQMGTHILGIKDMAGTMKPAAAKLLIGSLRAKY
700 710 720 730 740 750

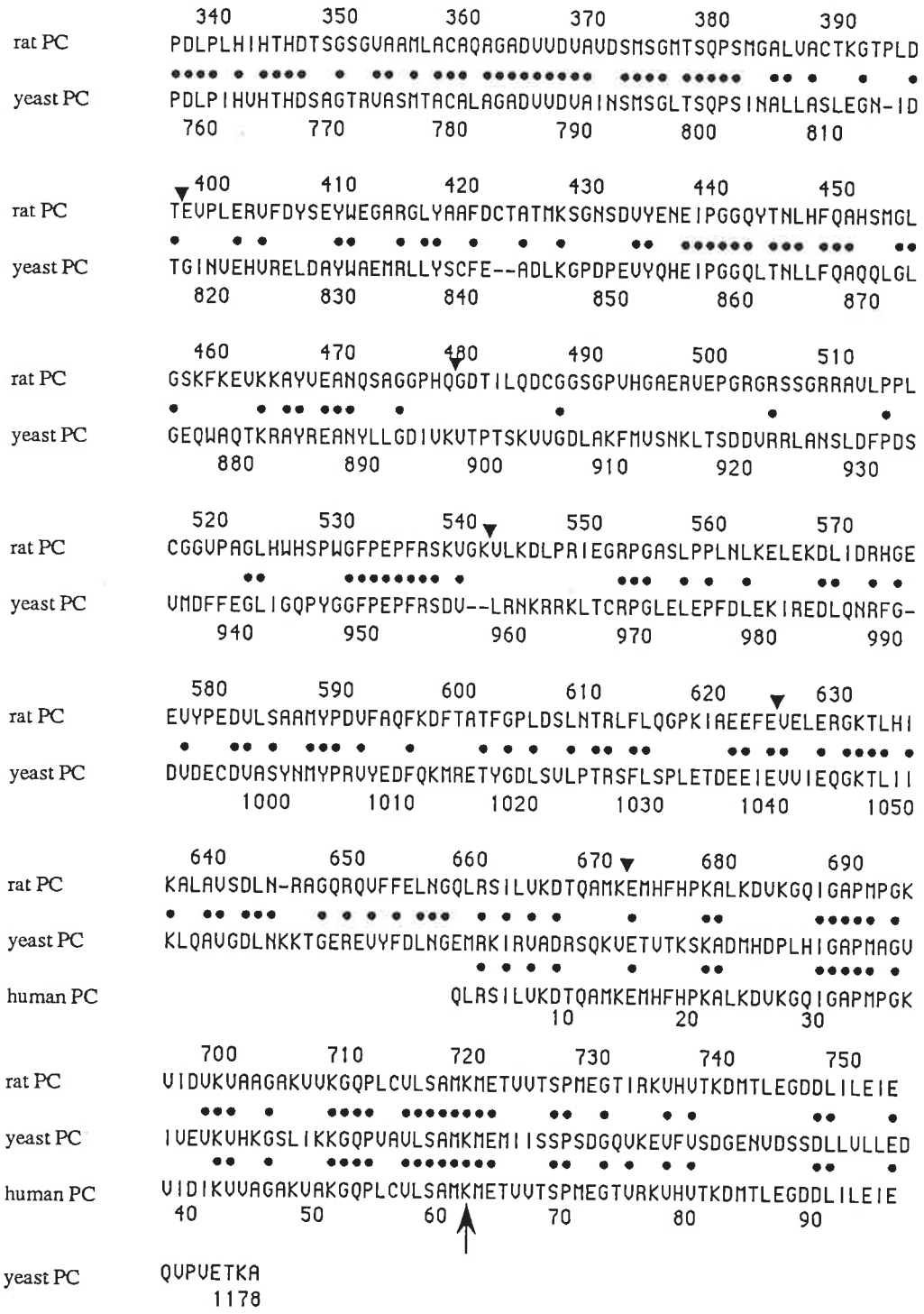


Figure 3.10 - Sequence comparison between rat PC, yeast PC and human PC. The partial amino acid sequence of rat PC, the complete sequence of yeast PC (Lim *et al.* (1988) and partial sequence of human PC (Lamhonwah *et al.*, 1987) were aligned for maximum similarity by the program Fasta. Residues which are identical between yeast PC and the other two sequences are denoted by (●). The lysine residue to which biotin is attached is marked by an arrow. Also indicated (▼) are the relative positions where introns interrupt the coding sequences of the rat PC gene. For rat and human PC, the residue numbers relate to the most N-terminal amino acid presently determined. The underlined sequence of rat PC represents the residues confirmed by peptide sequencing in section 4.2.3.

Figure 3.11 Comparison of the Amino Acid Sequences of the Pyruvate-binding Regions of Rat and Yeast Pyruvate Carboxylase, Oxaloacetate Decarboxylase and Transcarboxylase

The amino acid sequences of rat pyruvate carboxylase (RPC), yeast pyruvate carboxylase (YPC; Lim *et al.*, 1988), α-subunit of oxaloacetate decarboxylase from *Klebsiella pneumoniae* (ODC; Schwarz *et al.*, 1988) and 5S subunit of transcarboxylase from *Propionibacterium shermanii* (TC5S; Samols *et al.*, 1988) were aligned for maximum homology using the program SEQHP. Residues which are conserved in at least three of the proteins are highlighted. Gaps introduced into the sequences to maximise the alignment are indicated by dashes. All four proteins are known to bind pyruvate and divalent metal ions. Also indicated in the figure by (•) is the tryptophan residue of TC5S identified by Kumar *et al.* (1988b) as being associated with the pyruvate binding-site.

SEQUENCE SIMILARITIES IN THE PYRUVATE DOMAIN OF BIOTIN CARBOXYLASES

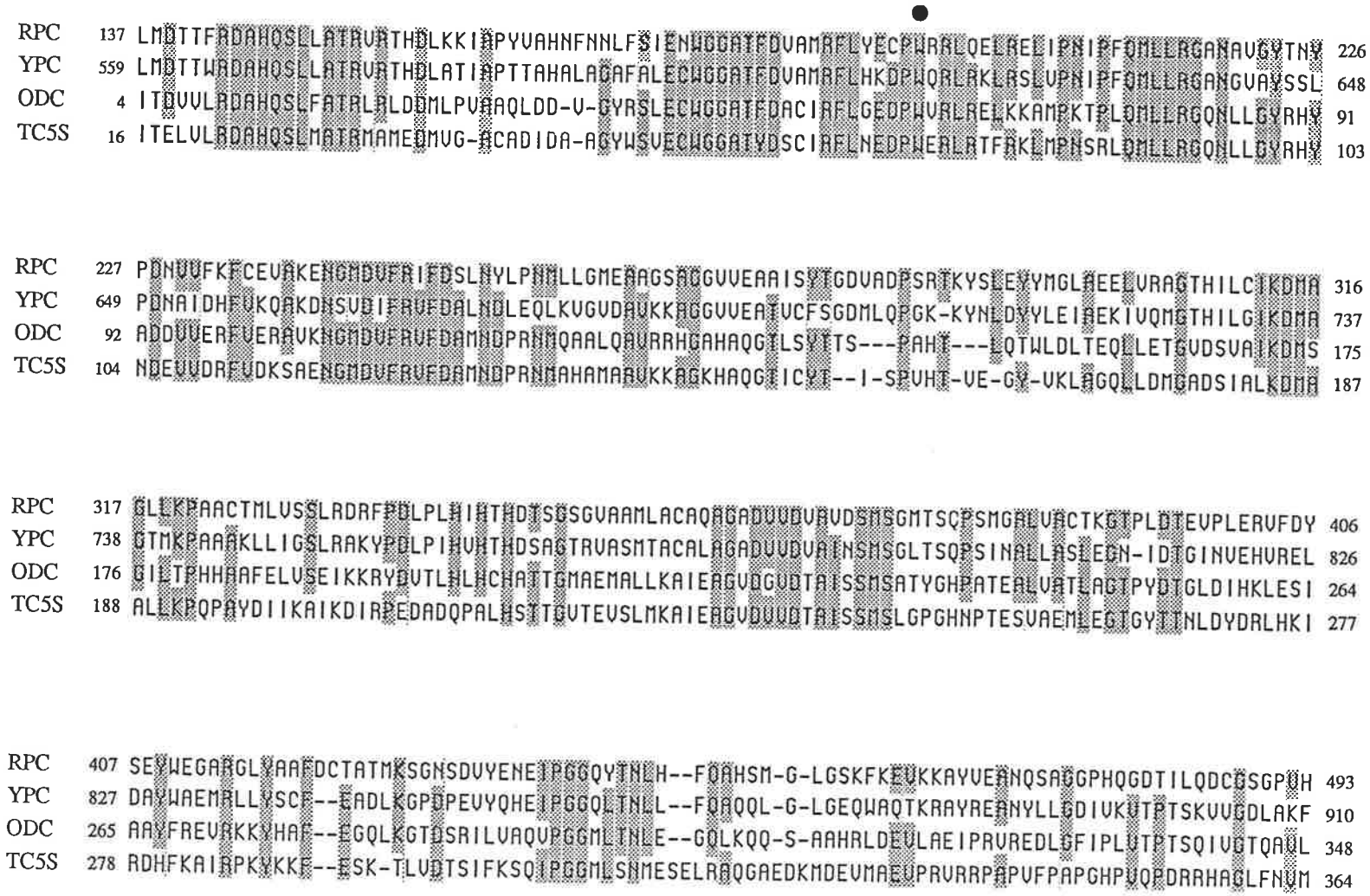
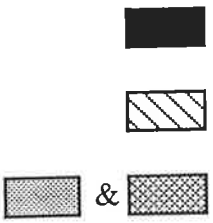


Figure 3.12 Schematic Representation of Sequence Conservations Observed between Biotin Carboxylases and Other Enzyme Groups

Regions of sequence homology between members of the biotin-dependent carboxylases and other enzyme classes, updated and adapted from Lim *et al.* (1988). Sequences are from:
PC (yeast) - Lim *et al.* (1988)
rat acetyl-CoA carboxylase - Lopez-Casillas *et al.* (1988)
chicken acetyl-CoA carboxylase - Takai *et al.* (1988)
rat propionyl-CoA carboxylase (α) Browner *et al.* (1989)
& (β) Kraus *et al.* (1986)
human propionyl-CoA carboxylase - (α) Lamhonwah *et al.* (1989)
& (β) Lamhonwah *et al.* (1986)
oxaloacetate decarboxylase - (α) Schwarz *et al.* (1988)
& (β,γ) Laußermair *et al.* (1989)
Transcarboxylase - (12S) Thornton *et al.* (1987), (5S) Samols *et al.* (1988) & (1.3S) Maloy *et al.* (1979)
yeast carbamyl phosphate synthetase - (large) Lusty *et al.* (1983) & (small) Nyunoya & Lusty (1984)
E. coli carbamyl phosphate synthetase - (large) Nyunoya *et al.* (1983)
rat carbamyl phosphate synthetase - Nyunoya *et al.* (1985)
E. coli pyruvate dehydrogenase - Guest *et al.* (1985)



represents biotin/lipoyl-carrying domain

represents pyruvate-binding region

represent ATP-binding regions

represents the coenzyme-A-binding region

represents the transcarboxylation region

represents a conserved region of undetermined function

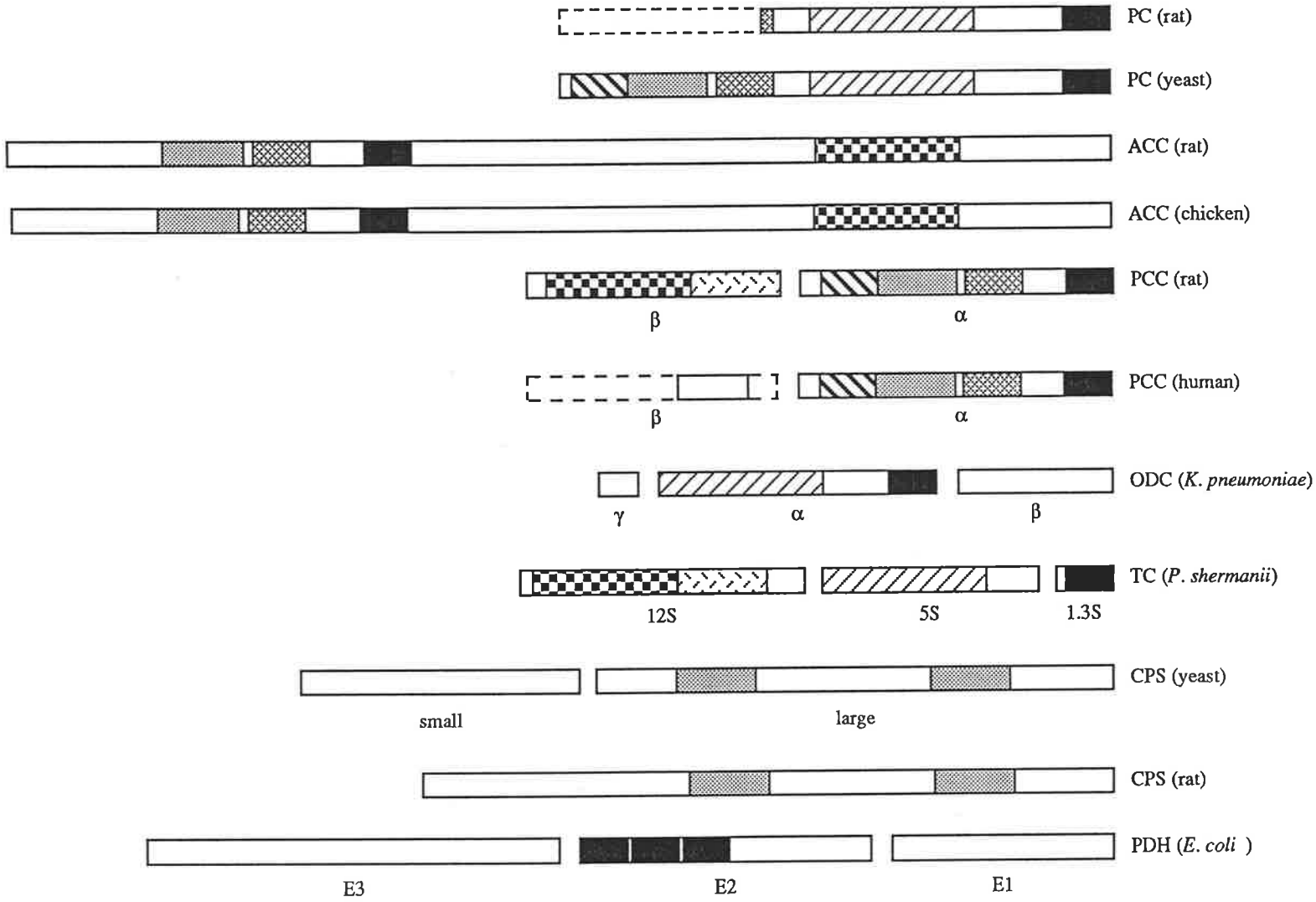

PC (rat)
PC (yeast)
ACC (rat)
ACC (chicken)
PCC (rat)
β
α
PCC (human)
β
α
ODC (K. pneumoniae)
γ
α
β
TC (P. shermanii)
12S
5S
1.3S
CPS (yeast)
small
large
CPS (rat)
PDH (E. coli)
E3
E2
E1

Figure 3.13 Comparison of the Amino Acid Sequences of the ATP/HCO_3^--binding Regions of Rat and Yeast Pyruvate Carboxylase, the α Subunit of Rat Propionyl-CoA Carboxylase and rat acetyl-CoA Carboxylase

The amino acid sequences of rat pyruvate carboxylase (RPC), yeast pyruvate carboxylase (YPC; Lim *et al.*, 1988), the α-subunit of rat propionyl-CoA carboxylase (ratPCC; Browner *et al.*, 1989) and rat acetyl-CoA carboxylase (ratACC; Lopez-Casillas *et al.*, 1988) were aligned for maximum homology using the program Fasta. Residues which are conserved between the proteins are highlighted. Gaps introduced into the sequences to maximise the alignment are indicated by dashes. All four proteins are known to bind ATP and HCO_3^-. Also indicated in the figure (•) is the conserved cysteine residue implicated as the catalytic base.

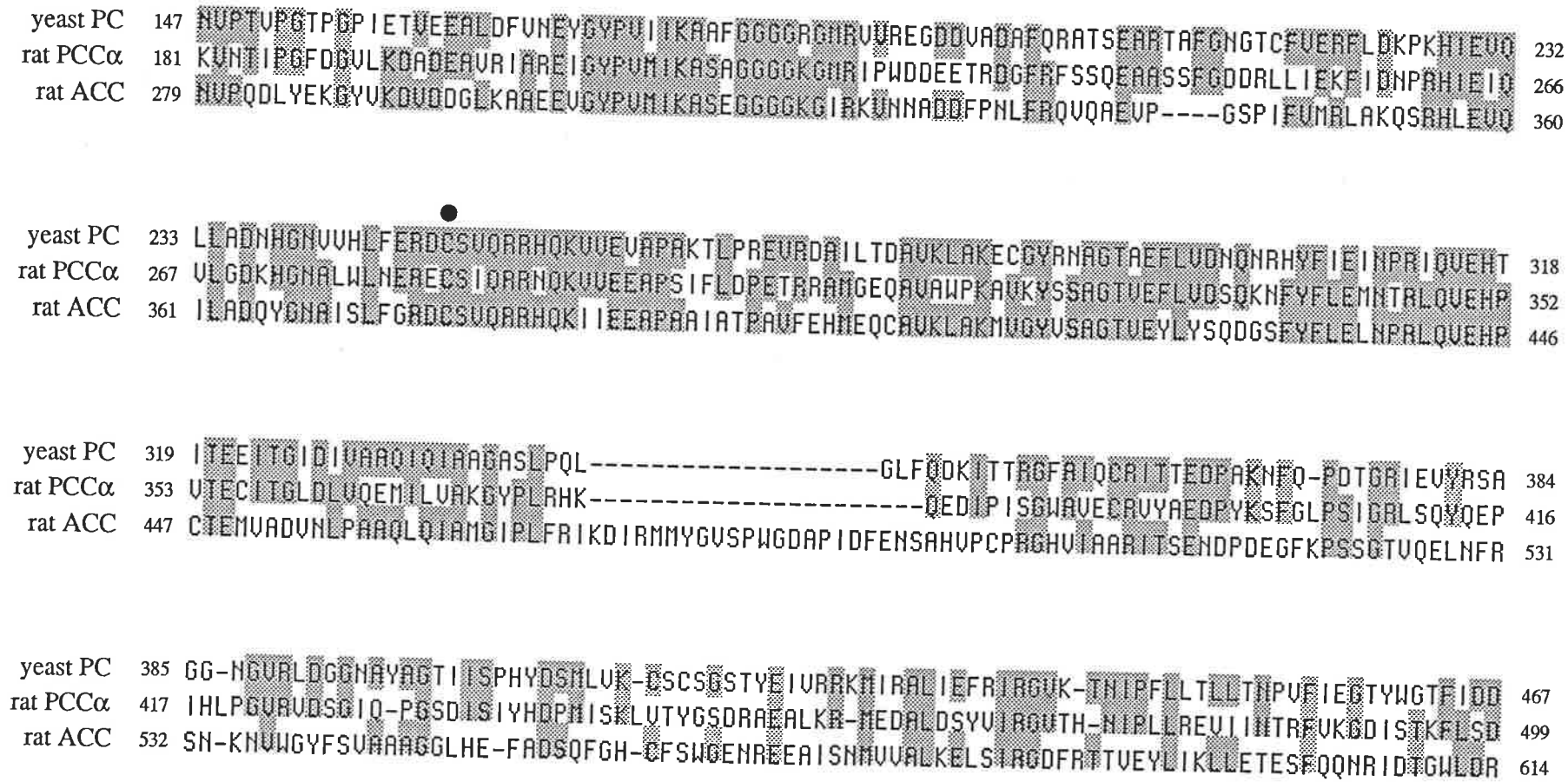
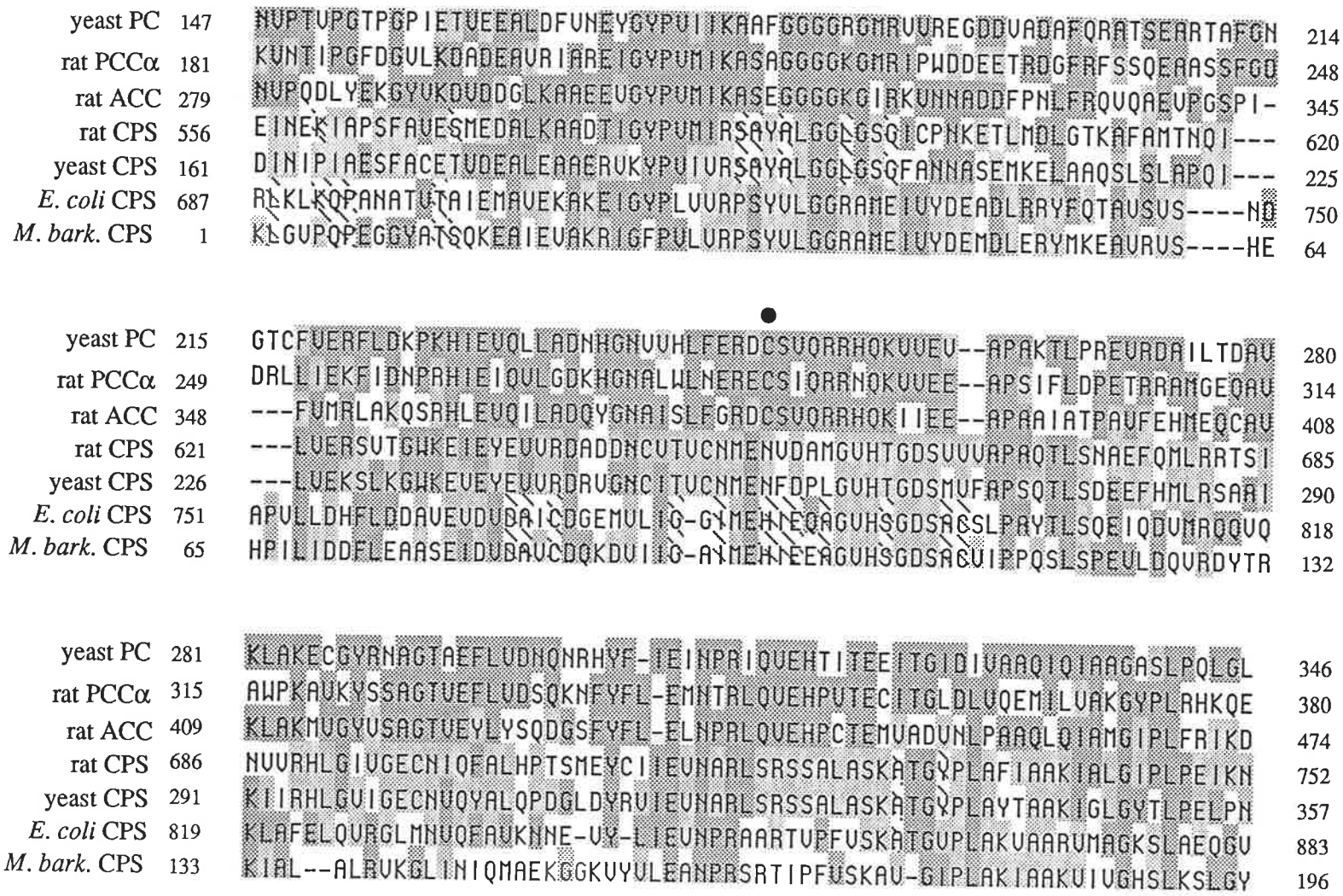
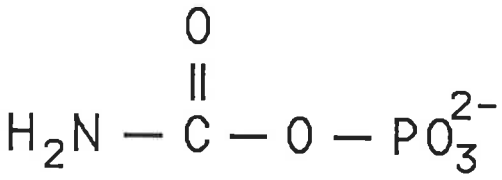

yeast PC 147 NVPTVPGTPGPIETVEEALDFVNEYGYPVIIKAAFGGGGRGMRVVREGDDVADAFQRATSEARTAFGNGTCFVERFLDKPKHIEVQ 232
rat PCCα 181 KVNTIPGFDGVLKDADEAVRIAREIGYPVMIKASAGGGGKGMRIPWDDEETRDGFRFSSQEAASSFGDDRLLIEKFIDNPRHIEIQ 266
rat ACC 279 NVPQDLYEKGYVKDVDDGLKAAEEVGYPVMIKASEGGGGKGIRKVNNADDFPNLFRQVQAEVP----GSPIFVMRLAKQSRHLEVQ 360
yeast PC 233 LLADNHGNVVHLFERDCSVQRRHQKVVEVAPAKTLPREVRDAILTDAVKLAKECGYRNAGTAEFLVDNQNRHYFIEINPRIQVEHT 318
rat PCCα 267 VLGDKHGNALWLNERECSIQRRNQKVVEEAPSIFLDPETRRAMGEQAVAWPKAVKYSSAGTVEFLVDSQKNFYFLEMNTRLQVEHP 352
rat ACC 361 ILADQYGNAISLFGRDCSVQRRHQKIIEEAPAAIATPAVFEHMEQCAVKLAKMVGYVSAGTVEYLYSQDGSFYFLELNPRLQVEHP 446
yeast PC 319 ITEEITGIDIVRAQIQIAAGASLPQL----------GLFQDKITTRGFAIQCRITTEDPAKNFQ-PDTGRIEVYRSA 384
rat PCCα 353 VTECITGLDLVQEMILVAKGYPLRHK----------QEDIPISGWAVECRVYAEDPYKSFGLPSIGRLSQYQEP 416
rat ACC 447 CTEMVADVNLPAAQLQIAMGIPLFRIKDIRMMYGVSPWGDAPIDFENSAHVPCPRGHVIAARITSENDPDEGFKPSSGTVQELNFR 531
yeast PC 385 GG-NGVRLDGGNAYAGTIISPHYDSMLVK-CSCSGSTYEIVRRKMIRALIEFRIRGVK-TNIPFLLTLLTNPVFIEGTYWGTFIDD 467
rat PCCα 417 IHLPGVRVDSGIQ-PGSDISIYHDPMISKLVTYGSDRAEALKR-MEDALDSYVIRGVTH-NIPLLREVIIHTRFVKGDISTKFLSD 499
rat ACC 532 SN-KNVWGYFSVAAAGGLHE-FADSQFGH-CFSWGENREEAISNMVVALKELSIRGDFRTTVEYLIKLLETESFQQNRIDTGWLDR 614

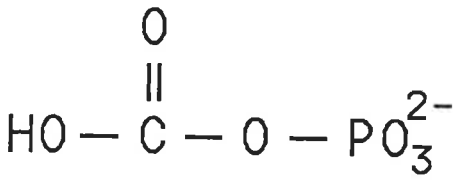
Figure 3.14 Comparison of the Amino Acid Sequences of the ATP/HCO_3^--binding Regions of the Biotin Carboxylases with Carbamyl Phosphate Synthetase

The amino acid sequences of yeast pyruvate carboxylase (YPC; Lim *et al.*, 1988), the α-subunit of rat propionyl-CoA carboxylase (ratPCC; Browner *et al.*, 1989), rat acetyl-CoA carboxylase (ratACC; Lopez-Casillas *et al.*, 1988) and carbamyl phosphate synthetase from rat (rat CPS; Nyunoya *et al.*, 1985), yeast (yeast CPS; Lusty *et al.*, 1983), *Escherichia coli* (*E. coli* CPS; Nyunoya and Lusty, 1983) and *Methanosarcina barkeri* (*M. barkeri* CPS; Morris and Reeve, 1988) were aligned for maximum homology using the program FASTA. Residues which are conserved between the proteins are highlighted. Also indicated in the figure (•) is the cysteine residue conserved between the biotin carboxylases. Gaps introduced into the sequences to maximise the alignment are indicated by dashes.


yeast PC 147 214
rat PCCα 181 248
rat ACC 279 345
rat CPS 556 620
yeast CPS 161 225
E. coli CPS 687 750
M. bark. CPS 1 64
yeast PC 215 280
rat PCCα 249 314
rat ACC 348 408
rat CPS 621 685
yeast CPS 226 290
E. coli CPS 751 818
M. bark. CPS 65 132
yeast PC 281 346
rat PCCα 315 380
rat ACC 409 474
rat CPS 686 752
yeast CPS 291 357
E. coli CPS 819 883
M. bark. CPS 133 196



A. carbamyl phosphate



B. carboxy phosphate

Figure 3.15 Chemical structures of carbamyl phosphate and carboxy phosphate (proposed reaction intermediate in the carboxylation of biotin, Knowles, 1989)

Figure 3.16 Comparison of the Amino Acid Sequences of Yeast Pyruvate Carboxylase and the α Subunit of Rat Propionyl-CoA Carboxylase at Their N-Termini

The amino acid sequences of yeast pyruvate carboxylase (YPC; Lim *et al.*, 1988) and the α-subunit of rat propionyl-CoA carboxylase (ratPCC; Browner *et al.*, 1988) were aligned for maximum homology using the program FASTA. The region shown is that portion of the total conserved region which is not also conserved in rat acetyl-CoA carboxylase (ratACC; Lopez-Casillas *et al.*, 1988). Residues which are conserved between the two proteins are highlighted in the figure by (•). Gaps introduced into the sequences to maximise the alignment are indicated by dashes.

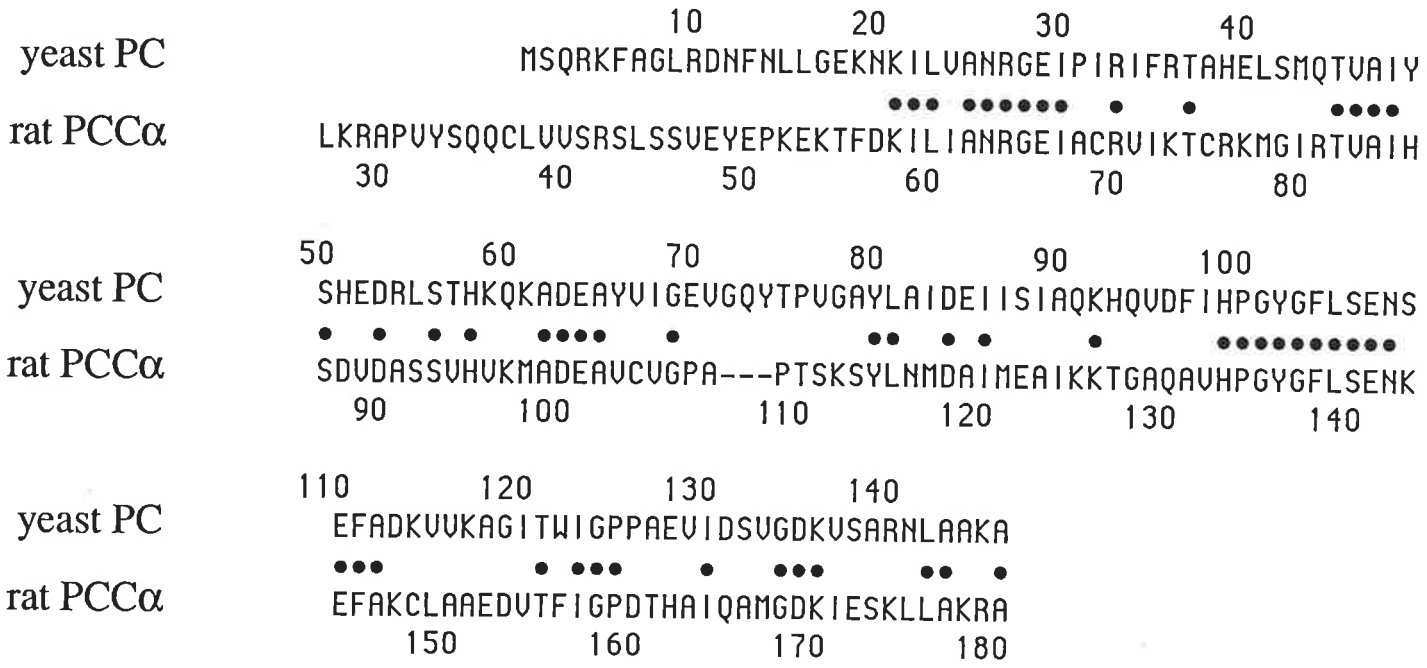
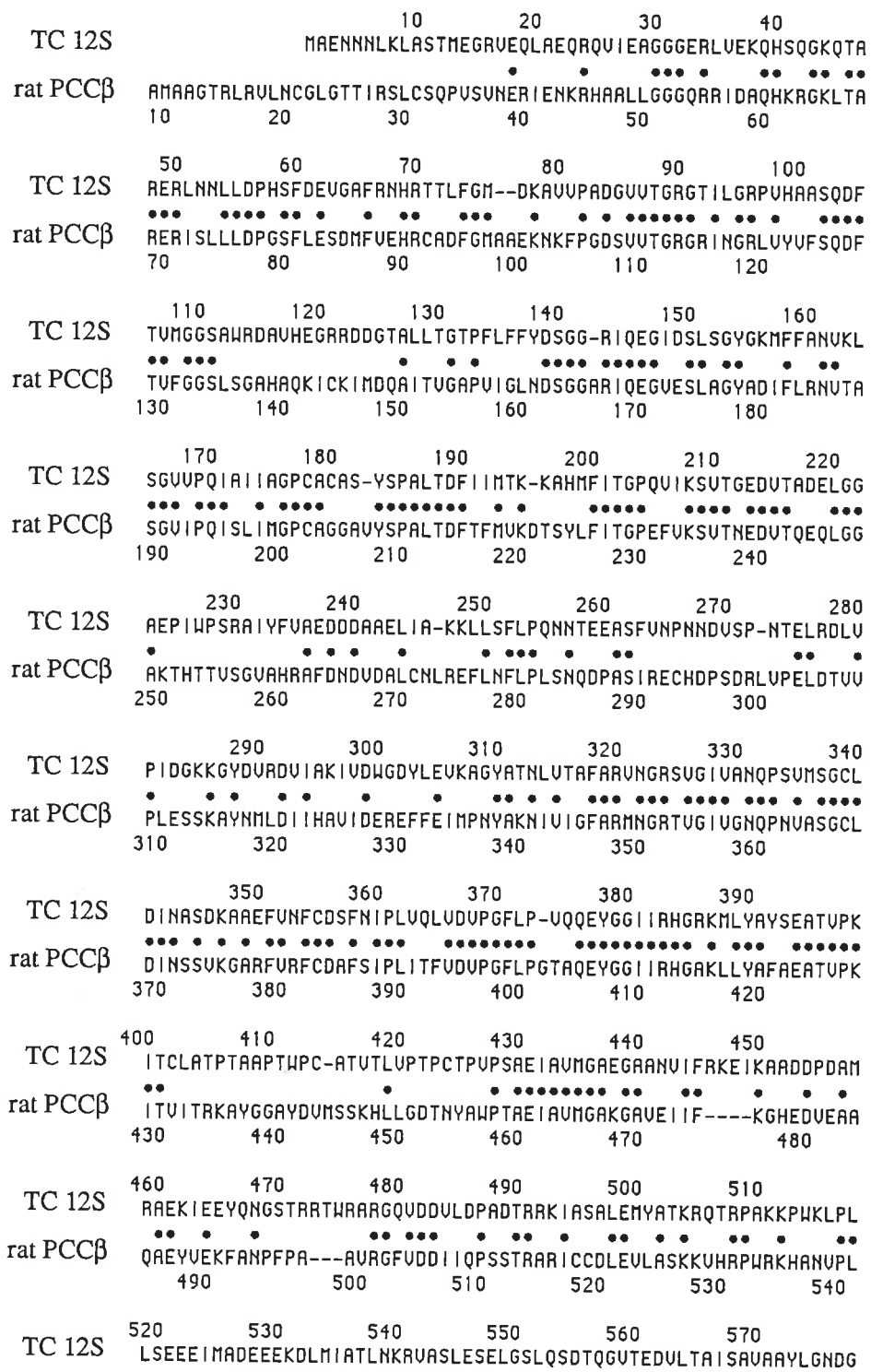

yeast PC
rat PCCα
yeast PC
rat PCCα
yeast PC
rat PCCα

Figure 3.17 Comparison of the Amino Acid Sequences of the Rat Propionyl-CoA Carboxylase β-subunit and Transcarboxylase 12S-monomer

Regions of the amino acid sequences of rat propionyl-CoA carboxylase β-subunit (Rat PCCβ; Kraus *et al.*, 1986) and 12S monomer of transcarboxylase from *Propionibacterium shermanii* (TC12S; Thornton *et al.*, 1987) were aligned for maximum homology using the program FASTA. Residues which are conserved between the proteins are indicated in the figure by (•). Gaps introduced into the sequences to maximise the alignment are indicated by dashes.


TC 12S
rat PCCβ
TC 12S
rat PCCβ
TC 12S
rat PCCβ
TC 12S
rat PCCβ
TC 12S
rat PCCβ
TC 12S
rat PCCβ
TC 12S
rat PCCβ
TC 12S

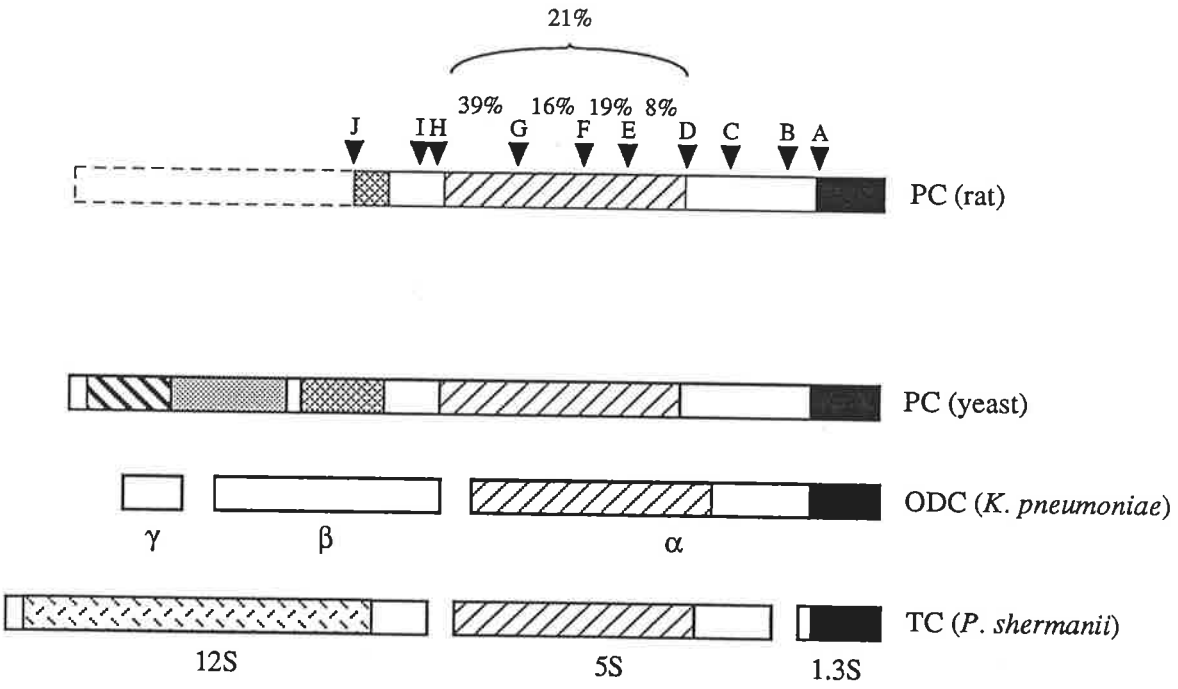


Figure 3.18 Relative Positions of Introns in Rat PC. The sequence of rat PC is shown schematically, indicating the regions of homology with other members of the biotin carboxylase family. The position of the corresponding intronic sequences in the gene coding for rat PC are indicated by the letters. Also shown is the level of identity between the four protein sequences shown in figure 3.11 (expressed as a %). ■ represents biotin-carrier domain, ▨ represents pyruvate-binding region, ▩ & ▩ represent ATP-binding regions, ▨ represents the coenzyme-A/transcarboxylation subsite and ▨ represents a region of undetermined function. Sequences are from: YPC - Lim *et al.* (1988); ODC - Schwarz *et al.* (1988) & Laußermair *et al.* (1989); TC - Thornton *et al.* (1987), Samols *et al.* (1988) & Maloy *et al.* (1979).

CHAPTER 4

THE DOMAIN STRUCTURE

4.1 INTRODUCTION

At present, two methods are available to provide high resolution 3-dimensional structures of proteins and enzymes, X-ray crystallography and 2-dimensional NMR spectroscopy. X-ray diffraction of protein crystals has been applied to the analysis of a number of enzymes including lysozyme [EC 3.2.1.17] (Blake *et al.*, 1965), citrate synthase [EC 4.1.3.7] (Remington *et al.*, 1982), lactate dehydrogenase [EC 1.1.1.27] (Musik and Rossman, 1979) and aspartate carbamoyltransferase [EC 2.1.3.2] (Krause *et al.*, 1987), providing precise atomic coordinates for the peptide backbone and amino acid side chains. In some cases, such as aspartate carbamyltransferase and citrate synthase, crystals have been isolated that incorporate substrate or product molecules, thereby providing detailed evidence for the catalytic mechanisms involved. More recently, with the advent of 2-dimensional NMR spectroscopy, detailed structures of relatively small (≤ Mr 40,000) molecules can be determined *in solution*, including interactions with substrates and other effector molecules (for a review see Bax, 1989).

Both these techniques have limitations however, in that relatively large quantities of homogeneous preparations are required, and in the instance of X-ray diffraction, these preparations have to be encouraged to form high quality crystals. Attempts to crystallise PC from various sources, under a variety of regimes, have so far failed to produce crystals of the enzyme suitable for such analysis (Lim, 1988; J. Wallace *pers. comm.*). In order to propose 3-dimensional structures which include the positions of the amino acid side chains within the protein molecules, diffraction patterns giving a resolution to at least 3Å are required (Fersht, 1985). PC, at Mr 520,000 per tetramer, is a large molecule compared to most proteins so far studied by X-ray diffraction, suggesting that even if crystal formation could be achieved, determination of the structure to such a resolution would not be guaranteed.

At high concentrations (10mg/ml) PC from yeast has been observed in the electron microscope to form long ordered chains, much like a row of fallen dominoes. No such ordered arrays were observed, however, to extend into the second or third dimensions (F. Mayer, *pers. comm.*). The most likely explanation for this is that the outer tip of each monomer is too flexible to allow these ordered rows to combine to make regular arrays and finally crystals. In any case it seems unlikely at this point that crystals suitable for X-ray diffraction can be formed readily by PC, and especially not within the scope of this study.

The large size of the enzyme and its tetrameric nature have so far limited the use of 2-dimensional NMR as a tool to study the structure of PC. In general, high resolution 2-dimensional studies are limited to monomeric or dimeric proteins due to an experimental

requirement for a correlation time (τ_c) of less than 10 ns (Bax, 1989). Although some proton-relaxation NMR studies have been undertaken into the interactions between PC and coordinating metal ions (Fung *et al.*, 1973; Reed and Scrutton, 1974), these experiments can only provide estimates of the distances between various protons and the paramagnetic metal atoms Mn(II) and Co(II). Therefore, in the absence of high resolution 3–dimensional structures, other methods of providing useful structural information must be found.

It has become widely accepted that large globular proteins are made up of a number of discrete structural (and sometimes functional) units called domains. This term was first used to describe the repeating units of sequence observed in immunoglobulin molecules (Hill *et al.*, 1966), but has since been applied almost universally. Rossman and Argos (1981) described a number of characteristics which define a domain:

(1) Similar domain structures or their amino acid sequences can be found either within the same polypeptide or in a different molecule.
(2) Domains within a polypeptide are spatially separated from each other, or at least form a compact unit or cluster of residues.
(3) Domains may have a specific function, such as binding a nucleotide or other ligands.
(4) The active centre of a molecule is at the interface between domains, which permits the simple function of each to be brought together to form a more sophisticated molecule.

A corollary to these definitions, is that delineation of the domain structure of a protein, may provide useful information regarding the mechanisms of action of that protein. The "spatial separation" of these domains implies that in many instances proteolytic cleavage under native conditions may be a useful method of domain identification, and this technique has been successfully applied by many workers in defining domain structures for globular proteins. Porter (1959) was able to show that IgG molecules consisted of two types of domains, Fab and Fc, using limited cleavage with papain. More recently, glutamate dehydrogenase (Haberland *et al*, 1980), myosin (Mornet and Ue, 1984) and citrate synthase (Bloxham *et al.*, 1980) have been investigated using this technique.

This last study is of particular interest and relevance, in that the X-ray diffraction structure for dimeric pig heart citrate synthase was subsequently determined (Remington *et al.*, 1982), providing an elegant example of the power of limited proteolysis as a tool for structural analysis. Treatment of pig citrate synthase with trypsin or subtilisin for limited periods in the presence of propionyl-CoA, resulted in a large N-terminal fragment, and a

small, less stable C-terminal fragment. Incubation in the presence of citrate or oxaloacetate abolished this proteolytic fragmentation. The site of cleavage was determined by sequence analysis of the fragments and when compared with the subsequent 3-dimensional structure, the large and small fragments can be seen clearly to represent distinct domains. Upon binding citrate or oxaloacetate, citrate synthase forms a tight, active conformation through rotation of the small domain relative to the large domain. This clearly explains the mechanism of protection against proteolysis by the above ligands, since in this conformation, the area susceptible to cleavage is no longer accessible to the protease.

A previous member of this laboratory, Dr. Y-S. Khew-Goodall (1985), demonstrated that limited treatment of chicken PC with chymotrypsin, resulted in the formation of a large, stable C-terminal fragment, of which she determined the N-terminal sequence. This fragmentation was shown to result in the loss of the overall activity of chicken PC with the rate of cleavage being lowered by the addition of ATP or acetyl-CoA. Oxaloacetate (2.0 mM), pyruvate (2.0 mM) and $MgCl_2$ (8.0 mM) were unable to alter the rate of loss of activity caused by chymotrypsin, while acetyl-CoA (0.2 mM) and ATP/$MgCl_2$ (2.0 mM) increased the $t_{0.5}$ of chicken PC activity by 15-20 fold and 5-6 fold respectively. Avidin was also observed to slow the rate of chicken PC digestion by chymotrypsin, although none of these treatments altered the resulting fragmentation pattern.

In this chapter, the combination of limited proteolysis with other techniques such as Western analysis and electron microscopy has been used to refine the domain model for PC with respect to the quaternary structure of the enzyme. Initially, the study was conducted on PC samples from mammalian sources (sheep and rat), however during the course of the work the complete sequence of PC from yeast was determined in this laboratory (Lim *et al.* 1988). This indicated that study of the structure of PC from yeast was likely to be the most beneficial in that the potential existed to test any hypotheses by site-directed mutagenesis. In addition, the sequence conservation reported in the previous chapter suggested that the overall structure of PCs from mammals and yeast were likely to be very similar.

The electron microscopy reported in this chapter was performed on my samples by Professor Frank Mayer, a visiting scientist from the University of Göttingen, FRG.

4.2 RESULTS

4.2.1 Proteolysis with Chymotrypsin

The effect of limited treatment of rat PC with chymotrypsin under non-denaturing conditions was tested as described in section 2.2.8. At various times, aliquots of the

reaction were removed and proteolysis terminated by the addition of PMSF to 10 mM. Subsamples of each of these time points were then subjected to either polyacrylamide gel electrophoresis, acetyl-CoA dependent activity assay or acetyl-CoA-independent activity assay. The resultant gel, after staining with Coomassie blue R_{250}, is shown in figure 4.1-a. Chymotrypsin treatment can be seen to cause degradation of the native band of apparent Mr 117,000, with concomitant production of two smaller bands, labelled rCHY-a and rCHY-b. Measuring the overall and acetyl-CoA independent activities by the incorporation of $^{14}CO_2$ into oxaloacetate, chymotrypsin was shown to cause loss of both these activities simultaneously (figure 4.1b). This result implies that chymotrypsin cleavage destroys the structure of the reaction centre of rat PC, as distinct from simply reducing its capacity to respond to the allosteric activator acetyl-CoA.

The data derived from densitometry measurements of the polyacrylamide gel from figure 4.1-a, are shown in figure 4.1-c, and clearly indicate the degradation of the monomers of PC to form two smaller species, the larger of which is more stable to protease. The native species is greatly depleted within 20 minutes at 30°C, with the concomitant appearance of the ~Mr 75,000 rCHY-a and ~Mr 50,000 rCHY-b fragments.

Comparing the data from figures 4.1-b and 4.1-c, the loss of activity approximately coincides with the loss of native PC monomers, implying that only intact PC subunits are able to participate in catalysis. However, in interpreting these experiments, it is important to remember that the PC tetramers undergo dissociation during dissolution in SDS-loading buffer for SDS-polyacrylamide gel electrophoresis, and that for this reason, the results cannot be extrapolated to provide information on the quaternary structure of the enzyme.

Incubation of proteolytically-treated PC samples with a molar excess of avidin prior to electrophoresis, causes retardation of biotin-containing bands during electrophoresis. Figure 4.2 clearly shows retardation of the parent rat PC and fragment rCHY-a in tracks where samples were preincubated with avidin. The transient species rCHY-b, however, was unaffected by preincubation with avidin, indicating that this fragment may represent the N-terminal portion of the subunit. Taken together with the coincidental production of the two fragments indicated in figure 4.1c, these results strongly suggests that each subunit of rat PC is made up of at least 2 distinct domains, a small N-terminal and a large C-terminal domain.

4.2.2 Proteolysis with Proteinase K

Based on the apparent stability of the large C-terminal domain to treatment with chymotrypsin, less selective proteases were screened in order to further delineate the domain structure.

When sheep PC was treated with proteinase K (a non-specific protease) distinct product bands could be observed after gel electrophoresis, as shown in figure 4.3. Initially two large fragments are produced, however the larger of the two (sPK-a) appears to be relatively unstable and is degraded within 20 minutes under these conditions. The smaller of the two species (sPK-b) appears to be quite stable, and accumulates with time. Even after 80 minutes, this fragment can be readily detected in sheep PC samples treated with proteinase K (data not shown). Avidin retardation experiments, represented by tracks 5(+) and 20(+) in figure 4.3, indicate that sPK-a (Mr 75,000) is associated with biotin, while sPK-b (Mr 69,000) is not. PC from all species so far investigated is neither phosphorylated nor glycosylated (Wallace and Easterbrook-Smith, 1985), therefore any alteration in observed molecular weight between two PC samples during electrophoresis, is most likely the result of different polypeptide chain lengths. This strongly suggests therefore, that sPK-b results from cleavage of approximately Mr 6,000 from the C-terminus of the less stable sPK-a.

Unlike the chymotrypsin treatment of rat PC shown above, proteinase K digestion of sheep PC did not result in a fragment representing the N-terminal portion of each subunit. This is consistent with the results observed when sheep PC is treated with chymotrypsin (Khew-Goodall, 1985), suggesting that this region of sheep PC is particularly susceptible to proteolysis and is rapidly degraded.

4.2.3 N-terminal Sequence Analysis of Proteolytic Products

In order to clearly define the identity of the proteolysis products, N-terminal sequence analysis using Edman degradation, was undertaken on samples transferred directly from polyacrylamide gels to derivatised-glass fibre sheet, by the method of Aebersold *et al.* (1986).

The sequences for fragments produced from sheep and rat PC by chymotrypsin, proteinase K and endoproteinase Arg-C, are shown in figures 4.4 and 4.5. The advantage of this method of sample preparation over others is demonstrated in figure 4.4-A, where the PTH values of each cycle of the rCHY-a sequencing experiment are shown. These sequences have been determined at the 5-20 pmole level, where as by more traditional methods of sample purification 50-100 pmoles have routinely been required for confident sequence determinations with this sequencer.

As can be seen from the sequence alignment (figure 4.5), fragments rCHY-a, sCHY-a, sPK-a and sPK-b all share common N-termini. This result clearly demonstrates that sPK-b is the result of cleavage from the C-terminus of sPK-a of a fragment of approximately Mr 6,000, as suggested in the previous section (4.2.2).

Figure 4.6 shows the position of the fragments with respect to the overall length of the polypeptide, as determined by comparison with the derived sequence data for yeast PC. The corresponding proteolysis products for yeast (Lim *et al.*, 1988) and chicken (Khew-Goodall, 1985) are shown, and are in remarkable agreement with the data from the mammalian PCs.

From these proteolysis studies, PC is proposed to consist of three domains, a large N-terminal domain, a larger internal domain and a small biotin carrying C-terminal domain.

4.2.4 Western Analysis of Monoclonal Antibody Epitopes

A number of monoclonal antibodies, raised against native sheep PC have been partially characterised in this laboratory (Carey, 1988), providing powerful probes of the structure of PC. The epitopes of five of these antibodies were defined within the proposed domain structure, by Western analysis of proteolysed sheep PC after electrophoresis. The results in figure 4.7 show that of the seven monoclonal antibodies tested, only three (#12, #42 & #60) bound strongly to denatured epitopes, while a further two (#6 & #18) bound weakly (as seen by binding to parent sheep PC). When proteolysed samples were probed, binding was detected to chymotrypsin treated enzyme in all five cases, but only the strongly binding antibodies (#12, #42 & #60) were observed to bind proteinase K treated enzyme. Antibodies #6, #12, #18, #42 & #60 bound to fragment sCHY-a, while #12, #42 & #60 bound to sPK-a, but only #60 bound to sPK-b. This localises the epitopes for #12 and #42 to the C-terminal Mr 6,000 of PC, #6 and #18 to the C-terminal Mr 75,000 of PC and #60 to an internal region of the subunit as shown in figure 4.8.

4.2.5 Electron Microscopic Analysis of Yeast PC Proteolysis

As the complete amino acid sequence of yeast PC had been determined, the fragmentation pattern caused by various proteases had been characterised (Lim *et al.*, 1988) and the protein was readily available, further studies into the structure of PC were conducted on samples from this source.

The aim of the experiment was to treat yeast PC with chymotrypsin or proteinase K, such that the fragmentation patterns observed upon gel electrophoresis were consistent with those found previously (Lim *et al.*, 1988; see figure 4.6). These samples could then be mounted and observed under the electron microscope. The data for the study was taken

from micrographs of two separate proteolysis experiments and figure 4.9 is typical of the gel electrophoresis results of both proteolysis experiments. That is, yeast PC was treated for 80 minutes with chymotrypsin as described in section 2.2.16, resulting mostly in production of a fragment representing the pyruvate and biotin-carrier domains, although a proportion of native species was also present. An additional fragment, observed in figure 4.9 to migrate at ~ Mr 40,000, did not contain biotin and although not well characterised is likely to represent the N-terminal portion (ATP-domain) of the molecule. Similar treatment with proteinase K resulted in two major species, one representing yeast PC truncated at the N-terminus by 17 amino acids, the other represented by an additional cleavage to remove the biotin-carrier domain. In this latter sample, a smaller proportion of native molecules was observed, compared to the chymotrypsin experiment.

Subsamples intended for electron microscopy, were cross-linked with Dithio-bis-(succinimidyl propionate) (DTSP) as described in section 2.2.16, prior to mounting on carbon support film. This procedure enhances the preservation of tetrameric particles (Rohde *et al.*, 1986), which dissociate upon dilution below 4 U/ml (Ashman *et al.*, 1972). The dimensions of the resultant particles were determined from prints at a final calibrated magnification of 250,000x. Tetrameric PC molecules mounted in this way are observed to have orthogonal axes of different length (Mayer *et al.*, 1980). The explanation for this is depicted in figure 4.10, where the collapse of the tetramer into itself results in a particle that has one short axis and one long axis. Also as indicated in the figure, the magnitude of the truncation upon proteolysis appears to differ between the long and short axes, such that the standard error of the mean is lessened in measurements of the long axis. The data set for this study, therefore, is based primarily on the dimensions of the long axis of tetrameric species.

Figure 4.11 represents example views of tetramers from untreated (a), chymotrypsin treated (b) and proteinase K treated (c) yeast PC. In both cases, protease treatment results in truncation of the outer tip of each subunit of the tetrameric molecule. In order to characterise this effect, the dimensions of a large number of particles were determined and categorised into size classes differing by 0.8 nm. When the number of particles in each class are graphed against the class dimensions (figure 4.12) it can be seen that chymotrypsin results mostly in truncation by 3-4 size classes and proteinase K by 4-5 size classes.

When the average dimension of truncated particles is calculated (ie. the average particle size excluding native particles - size class 9) it can be seen that chymotrypsin and proteinase K truncate the tetramer by approximately 2.7 nm (13%) and 3.0 nm (15%) respectively.

Table 4.1 shows a summary of the data, indicating the number of particles measured and the average particle dimensions.

4.3 DISCUSSION

4.3.1 Characterisation of the Domains of PC

The initial observation by Khew-Goodall (1985) that chicken PC could be readily fragmented under native conditions by a number of proteases, has led to a more detailed study of these cleavage products. The fragments produced by chymotrypsin treatment of rat PC, and chymotrypsin, proteinase K and endoproteinase arg-C treatment of sheep PC, have been isolated after gel electrophoresis and sequenced using Edman degradation. Taken together with the proteolysis data for yeast PC (Lim *et al.*, 1988) each monomer of PC is clearly composed of three globular masses, or domains, as evidenced by their separation upon treatment with various endoproteinases.

As shown in figure 4.6 these domains line up remarkably well with the linear assignment of function, derived from analysis of sequence conservation between related proteins or proteins with related function (chapter 3 - this thesis, Lim *et al.*, 1988). These domains, from the N-terminus, are thought to represent the ATP-binding (1^{st} partial reaction subsite), the pyruvate-binding (2^{nd} partial reaction subsite) and the biotin-carrier domains respectively.

A corollary of these assignments is that cleavage between the two domains (ATP and pyruvate) by chymotrypsin, would uncouple the two reaction subsites and therefore inactivate the enzyme. To test this hypothesis, the effect of chymotrypsin treatment on the overall and acetyl-CoA-independent activities of rat PC was measured (Figure 4.1-b). Loss of both these activities was observed, at essentially the same rate, indicating that the inactivation was not simply due to inhibition of the ability of PC to react to acetyl-CoA. Similar experiments conducted in yeast (Lim, 1988), confirm these results in that proteolysis causes the concurrent loss of the overall and acetyl-CoA-independent activities. Although other interpretations regarding the inhibitory effects of proteolysis are possible, given that the cleavage products might not necessarily dissociate until treated with denaturant, the view expressed here that cleavage separates the two reaction subsites is the simplest explanation (discussed further in the section 4.3.2).

The fact that inhibition of rat PC occurs while the two relatively stable fragments (rCHY-a and rCHY-b) can be detected in the solution, suggests that once cleavage occurs these two domains cannot interact in a catalytically competent manner. Although the N-termini of both these fragments are known, the positions of their C-termini are less certain. From avidin retardation experiments (figure 4.2) the large C-terminal fragment rCHY-a is

known to contain biotin, which is attached to a lysine residue 35 amino acids from the C-terminus of intact rat PC. Therefore, it is most likely that rCHY-a has an intact C-terminus and is 701 residues long. If rCHY-a and rCHY-b are separated by a single cleavage event, the C-terminus of rCHY-b would be represented by phe^{472} (all residue numbers refer to sequence of yeast PC, shown in figure 3.10), which would make rCHY-b approximately 430 residues long, based on the sequence comparison to yeast PC. The transient nature of rCHY-b, however, suggests that regions within this domain are susceptible to further proteolysis, such that the true extent of the C-terminus of rCHY-b is difficult to predict. Significant improvements in the technology involved in C-terminal sequencing would be required before the position of the C-terminus of rCHY-b could be adequately resolved. However, the potential for proteolysis to occur within the C-terminal region of rCHY-b may well explain the inability of rCHY-a and rCHY-b to interact catalytically, in that residues important for catalysis or inter-domain contact may have been lost.

Although the assignment of the two reaction subsites to enzymically separable domains is consistent with Lynen's hypothesis, more convincing evidence would be a demonstration that these domains can be combined in solution, as separate polypeptides, to result in catalytic activity. Such an experiment would require expression of the individual domains as separate gene products, followed by purification of the polypeptides to homogeneity. The advantages of such a system are obvious, in that interpretation of the results is not hampered by the inability to define the C-terminus of a polypeptide as described for rat PC above. Also, domains could be assayed individually or in pairs for the ability to bind substrates or catalyse partial reactions. The results presented here, that PC is comprised of three separable domains, compares well with some members of the biotin-dependent carboxylase family, for example TC and *E. coli* ACC, for which all three functions are encoded on three separate genes.

A comparison of the N-termini of the various proteolytic fragments (figure 4.5), shows that in PCs from rat, sheep, chicken and yeast, cleavage at the N-terminus of the pyruvate domain occurs at exactly analogous residues, corresponding to residue gln^{473} of yeast PC. The region of the protein joining the ATP and pyruvate domains must therefore, be quite accessible to the environment in that a number of protease cleavage sites are evident at or near this point. Endoproteinase arg-C, for instance, cleaves sheep PC some 37 residues further N-terminal to the site of chymotrypsin cleavage. Another conclusion to be drawn from these experiments is that the majority of the pyruvate domain is inaccessible to proteases.

Although secondary structure prediction programs can at best provide a measure of the probability of a given structure forming, it is interesting to compare the results of

such a prediction with experimental evidence. A secondary structure prediction, for the region of yeast PC between the ATP- and pyruvate-binding domains, was determined by the method of Garnier *et al.*, (1978) and is shown in figure 4.13. This prediction suggests that the known cleavage site of chymotrypsin (between residues phe^{472} and gln ^{473}) is likely to be within a region of random coil, with a high probability of being at the surface of the molecule. Although it is not known whether endo-proteinase arg-C cleavage of yeast PC occurs at the same site as has been observed in PC from sheep, the sequence conservation between the sequences of sheep and yeast PC around this region suggests that they may form similar structures. Again it is interesting that the secondary structure in this region of yeast PC (arg^{435}) is predicted to form random coil, with a high propensity towards being at the surface of the molecule.

These structure predictions are in accordance with the experimental findings, in that for cleavage to occur, these regions must be accessible to the protease and are therefore likely to be at or near the surface of PC. The fact that cleavage occurs mostly at one point to produce fragments with identical N-termini, even when treated with a non-specific protease (proteinase K), is consistent with the prediction that the sites of cleavage are relatively unstructured, but are flanked by more highly structured, less accessible regions.

It is also interesting to note that this inter-domain region has an unusually large percentage of proline residues, compared to the rest of the molecule. The approximately 140 amino acids between the designated ATP and pyruvate domains consist of 9% (rat) and 6.5% (yeast) proline compared with the rest of the molecule at 5% and 4% respectively. It is tempting therefore, given the very high sequence conservation between the PCs in this region, to speculate that the purpose of this inter-domain region is to allow the two domains (and therefore reaction subsites) to move relative to each other, facilitating formation of the reaction centre together with the biotin carrying domain.

A number of monoclonal antibodies, raised against native sheep PC, have been partially characterised (Carey, 1988). This analysis included measuring the effects of each antibody on the activity of PC, the ability of substrates to alter these effects and localisation of the various epitopes within the quaternary structure of sheep PC. Of these, 5 inhibitory (#6, #12, #18, #42 & #113) and 2 non-inhibitory (#60 & #80) antibodies were subjected to further analysis as a part of this study, in order to localise their respective epitopes within the overall length of the protein. As shown in figure 4.7, only three of these antibodies (#12, #42 & #60) bound strongly to denatured epitopes in a western analysis experiment, while two further antibodies (#6 & #18) bound weakly. Comparing the binding of these antibodies to the intact monomer as well as the chymotryptic and proteinase K fragments of sheep PC, the epitopes of #12 and #42 could be localised to the C-terminal Mr 6,000 of the

enzyme. The epitopes of antibodies #6, #18 and #60 could be localised to the C-terminal Mr 75,000 from these experiments.

4.3.2 Structural Arrangement of the Domains of PC

In the previous section, the effects of proteolysis on PCs from various sources were presented, providing evidence for the existence of three distinct domains. Due to the nature of the analysis, in that the proteolysis products were observed by denaturing gel electrophoresis, the conclusions were limited to defining the domains and their associated function. Information regarding how these domains combine to form the tertiary and quaternary structures must, therefore, be obtained by other means. Although the most precise method for analysing such structures is X-ray crystallography, this technique is not well suited to the study of PC, as detailed in the introduction to this chapter (section 4.1). Another powerful technique for observing a protein's overall structure is electron microscopy, which has been used extensively on PC in the past. Previous studies from this laboratory (Mayer *et al.*, 1980, Johannssen *et al.*, 1983, Rohde *et al.*, 1986) have shown that well preserved tetramers of PC can be visualised and their dimensions calculated using the electron microscope. Therefore, the questions asked as a part of this study were what effect do the proteases discussed above have on the ability of PC to remain as a tetramer, and also, what size particles result?

It has been observed previously that tetramers of PC mounted for electron microscopy undergo a flattening or collapsing process (Mayer *et al.*, 1980; Johannssen, *et al.*, 1983). The result of this collapse is shown in schematic form in figure 4.10, indicating that one axis of the tetramer appears to be longer than the other. As can also be seen from this diagram, the observed truncation at the tip of a molecule will be greater when measured along the long axis, compared to the short axis. In addition, the standard errors involved with measuring any truncation will be lessened along the long axis, due to the larger apparent magnitude of the truncation. The results, therefore, of the electron microscopy presented in figure 4.12 are derived solely from the dimensions of the long axis of tetramers.

As shown by figure 4.11, the immediate conclusion to be drawn from this experiment is that yeast PC treated with chymotrypsin or proteinase K under non-denaturing conditions is still able to form tetramers. In addition, these particles are truncated at their outer tip, compared to control samples. The regions cleaved from the tetramers appear to diffuse into the solution, as the chemical cross-linker (DTSP) does not cause cross-linking between the tetramers and the cleaved fragments.

A more detailed analysis, where the dimensions of a large number of tetrameric particles were determined is shown in figure 4.12. The particles were categorised into one of 11 size classes, each differing by 0.8 nm, on the basis of the dimensions of their long axes. When the number of particles in each size class are compared, a greater degree of heterogeneity can be seen in the samples treated with protease, compared to the control sample. Native sized PC particles are most commonly observed in size class nine, representing dimensions of 19.7 to 20.4 nm.

Chymotrypsin treatment causes truncation of tetramers by 3 - 4 size classes, and proteinase K results in truncation by 4 - 5 size cases. In the case of chymotrypsin this corresponds to an average truncation of 2.6 nm per tetramer, or 1.3 nm per subunit. For proteinase K, the average truncation is 3.0 nm per tetramer, or 1.5 nm per subunit. This represents truncation by 13% and 15% of the length of the molecule respectively.

Comparing the extent of truncation caused by these proteases with the known position of the cleavage sites within the polypeptide, two anomalies are evident. Chymotrypsin separates the ATP domain (which represents 40% of the mass of the protein) from the pyruvate and biotin-carrier domains, yet results in truncation by only 13 % of the length of the molecule. In addition, cleavage by proteinase K removes the first 17 amino acids from the N-terminus, along with the biotin-carrier domain (total of ~20 % of the mass), causing truncation of the tetramer by 15% of the length.

Table 4.1 shows the summary of the data from figure 4.12, and also includes the results calculated from the dimensions of the short axis. Comparing the data derived from the long and short axis shows an interesting trend. The observed magnitude of the truncation by the two proteases is greater along the long axis, yet the difference between the long and short axis is markedly larger in the proteinase K treated sample (7.6% for proteinase K, compared to 4.6% for chymotrypsin).

From these studies a domain model has been proposed, which accommodates the following experimental observations, and is shown schematically in figure 4.14.

(a) The ATP-binding, pyruvate-binding and biotin-carrier domains represent 40%, 42% and 18% of the mass of the protein respectively.

(b) Chymotrypsin treatment removes 40% of the protein, including the ATP-binding domain.

(c) This treatment does not interfere with tetramer formation, but results in truncation by 13% of the length of the tetramer.

(d) Proteinase K treatment removes the N-terminal 17 amino acids, and the C-terminal 20% of the protein including the biotin -carrier domain.

(e) This treatment does not interfere with tetramer formation, but results in truncation by 15% of the length of the tetramer.

(f) The short and long axes of truncated tetramers, resulting from collapsed particles, differ in the observed magnitude of truncation.

(g) This observed difference between the long and short axis is greater in the case of proteinase K treatment.

(h) The biotin prosthetic group has been localised to the central axis of each subunit, within 2-3 nm of the inter-subunit junction. (Johannssen, *et al.*, 1983)

(i) Both the ATP-binding (1st partial reaction subsite) and the pyruvate-binding (2nd partial reaction subsite) must be able to interact with the biotin moiety.

(j) Monoclonal antibodies #12 & #42 both bind to the biotin-carrier domain, and the epitopes have been localised to the central region of the tetramer (Carey, 1988).

(k) Together, these domains constitute a subunit of approximate dimensions, 9.5-10.0 nm x 7.5 nm x 3.5 nm (Mayer *et al.*, 1980; Rohde *et al.*, 1986).

This model shows the biotin-carrier domain at one end of the subunit, near the region of inter-subunit contact. The presence of tetramers in the proteolysis experiments where the biotin-carrier domain has been removed, shows that this region is not important for tetramer formation. However, contacts between adjacent biotin-carrier domains in the native tetramer may occur and may even be necessary for the formation or retension of active enzyme complexes.

Confirmation of these findings would be expected from experiments measuring the ability of truncated tetramers to form long chain-like complexes with the biotin-binding protein avidin (as described in section 1.4.2). These experiments proved to be practically difficult, however, due to the heterogeneity of the samples and the inability to reliably separate proteolysed tetramers from native tetramers. Although techniques for separating native from truncated tetramers could possibly be devised, the time scale of such experiments placed them beyond the scope of Professor Mayer's visit.

As can be seen from the schematic representation shown in figure 4.15, the proposed model can account for the dimensions of the particles observed in the electron microscope. Also evident is the difference between the observed length of the flattened molecules (long axis) compared to the intact particles (short axis). In the case of chymotrypsin treatment, removal of the ATP-binding domain can be seen to truncate the subunit length by 13% in the long axis, but the magnitude of this truncation appears less along the short axis. Similarly for the case of proteinase K treatment, where the long axis appears to be truncated by 15% of the length of the molecule, yet along the short axis, this truncation appears to be of the same order as that observed after chymotrypsin treatment.

Also shown in the figure is the further flattening of the molecule, with respect to chymotrypsin treated PC, that might be expected upon the loss of the biotin-carrier domain. This flattening explains how the truncation along the long axis increases with respect to chymotrypsin, yet remains virtually the same when measured along the short axis.

The schematic model presented also indicates that the regions important for intersubunit binding reside within the pyruvate-binding domain. The evidence for this comes from electron microscopic analysis of tetramers after chymotrypsin or proteinase K treatment. As indicated above, tetramers consisting of the C-terminal Mr 75,000 of each monomer can be observed after chymotrypsin treatment, indicating that the intersubunit junction is within this region.

Comparison of the epitope mapping results in figure 4.8 with those of Carey (1988) (shown in figure 4.16), provides further evidence consistent with the proposed structural model for PC. Antibodies #12 and #42 bind to a region near the centre of the tetramer, as would be expected of epitopes contained within the biotin-carrier domain. Binding of these antibodies can be excluded by preincubation of PC with avidin (Carey, 1988) confirming that antibodies #12 and #42 bind to epitopes within 1-2 nm of the biotin moiety in the native molecule. Localisation of the epitope of antibody #6 to the C-terminal Mr 75,000 is also consistent with the domain model, in that Carey (1988) proposed that this antibody binds near the inter subunit junction as shown in figure 4.16.

Antibodies #18 and #60 bind to a region approximately two-thirds the way along the subunit, consistent with the Mr 75,000 C-terminal fragment protruding towards the tip of each subunit. Binding of these antibodies is not influenced by avidin binding or co-incubation with polyclonal antibodies raised against biotin (Carey, 1988). These results are consistent with #18 and #60 binding to epitopes which are either more than 3-4 nm from the biotin moiety on the same face of the molecule, or on the convex face of the monomer.

Measuring the effects of substrates on the inhibition of sheep PC by these monoclonal antibodies, Carey (1988) found that oxaloacetate, at concentrations between 1-10 mM was able to significantly reduce the level of inhibition caused by antibody #12. Although the mechanisms by which these antibodies inhibit PC and the subsequent influence of oxaloacetate on such mechanisms are not clear, this observation is further evidence to support the structural model, in that the binding of oxaloacetate (at the pyruvate/oxaloacetate binding site) is able to influence the binding of an antibody directed to the biotin-carrier domain, implying that these two domains must be in close association.

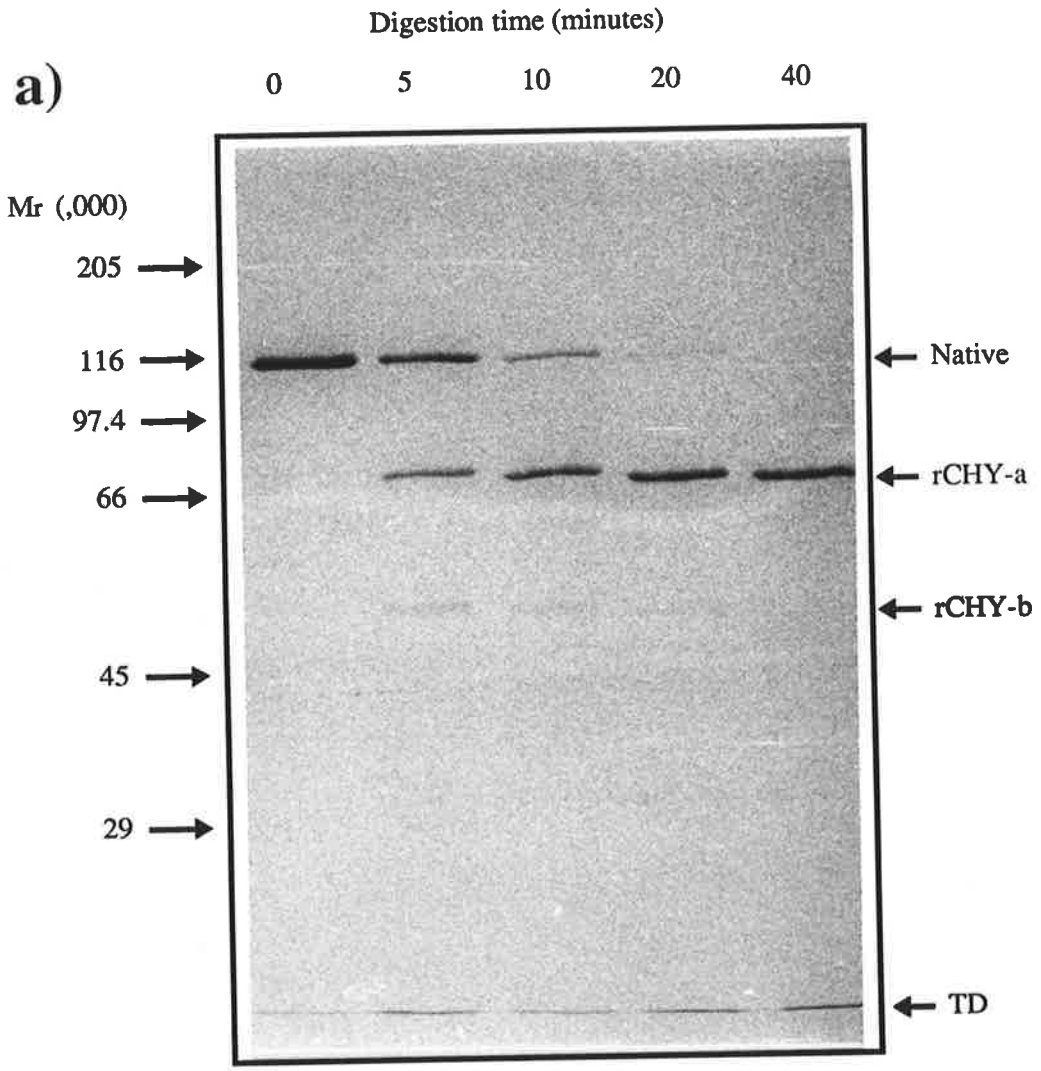
Figure 4.1 Limited Treatment of Rat Pyruvate Carboxylase with Chymotrypsin

Rat liver pyruvate carboxylase (14 U/mg), at a concentration of 1 mg/ml in 100 mM N-ethyl morpholine pH 8.0, was treated with chymotrypsin at a substrate to protease mass ratio of 150. After 0, 5, 10, 20 and 40 minutes at 30°C, subsamples were removed to a separate tube to which PMSF was added to a final concentration of 10 mM. This treatment terminated the proteolysis reaction.

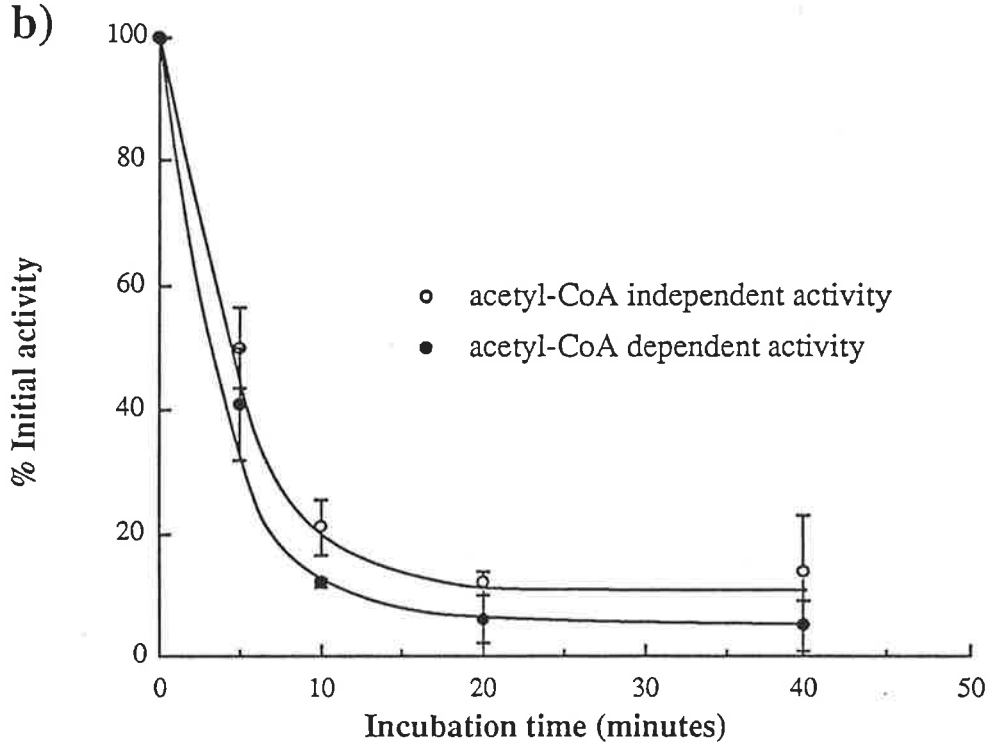
a) 5 μg of each subsample was fully reduced and denatured by boiling for 5 minutes in a loading buffer containing 0.1 % (w/v) SDS & 2.5% (v/v) β-mercaptoethanol. These samples were then electrophoresed through a 12.5% SDS-polyacrylamide gel and stained with Coomassie brilliant blue R_{250} as described in section 2.2.9. The extent of migration of the molecular weight markers is indicated along with the respective relative molecular mass (Mr). The position of the two major proteolytic fragments (rCHY-a & rCHY-b) are shown, as is the intact monomer (Native).

b) 3 μg of each subsample was subjected to the acetyl-CoA-dependent radiochemical assay of Keech and Farrant (1968) (section 2.2.2). In addition, 50 μg of each subsample was tested for acetyl-CoA-independent activity by the method of Ashman *et al.* (1972) (section 2.2.3). The figure shows the results for each time point expressed as a % of the initial activity.

c) The protein content of the 3 major bands observed after SDS-polyacrylamide gel electrophoresis (part a) was quantified by densitometry (100% defined as the protein content of the intact monomer band at time zero) and expressed as a function of time.


a)
Digestion time (minutes)
0
5
10
20
40
Mr (,000)
205
116
97.4
66
45
29
Native
rCHY-a
rCHY-b
TD

b)


100
80
60
40
20
0
% Initial activity
acetyl-CoA independent activity
acetyl-CoA dependent activity
0
10
20
30
40
50
Incubation time (minutes)

c)

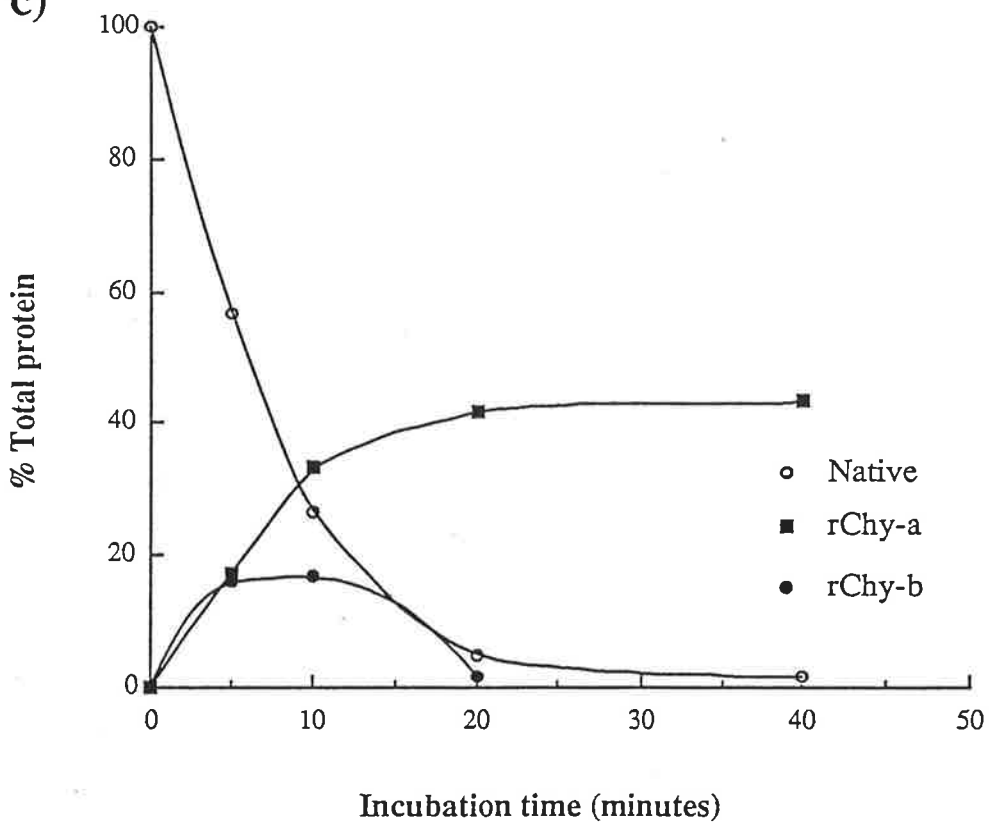

100
80
60
40
20
0
% Total protein
Native
rChy-a
rChy-b
0
10
20
30
40
50
Incubation time (minutes)

Figure 4.2 Effect of Avidin on the Products of Limited Treatment of Rat Pyruvate Carboxylase with Chymotrypsin

Rat liver pyruvate carboxylase (14 U/mg), at a concentration of 1 mg/ml in 100 mM N-ethyl morpholine pH 8.0, was treated with chymotrypsin at a substrate to protease mass ratio of 100 at 30°C. Subsamples were removed after 0, 5, 10 and 20 minutes and treated with PMSF (10 mM final concentration) to terminate proteolysis. 5 μg of each subsample was fully reduced and denatured by boiling for 5 minutes in gel-loading buffer containing 0.1 % (w/v) SDS & 2.5% (v/v) β-mercaptoethanol. These samples were then electrophoresed through a 12.5% SDS-polyacrylamide gel and stained with Coomassie brilliant blue R_{250} as described in section 2.2.9.

Samples indicated by (+) were treated with avidin (2 fold molar excess) for 10 minutes at room temperature prior to electrophoresis. The migration of proteins bands containing covalently bound biotin is significantly retarded by this treatment. The extent of migration of the molecular weight markers is indicated along with the respective relative molecular mass (Mr). The position of the two major proteolytic fragments (rCHY-a & rCHY-b) are shown, as is the intact monomer (Native).

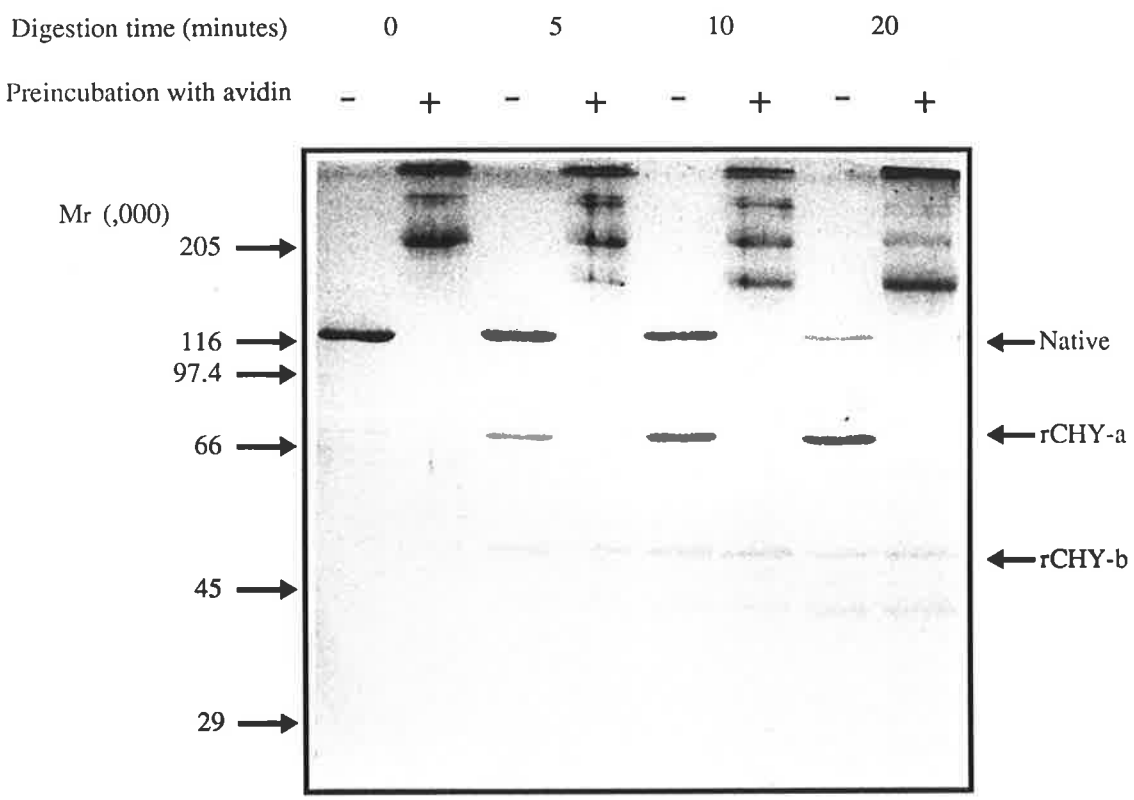

Digestion time (minutes)
0
5
10
20
Preincubation with avidin
- + - + - + - +
Mr (,000)
205
116
97.4
66
45
29
Native
rCHY-a
rCHY-b

Figure 4.3 Effect of Avidin on the Products of Limited Treatment of Sheep Pyruvate Carboxylase with Proteinase K

Sheep liver pyruvate carboxylase (20 U/mg), at a concentration of 1 mg/ml in 100 mM N-ethyl morpholine pH 8.0, was treated with proteinase K at a substrate to protease mass ratio of 50 at 30°C. Subsamples were removed after 0, 5, 10 and 20 minutes and treated with PMSF (10 mM final concentration) to terminate proteolysis. 5 μg of each subsample was fully reduced and denatured by boiling for 5 minutes in gel-loading buffer containing 0.1 % (w/v) SDS & 2.5% (v/v) β-mercaptoethanol. These samples were then electrophoresed through a 12.5% SDS-polyacrylamide gel and stained with Coomassie brilliant blue R_{250} as described in section 2.2.9.

Samples indicated by (+) were treated with avidin (2 fold molar excess) for 10 minutes at room temperature prior to electrophoresis. The migration of proteins bands containing covalently bound biotin is significantly retarded by this treatment. The extent of migration of the molecular weight markers is indicated along with the respective relative molecular mass (Mr). The position of the two major proteolytic fragments (sPK-a & sPK-b) are shown, as is the intact monomer (Native).

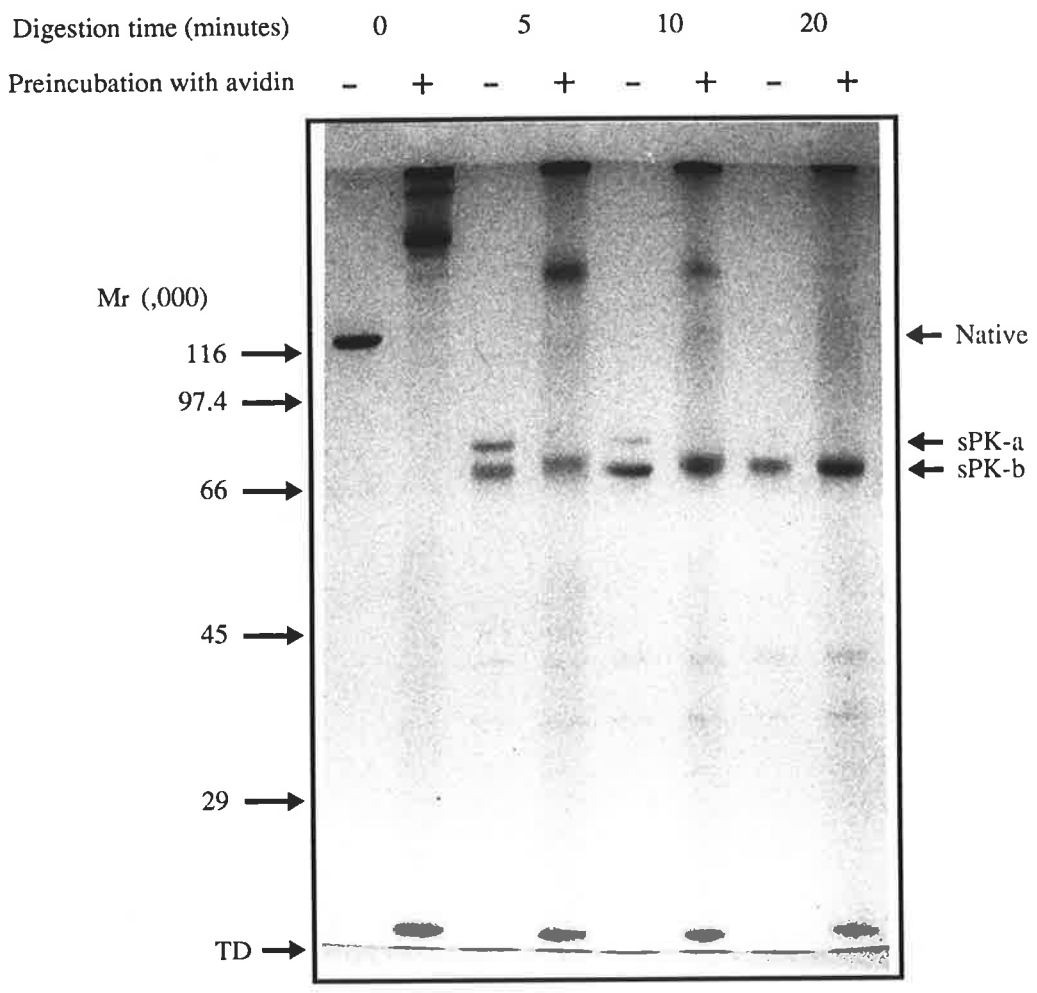

Digestion time (minutes)
0
5
10
20
Preincubation with avidin
- + - + - + - +
Mr (,000)
116
97.4
66
45
29
TD
Native
sPK-a
sPK-b

Figure 4.4 Yield of PTH-derivatives of Amino Acids from Samples Subjected to Automated Edman Degradation after Direct Transfer onto Derivatised Glass Fibre Sheet.

Proteolytic fragments of rat and sheep liver pyruvate carboxylase were isolated from polyacrylamide gels by electrotransfer to 3-amino-propyl-triethoxy silane-derivatised glass fibre sheet by the method of Aebersold *et al.* (1986) as described in section 2.2.10. The derivatised glass fibre sheets were then placed directly into the reaction chamber of an Applied Biosystems gas-phase automated peptide sequencer (Model 470). The resultant PTH-amino acids evolved during each cycle of Edman degradation were analysed by rp-HPLC and the yield of each cycle quantified by comparison to derivatised standards chromatographed under the same conditions.

Shown here, plotted as histograms, are the yields of the PTH-derivatised amino acids identified in each cycle of Edman degradation (expressed in picomoles) from the analysis of six peptides. The amino acids at each cycle are indicated using the one letter code.

A sCHY-a (~ 75,000 Mr)

B rCHY-a (~ 75,000 Mr)

C sARG-a (~ 78,000 Mr)

D sPK-a (~ 75,000 Mr)

E sPK-b (~ 69,000 Mr)

F rCHY-b (~ 50,000 Mr)

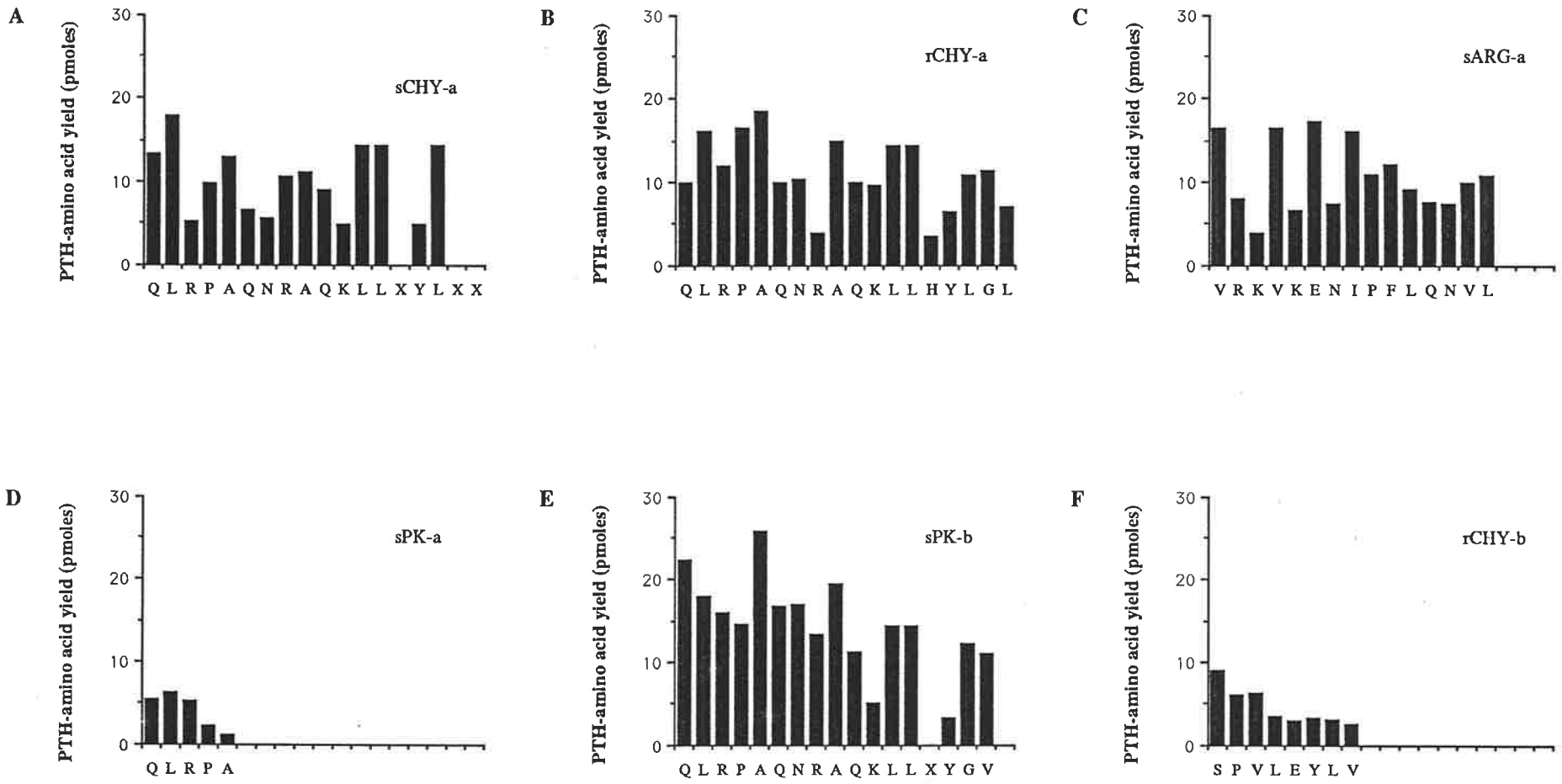

A
PTH-amino acid yield (pmoles)
30
20
10
0
sCHY-a
Q L R P A Q N R A Q K L L X Y L X X
B
PTH-amino acid yield (pmoles)
30
20
10
0
rCHY-a
Q L R P A Q N R A Q K L L H Y L G L
C
PTH-amino acid yield (pmoles)
30
20
10
0
sARG-a
V R K V K E N I P F L Q N V L
D
PTH-amino acid yield (pmoles)
30
20
10
0
sPK-a
Q L R P A
E
PTH-amino acid yield (pmoles)
30
20
10
0
sPK-b
Q L R P A Q N R A Q K L L X Y G V
F
PTH-amino acid yield (pmoles)
30
20
10
0
rCHY-b
S P V L E Y L V

Figure 4.5 Alignment of N-terminal Sequences of Fragments of Pyruvate Carboxylase Produced by Proteolysis.

Proteolytic fragments of rat and sheep liver pyruvate carboxylase were isolated from polyacrylamide gels by electrotransfer to 3-aminopropyl-triethoxy silane-derivatised glass fibre sheet by the method of Aebersold *et al.* (1986) as described in section 2.2.10. The derivatised glass fibre sheets were then placed directly into the reaction chamber of an Applied Biosystems gas-phase automated peptide sequencer (Model 470). The resultant PTH-amino acids evolved during each cycle of Edman degradation were analysed by rp-HPLC and compared to derivatised amino acid standards chromatographed under the same conditions.

Shown here, using the one letter code, are the N-terminal sequences of the various fragments of rat and sheep pyruvate carboxylase determined as described above. In addition, the corresponding fragments (indicated by *) isolated from chicken (Khew-Goodall, 1985) and yeast (Lim *et al.*, 1988) are shown for comparison. Regions of identical sequence are indicated by shading.

The alignment of fragments rCHY-b and sARG-a with the deduced amino acid sequence of yeast pyruvate carboxylase (Lim *et al.*, 1988) are also shown, with the numbers indicating the position of the first residue within the overall sequence of yeast pyruvate carboxylase.

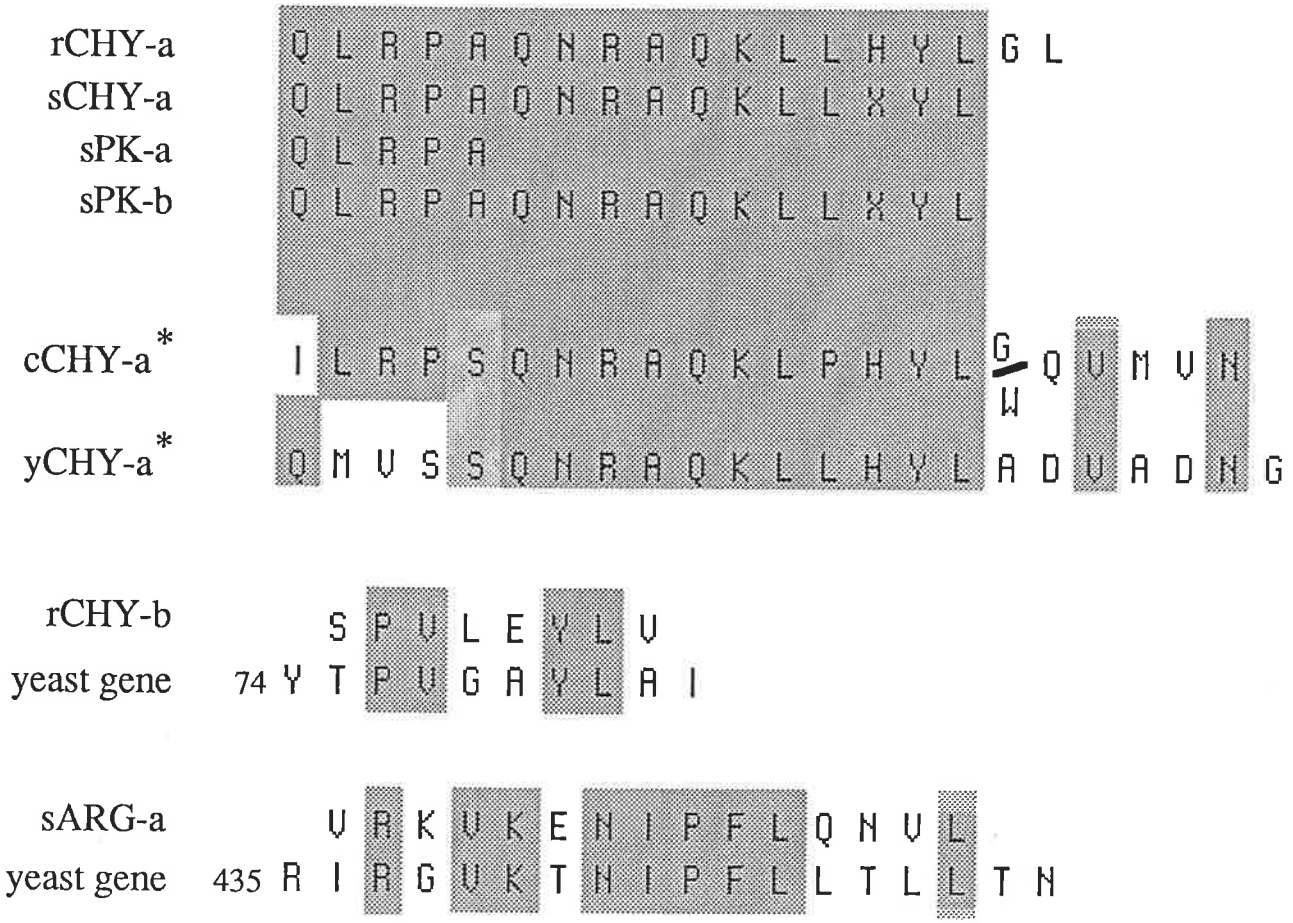

rCHY-a
sCHY-a
sPK-a
sPK-b
cCHY-a*
yCHY-a*
rCHY-b
yeast gene
74
sARG-a
yeast gene
435

Figure 4.6 Schematic Alignment of the Proteolytic Fragments of Pyruvate Carboxylase with Respect to the Complete Yeast Pyruvate Carboxylase Sequence

The N-terminal amino acid sequences of the various proteolytic fragments of sheep and rat pyruvate carboxylase indicated in figure 4.5 allow the fragments to be aligned with respect to the overall length of the complete sequence for yeast pyruvate carboxylase (Lim *et al.*, 1988). Also indicated in the figure are the three regions of assigned function in yeast pyruvate carboxylase, that of the ATP-binding, pyruvate-binding and biotin-carrier domains.

* The corresponding proteolytic fragments from chicken pyruvate carboxylase (Khew-Goodall, 1985) and yeast pyruvate carboxylase (Lim *et al.*, 1988) are shown for comparison.

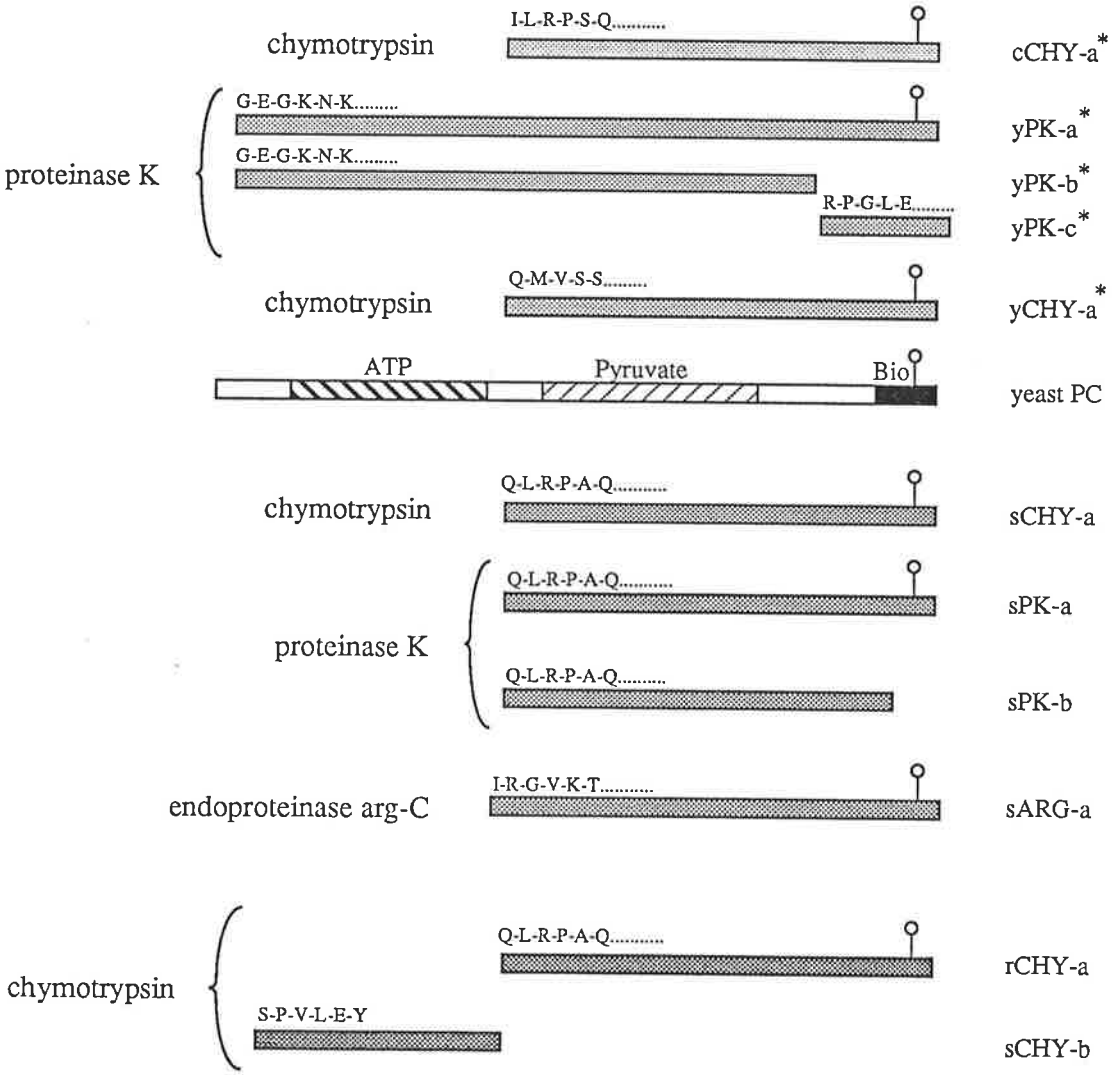

chymotrypsin
I-L-R-P-S-Q.....
cCHY-a*
proteinase K
G-E-G-K-N-K.....
yPK-a*
G-E-G-K-N-K.....
yPK-b*
R-P-G-L-E.....
yPK-c*
chymotrypsin
Q-M-V-S-S.....
yCHY-a*
ATP
Pyruvate
Bio
yeast PC
chymotrypsin
Q-L-R-P-A-Q.....
sCHY-a
proteinase K
Q-L-R-P-A-Q.....
sPK-a
Q-L-R-P-A-Q.....
sPK-b
endoproteinase arg-C
I-R-G-V-K-T.....
sARG-a
chymotrypsin
Q-L-R-P-A-Q.....
rCHY-a
S-P-V-L-E-Y
sCHY-b

Figure 4.7 Western Analysis of Protease-Treated Sheep Pyruvate Carboxylase Probed with Monoclonal Antibodies

Sheep liver pyruvate carboxylase (20 U/mg), at a concentration of 1 mg/ml in 100 mM N-ethyl morpholine pH 8.0, was treated with chymotrypsin (substrate to protease mass ratio of 100) or proteinase K (substrate to protease mass ratio of 50) at 30°C for 10 minutes. The reactions were terminated by the addition of PMSF to a final concentration of 10 mM. Samples of 18 μg of control and proteolysed pyruvate carboxylase were then electrophoresed through 12.5% SDS-polyacrylamide gels, in tracks 3 cm wide. Biotinylated molecular weight markers (N-hydroxysuccinimide-biotin-treated Sigma SDS-6H) were electrophoresed in adjacent tracks (1 cm wide) to provide a positive control for the experiment and to calibrate the gel.

The separated proteins were then transferred to nitrocellulose membrane and probed with various monoclonal antibodies followed by an anti-mouse antibody-alkaline phosphatase conjugate as described in section 2.2.11. The probing apparatus used in this experiment allowed a single filter to be probed with up to 45 individual antibody solutions at the same time without cross-contamination. The extent of migration of the biotinylated molecular weight markers are indicated along with the respective relative molecular mass (Mr). The position of the major proteolytic fragments (sCHY-a, sPK-a & sPK-b) are shown, as is the intact monomer (Native).

Track 1 - probed with antibody #6
Track 2 - probed with antibody #12
Track 3 -probed with antibody #18
Track 4 - probed with antibody #42
Track 5 - probed with antibody #60
Track 6 - probed with antibody #80
Track 7 - probed with antibody #113

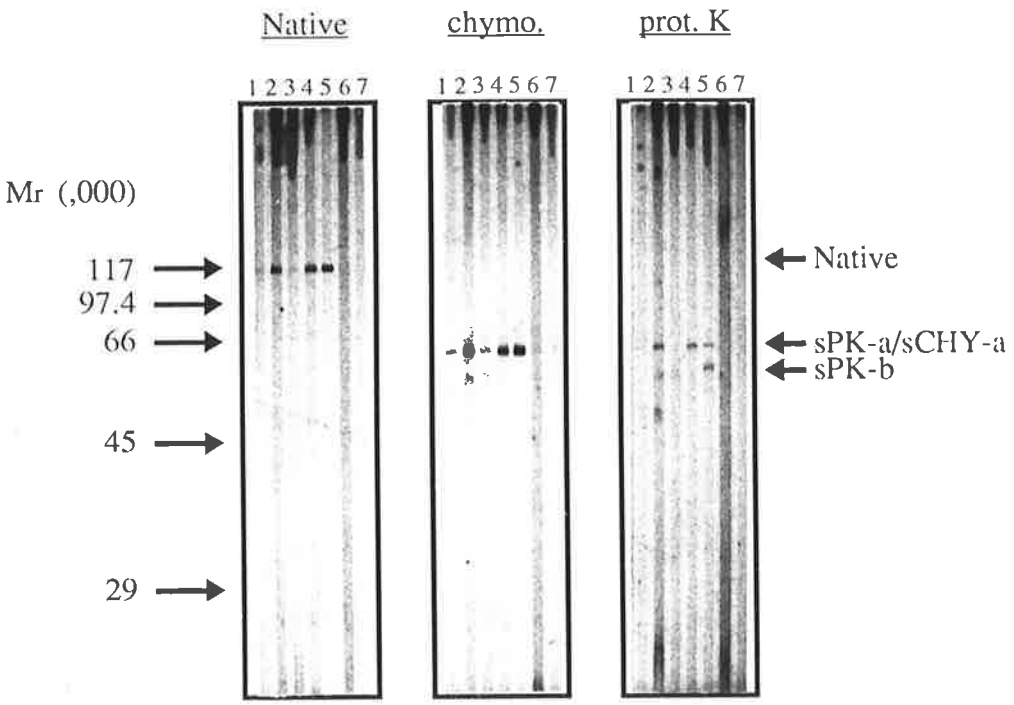

Native
chymo.
prot. K
1 2 3 4 5 6 7
1 2 3 4 5 6 7
1 2 3 4 5 6 7
Mr (,000)
117
97.4
66
45
29
Native
sPK-a/sCHY-a
sPK-b

Figure 4.8 Schematic Alignment with Respect to the Complete Yeast Pyruvate Carboxylase Sequence of the Regions Incorporating Monoclonal Antibody Epitopes

Shown here is a schematic representation of the alignment of three proteolytic fragments of sheep pyruvate carboxylase with respect to the complete sequence of yeast pyruvate carboxylase (Lim *et al.*, 1988). Also indicated in the figure are the regions of sheep pyruvate carboxylase to which five of the monoclonal antibodies bind as determined from Western analysis (figure 4.7)

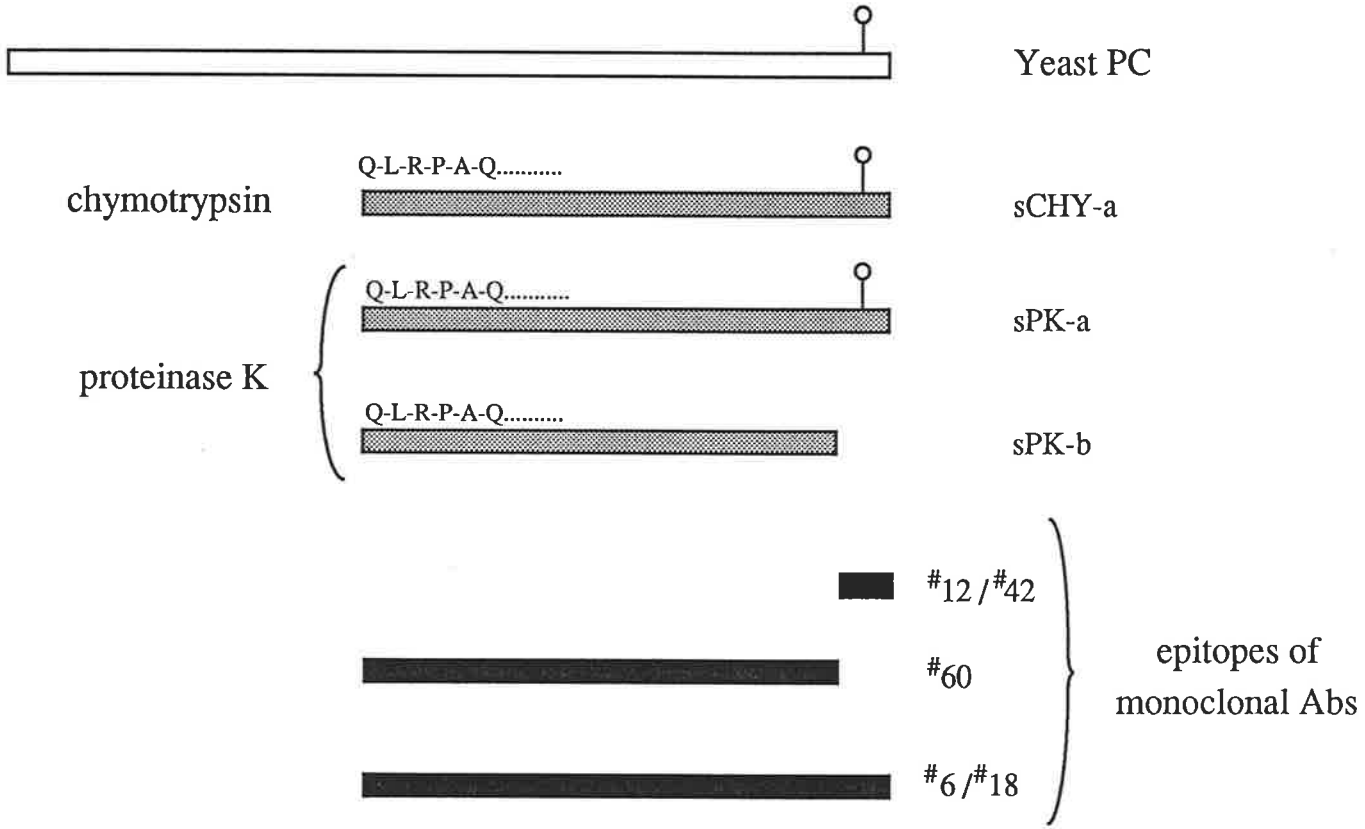

Yeast PC
chymotrypsin
Q-L-R-P-A-Q.....
sCHY-a
proteinase K
Q-L-R-P-A-Q.....
sPK-a
Q-L-R-P-A-Q.....
sPK-b
#12 /#42
#60
#6 /#18
epitopes of
monoclonal Abs

Figure 4.9 Limited Treatment of Yeast Pyruvate Carboxylase with Chymotrypsin and Proteinase K

Yeast pyruvate carboxylase (specific activity, 21 U/mg), at a concentration of 1 mg/ml in 10 mM $MgCl_2$ / 50 mM Tris-Cl pH 7.2, was treated with chymotrypsin or proteinase K at a substrate to protease mass ratio of 12.5 at 30°C for 80 minutes. Proteolysis was terminated by addition of PMSF to a final concentration of 10 mM. Subsamples (5 μg) of reaction and untreated control were fully reduced and denatured by boiling for 5 minutes in gel-loading buffer containing 0.1 % (w/v) SDS & 2.5% (v/v) β-mercaptoethanol. These samples were then electrophoresed through a 12.5% SDS-polyacrylamide gel and stained with Coomassie Brilliant Blue R_{250} as described in section 2.2.9. The extent of migration of the molecular weight markers are indicated along with the respective relative molecular mass (Mr).

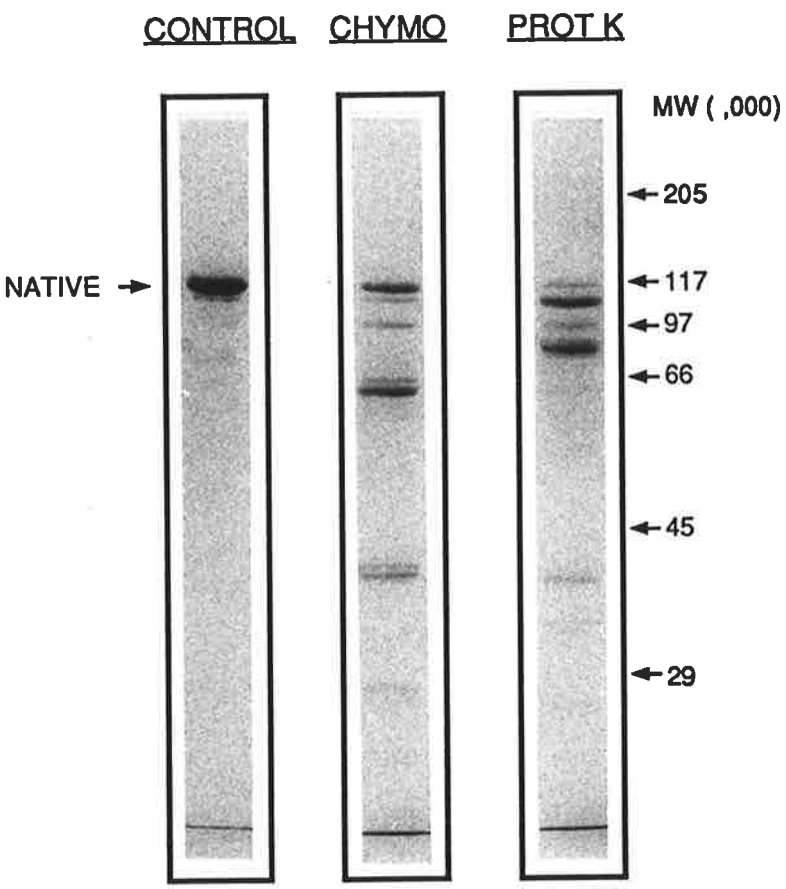

CONTROL
CHYMO
PROT K
MW (,000)
NATIVE
205
117
97
66
45
29

Figure 4.10 Schematic Representation of the Flattening of Tetramers of Pyruvate Carboxylase Observed After Mounting for Electron Microscopy

Shown here in schematic form is the rationale behind the observations of Mayer *et al.* (1980), Johannssen *et al.* (1983) & Rohde *et al.* (1986) that tetramers viewed from above appeared to have orthogonal axes of different length. This is considered to be an artifact caused by a flattening process during mounting of samples of pyruvate carboxylase for electron microscopy (see figure 1.5). Panel A represents the side view of tetrameric pyruvate carboxylase and panel B represents the corresponding particles viewed from above.

Sample I represents native tetrameric pyruvate carboxylase.

Sample II represents a native tetrameric particle where one face of the tetramer has collapsed onto the orthogonal face.

Sample III represents the case where tetrameric pyruvate carboxylase has undergone a further process of flattening after the initial collapse.

Sample IV indicates the view when the outer tip of each subunit is truncated. Also indicated in this field is the apparent difference between the dimensions of such a truncation when measured along the long (D_1) and short (D_2) axes.

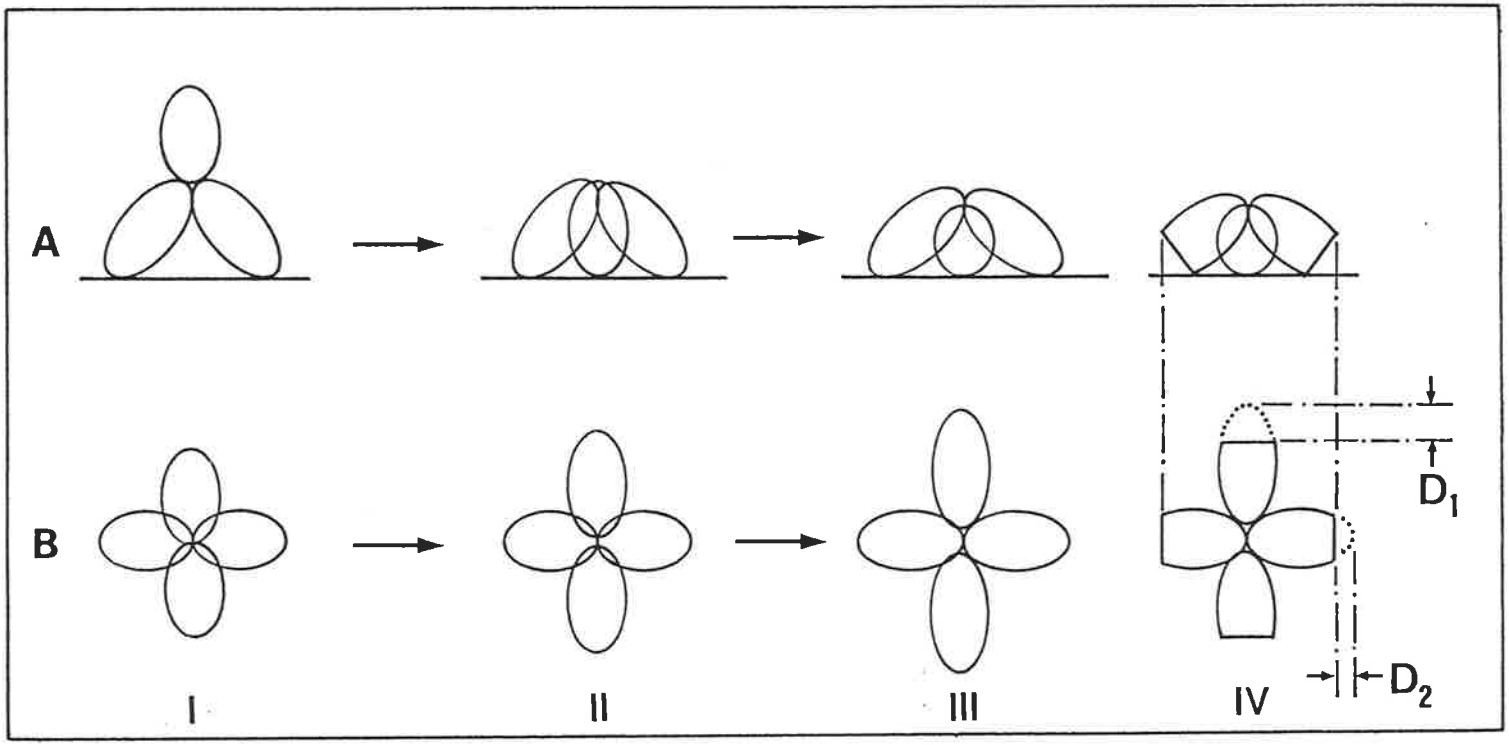

A
B
I
II
III
IV
D_1
D_2

Figure 4.11 Electron Micrographs of Native and Truncated Tetramers of Pyruvate Carboxylase

Yeast pyruvate carboxylase (specific activity, 21 U/mg), at a concentration of 1 mg/ml in 10 mM $MgCl_2$ / 50 mM Tris-Cl pH 7.2, was treated with chymotrypsin or proteinase K at a substrate to protease mass ratio of 12.5 at 30°C for 80 minutes. Proteolysis was terminated by addition of PMSF to a final concentration of 10 mM. Samples of proteolysed and control pyruvate carboxylase were treated with a cross-linking reagent (1mM DTSP) for 10 minutes at 30°C, after which the reaction was quenched by the addition of L-lysine (10 mM). This treatment enhances the preservation of any tetrameric particles during the mounting procedure for electron microscopy (Mayer *et al.*, 1980). The samples were mounted and stained with uranyl acetate as described in section 2.2.16.

Shown here are examples views taken from micrographs of samples treated as follows:

a) Control

b) Chymotrypsin-treated

c) Proteinase K-treated

The scale of the micrographs is indicated by the solid bar, which represents 15 nm. Also shown under the micrographs are corresponding line diagrams, with equivalent centres, to indicate the extents of the particles. The dashed line represents the outer edge of the native particle.

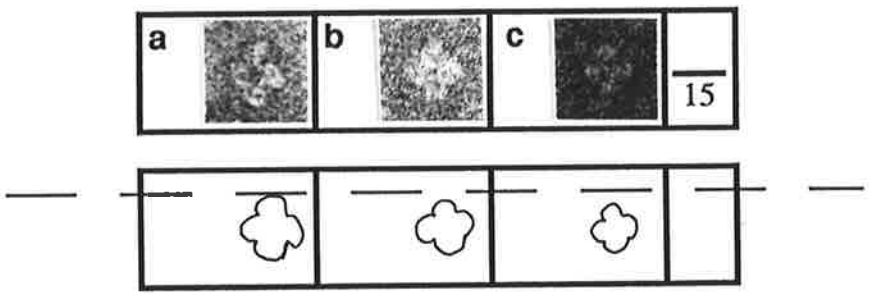
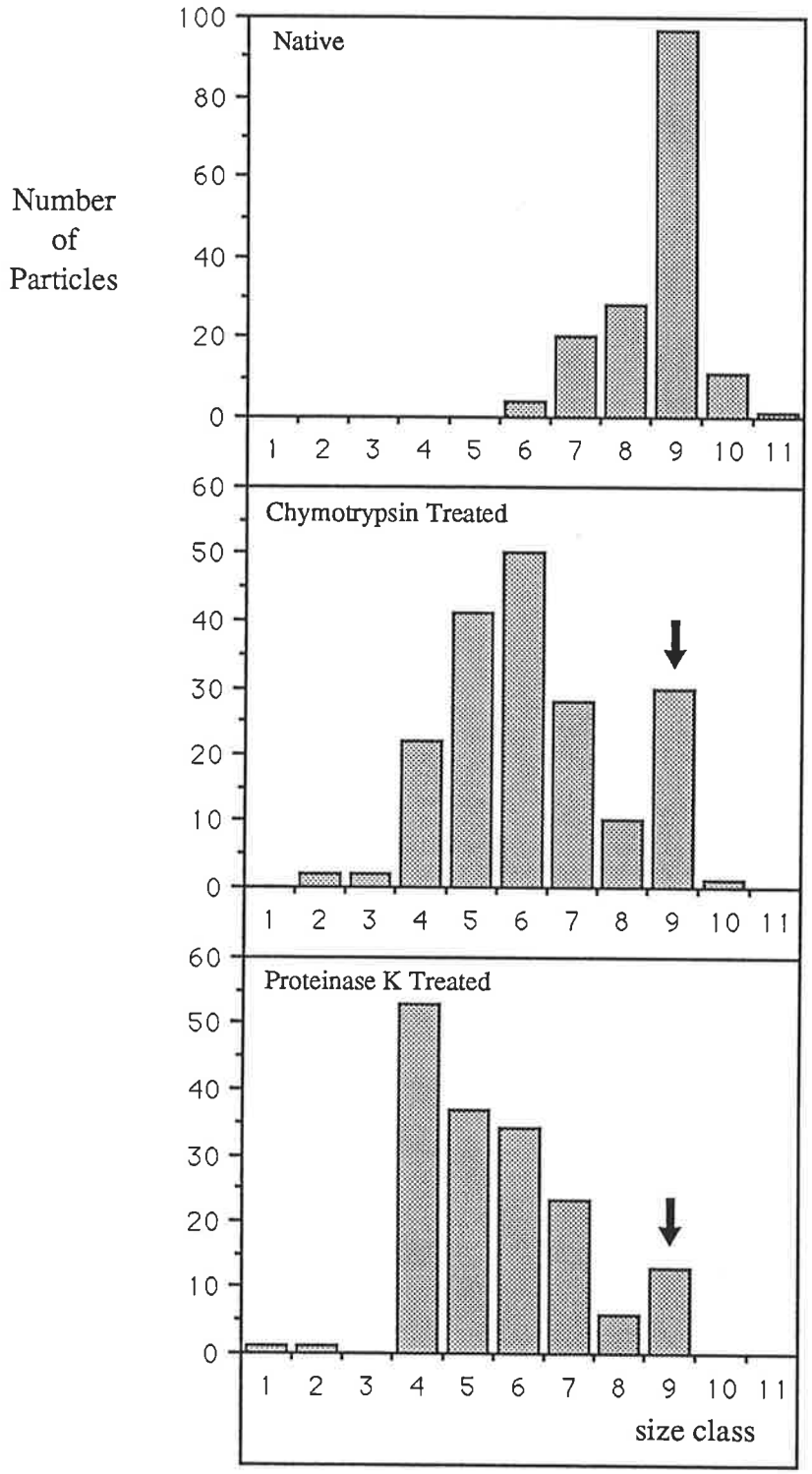

a
b
c
15

Figure 4.12 The Effect of Proteolysis on Tetramers of Pyruvate Carboxylase as Determined by Electron Microscopy

Yeast pyruvate carboxylase (specific activity, 21 U/mg), at a concentration of 1 mg/ml in 10 mM $MgCl_2$ / 50 mM Tris-Cl pH 7.2, was treated with chymotrypsin or proteinase K at a substrate to protease mass ratio of 12.5 at 30°C for 80 minutes. Proteolysis was terminated by addition of PMSF to a final concentration of 10 mM. Samples of proteolysed and control pyruvate carboxylase were treated with a cross-linking reagent (1mM DTSP) for 10 minutes at 30°C, after which the reaction was quenched by the addition of L-lysine (10 mM). This treatment enhances the preservation of any tetrameric particles during the mounting procedure for electron microscopy (Mayer *et al.*, 1980). The samples were mounted and stained with uranyl acetate as described in section 2.2.16.

The dimensions of the long axis of a large number of tetrameric particles were measured from electron micrographs of control and protease-treated samples. These particle were then categorised into size classes differing by 0.8 nm. The number of particles observed within each size class are presented as a histogram. The vertical arrow indicates the size class in which the majority of native sized particle are found.

Size class	Range of long axis (nm)
1	13.3 – 14.0
2	14.1 – 14.8
3	14.9 – 15.6
4	15.7– 16.4
5	16.5 – 17.2
6	17.3 – 18.0
7	18.1 – 18.8
8	18.9 – 19.6
9	19.7 – 20.4
10	20.5 – 21.2
11	21.3 – 22.0


Number of Particles
Native
Chymotrypsin Treated
Proteinase K Treated
size class
1 2 3 4 5 6 7 8 9 10 11
0 20 40 60 80 100
0 10 20 30 40 50 60

Treatment	No. of Particles	Average remaining length (% of Control)	
		short axis	long axis
Control	163	100	100
chymotrypsin	177	90.1	95.7
proteinase K	144	86.9	94.5

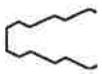
Table 4.1 Summary of the Effect of Protease Treatment on Tetramer Dimensions as Determined by Electron Microscopy The number of tetrameric particles observed in electron micrographs of control and protease-treated pyruvate carboxylase are shown, along with the average calculated dimensions of each axis of the tetramers expressed as a percentage of the control.

Figure 4.13 The Computer Predicted Secondary Structure of the Protease Susceptible Region of Yeast Pyruvate Carboxylase

The amino acid sequence of the region of yeast pyruvate carboxylase between the designated ATP-binding and pyruvate-binding domains (residues 371 and 490) was subjected to the secondary structure prediction program **Peptidestructure** (Devereux *et al.*, 1984) using the method of Garnier *et al.* (1978). In addition, this program calculates the probability that each residue will be exposed to the solvent. Shown here in schematic form is the graphical output of the prediction using the related program **Plotstructure**. The site of proteolytic cleavage by chymotrypsin is indicated by the arrow.


represents regions predicted to form α-helix

represents regions predicted to form β-pleated sheet


represents regions predicted to form β-turns

represents regions predicted to form random coil


represents residues with a surface probability of ≥ 0.7

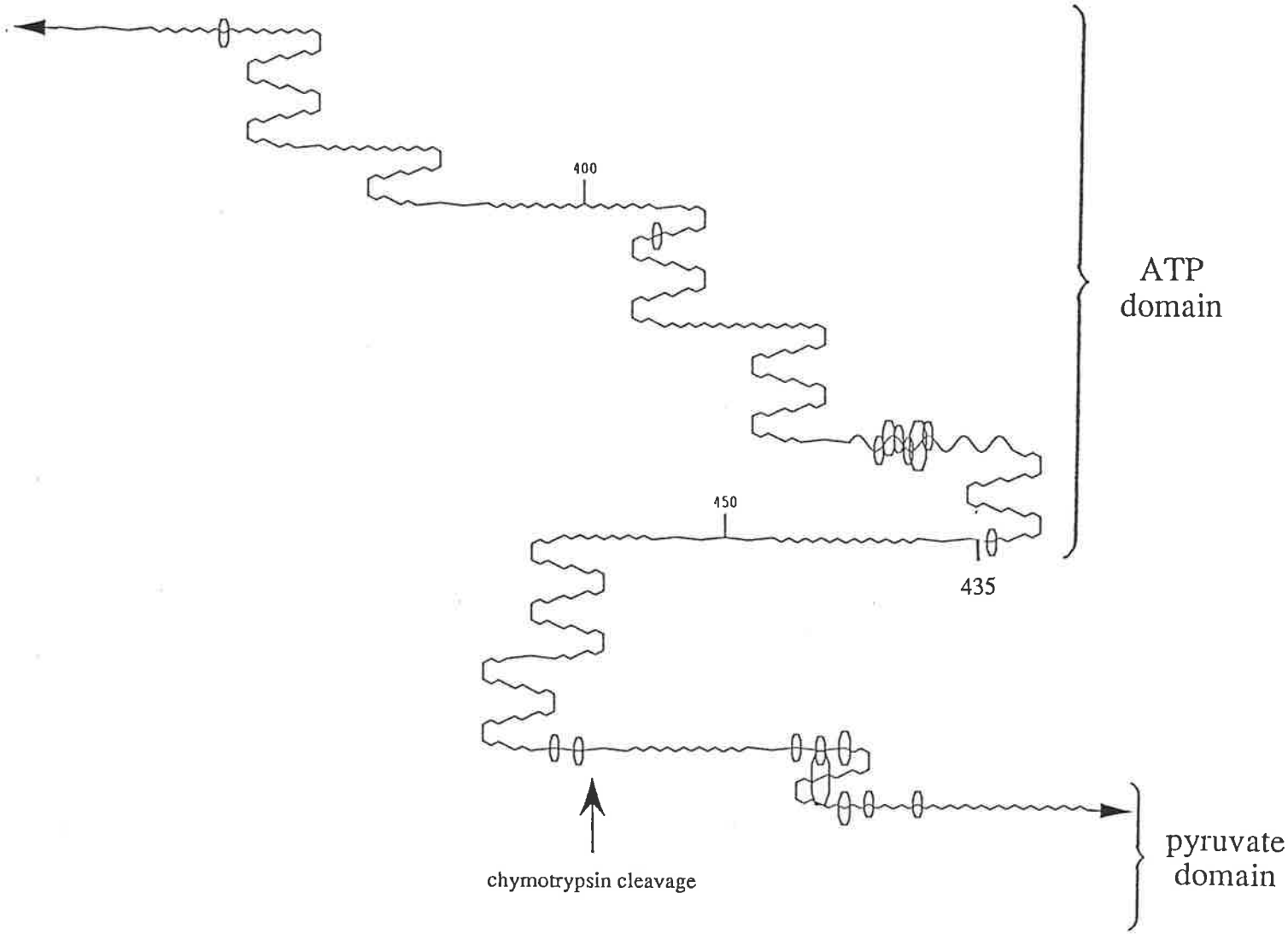

400
450
435
ATP
domain
chymotrypsin cleavage
pyruvate
domain

Figure 4.14 A Schematic Representation of a Model Consistent with the Structural Information for Pyruvate Carboxylase

Shown here in schematic form is a plausible structural model for pyruvate carboxylase which is consistent with the data presented. The arrangement of the ATP-binding (ATP), pyruvate-binding (PYR) and biotin-carrier (BIO) domains within one subunit of the tetramer are highlighted. The position of the biotin moiety and hence the reaction centre is indicated by the solid-filled circle.

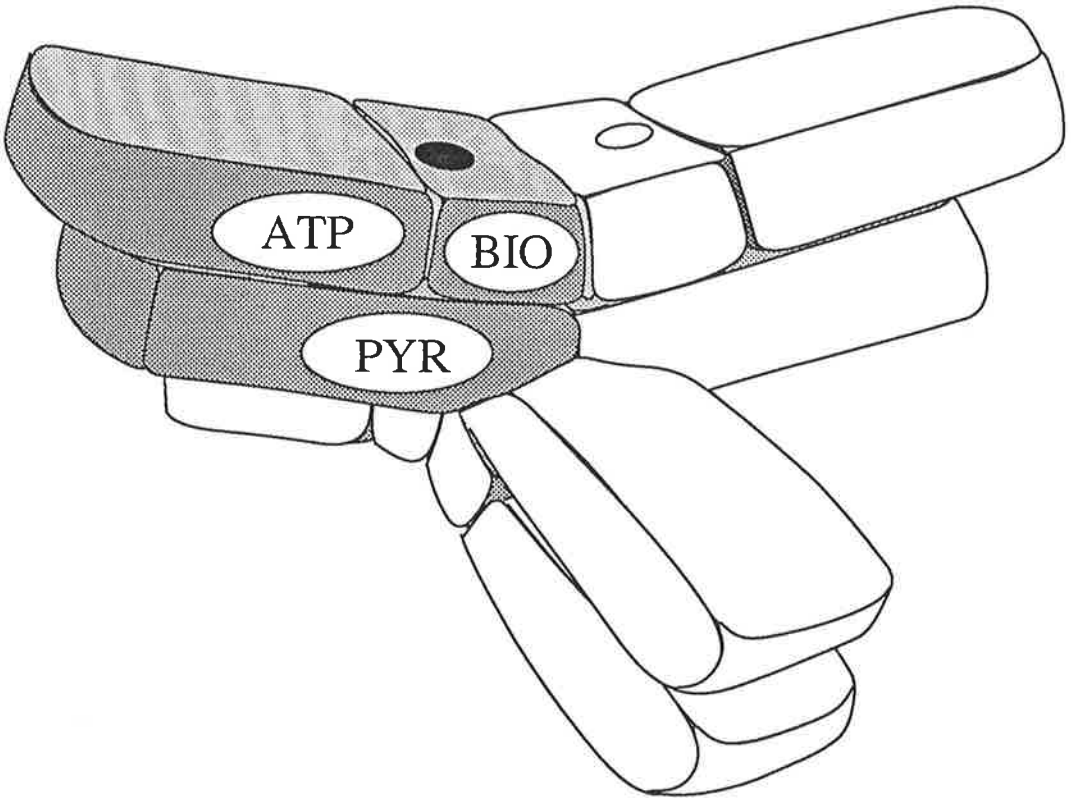

ATP
BIO
PYR

Figure 4.15 A Schematic Explanation for the Observed Differences Between Chymotrypsin and Proteinase K Treated Pyruvate Carboxylase

Shown here in schematic form is an explanation for the different magnitudes of the truncation observed along the two alternate axes of the tetramers (indicated in table 4.1). Also indicated is the reasoning for the magnitudes of the respective truncations (chymotrypsin versus proteinase K).

The dashed vertical lines indicate the observed dimensions of the full length and truncated particles, the magnitude of which is indicated by the solid horizontal lines in each case. The three domains BIO (biotin-carrier domain), PYR (pyruvate-binding domain) and ATP (ATP-binding domain) are indicated in side-on views of the pyruvate carboxylase structural model as proposed in figure 4.14. The support grid for the mounting of samples is also indicated (EM grid). The biotin moiety is represented by the black circle attached to the biotin-carrier domain by a flexible linker. In the case of the long axis view, only one subunit of the tetramer is shown.

The native example represents full length tetramers and the observed dimensions (along each axis) are designated 100% as indicated. The difference between the long and short axes results from flattening of the tetramers as indicated in figures 1.5 & 4.10.

The chymotrypsin and proteinase K examples represent views of the particles after truncation with the appropriate protease. The observed magnitude of the truncations are indicated in each case. Note that in the case of proteinase K treatment the tetramers have undergone further flattening with respect to the native or chymotrypsin treated samples.

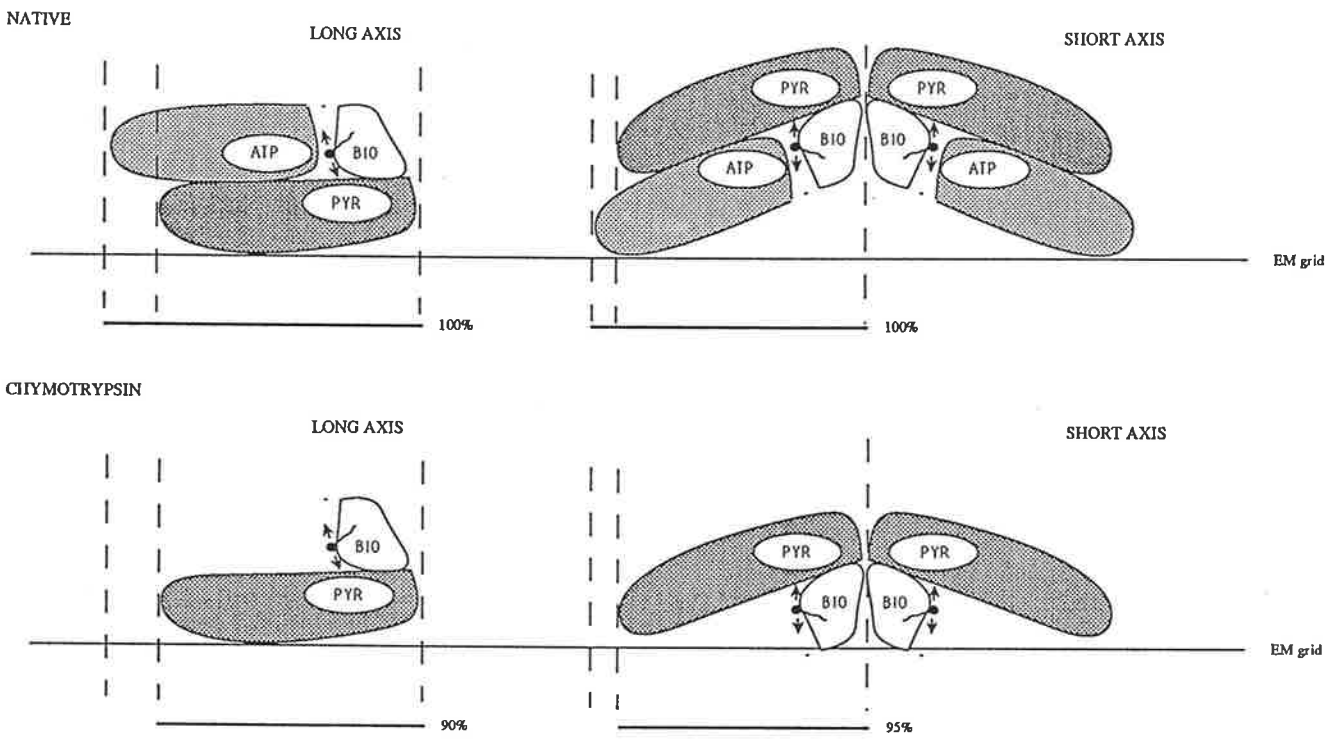
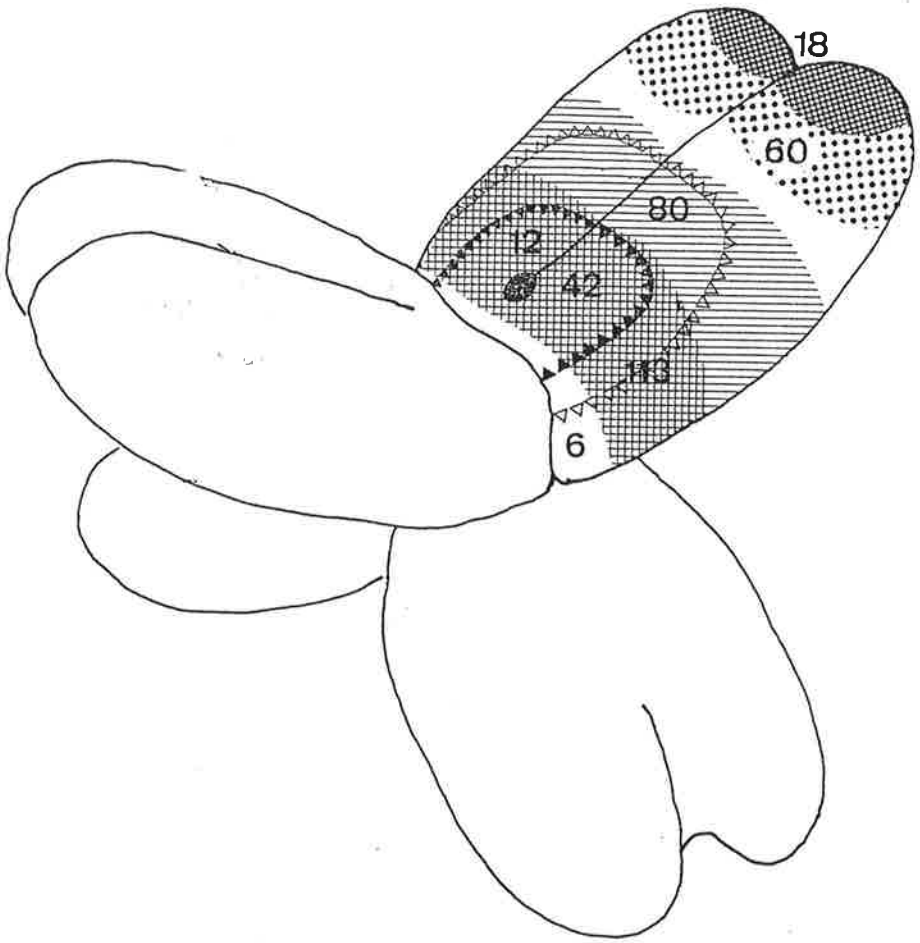

NATIVE
LONG AXIS
SHORT AXIS
ATP
BIO
PYR
PYR
PYR
BIO
BIO
ATP
ATP
EM grid
100%
100%
CHYMOTRYPSIN
LONG AXIS
SHORT AXIS
BIO
PYR
PYR
PYR
BIO
BIO
EM grid
90%
95%
PROTEINASE K
LONG AXIS
SHORT AXIS
ATP
PYR
PYR
PYR
ATP
ATP
EM grid
87%
95%

Figure 4.16 A Schematic Summary of the Position of Monoclonal Antibody Epitopes Determined from Electron Micrographs of Pyruvate Carboxylase-Antibody Complexes

Shown here as shaded regions are the proposed locations (on one subunit) of the epitopes of 7 monoclonal antibodies (#6, #12, #18, #42, #60, #80 & #113) as determined by electron microscopy of pyruvate carboxylase-IgG complexes (Carey, 1988). The reaction site is shown by the large solid-filled circle, which represents the biotin moiety. Based on the relative dimensions of avidin and IgG, the solid triangles represent the extent of exclusion by bound avidin and the open triangle represent the exclusion limit of a bound anti-biotin IgG molecule.


18
60
80
12
42
43
6

CHAPTER 5

THE ACETYL-CoA BINDING SITE

5.1 INTRODUCTION

The mechanism of enzymic catalysis and its regulation are common themes of investigation in biologically significant proteins. While the detailed structure of an enzyme alone may in some instances reveal the mechanistic pathway, more often, such conclusions are drawn from a wide body of evidence, provided by experiments on functional aspects of the molecule, in addition to the study of its structure. This chapter, therefore deals with one of the functional aspects of PC, its activation by acetyl-CoA.

PC from most species is allosterically activated by acetyl-CoA, and in the absence of a detailed 3-dimensional structure, study of the mechanism of its influence has revolved around enzyme kinetics, (including activation by analogues of acetyl-CoA), affinity labelling and chemical modification. Although it is known that acetyl-CoA activates the carboxylation of biotin and enhances the preservation of the tetrameric state, the site of binding of acetyl-CoA has yet to be determined.

In general, affinity labels rely on the presence of a reactive residue in the ligand binding site or reaction centre, in order for covalent modification to occur. The acetyl-CoA binding site of PC, however, is an allosteric site and may contain insufficiently reactive residues to enable a traditional affinity label to work. However, on the premise that appropriate residues are within the acetyl-CoA binding site, Khew-Goodall (1985) attempted to modify the site with a number of coenzyme-A derivatives.

Oxidised coenzyme-A, which reacts with a sulphydryl group in a disulphide exchange, had been shown to inhibit chicken PC (Fung, 1972) and was therefore tested as a potential affinity label. This inactivation was protectable by 0.25 mM acetyl-CoA, but not by 5.0 mM oxaloacetate. To obtain an acceptable rate of reaction for this disulphide exchange, a pH above neutral was required, which in turn caused inactivation of the enzyme itself. Although Mg^{2+} may be used in the solution to protect the enzyme against pH-dependent inactivation, this treatment lowered the rate of inactivation by oxidised Coenzyme-A, reducing its effectiveness as an affinity label.

S-(2-formylethyl)-CoA was synthesised and shown to inactivate chicken PC in a rate saturating manner, although at low concentrations of the inhibitor, the reaction kinetics became difficult to interpret. The inactivation process requires formation of a Schiff base between a lysine residue and the formaldehyde group of the ligand and subsequent reduction with $NaBH_4$ to generate a covalant link between the enzyme and the affinity label. Acetyl-CoA (0.25 mM) was able to protect chicken PC from inactivation by this treatment,

suggesting that modification was occuring at the acetyl-CoA binding site. For this reaction to proceed satisfactorally, again a pH above neutral was required, which in turn caused inactivation of the enzyme. The usefulness of S-(2-formylethyl)-CoA as an affinity label was lessened by difficulties involved with synthesising the ligand from acrolein and coenzyme-A, and the unusual reaction kinetics observed at low concentrations of the inhibitor.

A photo-affinity analogue of acetyl-CoA, p-azidophenacyl-CoA, was also characterised by Khew-Goodall (1985) on the basis that the reaction should be less dependent on pH, or the presence of reactive species in the binding site. The reaction was initiated by exposure to ultraviolet light, creating a highly reactive nitrene moiety at the phenacyl end of the ligand, which could potentially form a covalent bond with any of the residues within close proximity to the binding site. Inactivation by p-azidophenacyl-CoA was shown to be saturable with respect to inhibitor concentration, protectable by acetyl-CoA (0.25 mM) and the stoichiometry of labelling could be extrapolated to one mole of phenacyl-CoA analogue bound per mole of biotin when residual activity reached zero. This ligand therefore satisfies all three criteria of an affinity label as described by Singer (1967). Using gel filtration, Khew-Goodall was able to show that after photolysis, p-azidophenacyl-[^{3}H]CoA could be detected, associated with the large C-terminal chymotryptic fragment of chicken PC. Although this result does not allow identification of the acetyl-CoA binding site, it does suggest that at least the acyl-binding portion may be contained within the C-terminal Mr 75,000 of the enzyme.

Although p-azidophenacyl-CoA appears to be an effective photoaffinity label for the acetyl-CoA binding site, in practice its use may be limited due to the potential complexity of the labelling pattern. The difficulty in interpreting photoaffinity labelling experiments is well demonstrated by studies using S-(p-azidophenacyl)thiocarnitine as a photoaffinity label for the carnitine/acyl carnitine binding site of carnitine acetyltransferase [EC 2.3.1.7] (Mauro *et al.*, 1986). These authors showed that at least six labelled tryptic peptides could be identified after exhaustive trypsin treatment of the photolabelled enzyme and that the addition of acetyl-DL-carnitine reduced the extent of labelling in all six peptides. Their results, however, can be interpreted in a number of ways:

(1) The labelled peptides are representative of the residues around the ligand binding site, thereby providing information regarding the conformation of the polypeptide chain at this site.

(2) An increased number of the labelled peptides are observed due to incomplete cleavage by trypsin as a result of the covalent labelling.

(3) Some number of these peptides are labelled as a result of non-specific binding to sites on the enzyme other than the carnitine binding site. The protection afforded to these sites by carnitine may be due to conformational changes in the enzyme, where upon these sites become hidden from the environment.

In addition to the potential difficulties in interpreting photo-affinity labelling experiments, the use of p-azidophenacyl-CoA to localise the acetyl-CoA binding site of PC requires suitable methods of detecting enzyme bound label. During this study, radioactive coenzyme-A was not readily available from commercial sources, therefore, attempts to incorporate ^{32}P into the 3´ position of coenzyme-A were made.

Labelling with ^{32}P had an advantage over other isotopes in that characterisation of the affinity label could be readily undertaken using gel electrophoresis and autoradiography, thereby avoiding the long exposures often required of fluorography. The greater sensitivity associated with ^{32}P in comparison with other isotopes was also expected to be advantageous, considering the potentially large number of modified proteolytic peptides, generated by complete fragmentation of photolabelled enzyme.

The relatively short half life of ^{32}P, required that the synthesis technique be simple and readily repeatable. A number of methods for labelling 3´-dephospho-coenzyme-A with ^{32}P at the 3´position were tried, including incubation with partially purified 3´-dephospho-coenzyme-A kinase [EC 2.7.1.24] , T4 RNA ligase and treatment with ^{32}P-phosphoric acid. In each case the reactions were analysed by HPLC-anion exchange chromatography, and although in some cases small quantities of radioactivity could be identified as associated with coenzyme-A, the yields of incorporation were wholly inadequate. Other methods of detecting enzyme bound coenzyme-A analogues were therefore investigated.

Attempts were made to raise polyclonal anti-sera to coenzyme-A, bound as a hapten to two different proteins. p-Azidophenacyl-CoA labelled citrate synthase, and coenzyme-A conjugated to BSA (adapted from King *et al.*, 1978) were injected into rabbits, and although antibodies were raised against the carrier molecule in each case, no significant immunological response was raised against the hapten. Although coenzyme-A is principly an intracellular molecule, its lack of immunogenicity observed here suggests that the

immune system of the rabbit is naturally exposed to coenzyme-A, perhaps as a consequence of passage through the serum attached to carrier proteins (Knudsen and Grunnet, 1982).

In the absence of a convenient detection method for enzyme-bound p-azidophenacyl-CoA and the potential for identification of multiple labelling points, the effect of limited trinitrobenzene sulphonate (TNBS) modification was used to study the interaction between acetyl-CoA and PC. As described in section 1.6.3, limited TNBS treatment causes the desensitisation of sheep PC to acetyl-CoA by modification of a single catalytically important lysine residue (dubbed Lys-A; Booker *et al.*, 1986), which can be protected against by coincubation with acetyl-CoA (Keech & Farrant, 1968, Ashman *et al..*, 1973). TNBS is a primary amino group specific reagent at neutral pH, and the mechanism of the reaction is shown in figure 5.1.

PC from yeast was chosen as the subject of this study, because

(1) inhibition by TNBS had been demonstrated in this species (Scrutton & White, 1973)

(2) the complete amino acid sequence was known

(3) the protein was readily available

(4) any resulting hypotheses could be tested by site directed mutagenesis.

5.2 RESULTS AND DISCUSSION

5.2.1 TNBS Modification of Yeast PC

As explained above, yeast PC was the enzyme of choice for this study, therefore, it was important to further characterise the effects of TNBS on this enzyme. Scrutton and White (1973) had demonstrated that treatment of yeast PC with 1.0 mM TNBS, caused the inactivation of the acetyl-CoA-dependent as well as the acetyl-CoA-independent activities of the enzyme. Although their report included results of the TNBS concentration dependence of the inactivation process in rat PC, no further details were included regarding these effects in PC from yeast.

The effects of TNBS, at various concentrations, on yeast PC overall activity were measured over a period of 80 minutes, as described in section 2.2.12. The inactivation profile (figure 5.2-a) shows that with TNBS, at concentrations up to 1.0 mM, the reaction is pseudo first order with respect to time, until 40 % of the initial activity remains. The results shown compare well with the TNBS inactivation studies of Scrutton and White

(1973) and Keech and Farrant (1968) in that inhibition to levels below 40 % of the initial activity results in modification at additional sites on the enzyme. However, the overall rate of inactivation is lower than for the vertebrate enzymes, necessitating considerably longer reaction times to achieve significant inhibition (80 minutes compared to 5-10 minutes for sheep). This increased reaction time suggests that the chemical environment surrounding Lys-A in yeast PC is slightly different to that present in the vertebrate enzymes (ie Lys-A in yeast PC is less nucleophilic).

From this data, as explained in appendix I, the rate of inactivation at each concentration can be calculated. A plot of the rate of inactivation (k') as a function of TNBS concentration (figure 5.2-b) shows that the reaction is pseudo first order with respect to TNBS concentration at concentrations up to 1.0 mM. This again is consistent with the results of Keech and Farrant (1968), who demonstrated that the inactivation of sheep PC by TNBS was pseudo first order with respect to time and TNBS concentration in the range of 0 - 0.67 mM TNBS. It can be seen from figure 5.1-b, that at TNBS concentrations above 1.0 mM, a divergence from linearity occurs, which can most likely be attributed to modification at additional sites on the enzyme.

A replot of this data, showing the log k' as a function of log [TNBS], is a straight line, with a slope = 0.7, suggesting that inactivation of yeast PC is a consequence of modification by 1 molecule of TNBS as described in appendix I. The inactivation of yeast PC observed upon TNBS treatment is therefore likely to be as a result of modification of a single catalytically essential lysine residue (Lys-A). Subsequent inactivation studies were performed using conditions derived from these experiments (ie. 0.5mM TNBS for 80 minutes) in order to minimise modification at any non-catalytic lysine resides.

In order to further characterise the inactivation process in yeast PC, the ability of acetyl-CoA and 3´5´ADP to protect the enzyme against TNBS induced activity loss was tested. Figure 5.3 demonstrates that modification of Lys-A can be protected against by 0.5mM acetyl-CoA, as is the case for PC from sheep (Ashman *et al.* , 1973), chicken and rat (Scrutton *et al.*, 1977). 3´5´ADP at a concentration of 5mM was also shown to protect Lys-A from modification by TNBS, although at 0.5mM no protection was observed. This compares with the results of Chapman-Smith (1981), who showed that sheep PC was protected from inactivation by co-incubation with 0.5mM acetyl-CoA and 0.5mM 3´5´ADP.

Although a similar study on rat PC was inconclusive (Scrutton *et al.*, 1977), comparing these results with those of Chapman-Smith (1981), suggest that Lys-A plays some role in binding acetyl-CoA through the 3´ phosphate of the adenosine moiety.

5.2.2 Acetyl-CoA Binding Assay

The mechanism by which TNBS modification at Lys-A causes inactivation of PC, has been limited to two alternatives, that of hindering the binding of acetyl-CoA, or disabling the mediation of the effect of acetyl-CoA, once bound (Ashman *et al.*, 1973; Scrutton *et al.*, 1973; Chapman-Smith, 1981). The way to resolve these alternatives is to determine whether acetyl-CoA can bind to the modified enzyme.

The measurement of acetyl-CoA binding is complicated by a deacylase activity apparently associated with the enzyme. Frey and Utter (1977) measured this activity in chicken PC, and found at a pH of 7.2 and at 23°C, 4 nmoles of acetyl-CoA was broken down per minute per mg of protein. Although this activity is low compared to the carboxylation of pyruvate (15 μmoles/minute/mg protein), it is sufficient to interfere with the ability to accurately determine the binding coefficient of acetyl-CoA in experiments taking longer than 2.5 minutes to perform at room temperature. These same authors devised a rapid flow dialysis technique, and measured the binding dissociation constant of acetyl-CoA as 13.9 μM. This figure is in good agreement with the activation constant for acetyl-CoA under these conditions (13.3 μM).

In order to minimise the influence of any deacylase activity on the detection of acetyl-CoA binding, the technique must be rapid, sensitive enough to allow sufficiently low concentrations of PC and preferably work at low temperatures. Although chicken and sheep PC are known to undergo a process of cold-inactivation at temperatures below 4°C (Irias *et al.*, 1969), yeast PC does not dissociate at low temperatures (4°C) and was therefore the enzyme of choice for this analysis.

Such a system has been developed and consists of measuring the ability of PC to retard the progress of [^{14}C]acetyl-CoA through an ultrafiltration membrane. This has been applied to the question of whether TNBS-modified PC can bind acetyl-CoA. As described in section 2.2.13, PC samples are mixed with a fixed concentration of acetyl-CoA, and centrifuged through a Mr 30,000 cut off membrane. The filrate is then counted and compared to samples where no PC was added.

As shown in figure 5.4, when TNBS-treated PC samples are compared to untreated PC in such an assay system, a significant difference is observed. Control samples of 0.5 mg/ml, 1.0 mg/ml and 2.0 mg/ml PC were able to retard radioactive material, whereas the same concentrations of TNBS-treated PC were not. Some binding in the 2.0 mg/ml TNBS-treated sample was detected, although this was less than the equivalent control.

Some interference by a deacylase activity can be seen from this data, particularly at higher concentrations of PC, in that at 0.5 mg/ml 51 % of binding sites are occupied, where as at 1.0 mg/ml and 2.0 mg/ml, only 43% and 22% of binding sites are occupied respectively. Given the difference in this assay between the samples modified by TNBS and untreated enzyme, two alternative explanations are possible. Modification by TNBS causes the deacylase activity of yeast PC to increase approximately 4 fold or modification by TNBS inhibits yeast PC from binding acetyl-CoA. Since Ashman *et al.* (1973) showed that the rate of deacylation in samples of sheep PC does not alter upon treatment by TNBS, this first explanation seems very unlikely.

The conclusion that TNBS-treatment of yeast PC inhibits acetyl-CoA binding allows, for the first time, delineation between the two proposed mechanisms of TNBS-mediated inactivation of PC. Although it is possible that the site of modification is distant from that of acetyl-CoA binding, and the inability to bind acetyl-CoA is caused by an irreversible conformational change, the more likely explaination is that lys-A is near the acetyl-CoA binding site and modification of this residues exerts its effects directly. Evidence for this comes from comparing the binding result (figure 5.4) with the effects of acetyl-CoA and 3´5´ADP on the TNBS-mediated inactivation of yeast PC (figure 5.3) and sheep PC (Chapman-Smith, 1981). Both of these ligands are able to protect against inactivation, while acetyl-3´ dephospho-CoA only protects sheep PC poorly and acetyl-phosphopantetheine does not protect this enzyme at all (Chapman-Smith, 1981). The most satisfactory way to demonstrate the involvement of Lys-A in binding acetyl-CoA is by site-directed mutagenesis, substituting a leucine, glutamine, glutamate or arginine residue for the native lysine, testing for the ability of acetyl-CoA to bind to and stimulate the altered enzyme.

In such an experiment, if Lys-A is important for acetyl-CoA binding, assuming these substitutions cause no folding problems, the leucine or glutamine substitutions may be expected to result in an active enzyme but with a lower affinity for acetyl-CoA.

Glutamate substitution at this position might be expected to abolish acetyl-CoA binding, due to the introduction of a negative charge in the binding site. As another coenzyme-A-binding enzyme, citrate synthase, utilises interactions between an arginine residue and the 3´phosphate of coenzyme-A (Remington *et al.*, 1982), substituting an arginine for Lys-A in PC might be expected to result in an active enzyme which can bind acetyl-CoA, although the affinity for this ligand may be altered. The ability to conduct such experiments depends entirely on the isolation and identification of Lys-A.

5.2.3 Analysis of Proteolysed TNBS-modified Yeast PC

Having determined that Lys-A represents a region of the PC polypeptide associated with the acetyl-CoA binding site, it was clearly important to try and determine the identity of Lys-A. Attempts were made to isolate short peptides containing Lys-A after TNBS treatment and complete proteolysis, using the distinctive absorbance at 340nm of TNP-lysine groups as a method of identifying modified peptides. The experiments were performed in duplicate, comparing the results of PC modified in the presence and absence of acetyl-CoA. Under these conditions, the sample modified in the absence of acetyl-CoA would be expected to contain a 345nm absorbing species that was not present in the sample modified in the presence of acetyl-CoA. This peptide would be likely to represent a Lys-A containing peptide. Due to the stabilising effect of acetyl-CoA on the structure of PC, the ligand was removed from samples by rapid gel filtration (section 2.2.) prior to proteolysis.

Proteolysis was effected as described in section 2.2.18 and although purifications were made from samples proteolysed with trypsin, endoproteinase Glu-C at pH 4.0 in ammonium acetate (cleavage C-terminal of glutamyl residues) and endoproteinase Glu-C at pH 7.8 in potassium phosphate (cleavage C-terminal of glutamyl and aspartyl residues) only the profiles from the trypsin-treated experiment are shown (figure 5.5). This figure shows typical absorbance profiles of peptides eluting from reverse-phase HPLC columns.

Figure 5.5-a represents the 345nm absorbing peptides found in trypsinised-yeast PC treated with TNBS in the presence of 0.5mM acetyl-CoA. A major peak, eluting at 26 minutes (37-40% buffer B) can be seen. When compared with figure 5.5-b, which represents a sample treated similarly except that acetyl-CoA was omitted from the TNBS-treatment, it can be seen that this sample contains an additional 345nm absorbing peak eluting at 16 minutes (20-23 % buffer B). These results, together with the activity assays discussed in section 5.2.1 above, demonstrate that TNBS treatment of yeast PC causes modification of mostly two residues, only one of which is catalytically important. These

residues have been dubbed Lys-A (catalytically competent) and Lys-B for convenience. TNBS treatment of sheep PC also results in modification of two lysine residues (Ashman *et al.*, 1973; Chapman-Smith, 1981), only one of which is essential for catalytic activity. Comparing the kinetic data for TNBS modification of sheep (Keech and Farrant, 1968; Ashman *et al.*, 1973), rat (Scrutton and White, 1973) and yeast PC (section 5.2.1), it is likely that the identity of Lys-A is conserved between all three. As Lys-B is not involved in catalysis it is quite possible that this lysine residue is not conserved in its position within the polypeptides from the various species. That is, although the elucidation of Lys-B within the yeast PC sequence will provide useful information regarding the conformation of the polypeptide at that position, comparisons cannot be readily drawn with PCs from other species.

The two semi-purified peptides containing either Lys-A or Lys-B (figures 5.4a &b) were collected and subjected to further chromatography, as shown in figures 5.5c and 5.5d respectively. As trypsin is unable to cleave peptide bonds C-terminal of ε-amino-modified lysines (Allen, 1981), Lys-A or Lys-B will not represent the C-terminus of a tryptic peptide. Therefore the sequence of a TNBS-modified peptide should proceed through the modified lysine, ceasing at the following lysine or arginine. In the absence of radiolabelled TNBS, this criterion represents the most complete method of demonstrating the identity of candidate peptides.

The putative Lys-A tryptic-peptide (100 pmoles) from figure 5.5c was collected and subjected to Edman degradation sequence analysis. Although a small level of contaminating peptide sequence (appendix II) was obtained from this experiment (5 - 10 pmoles), no sequence corresponding to the 100 pmoles of modified peptide was obtained. The identification of the contaminating peptide confirms that the sequencer was performing adequately. There are two possible explainations for the failure of this peptide to sequence, either the N-terminus of the peptide was chemically blocked to the sequencing reagent (PITC), or the peptide had a high propensity to wash off the sample filter within the sequencer during extraction cycles. The latter possibility seems less likely in that the contaminating peptide, which did sequence, eluted at the same position in the gradient during chromatography and is therefore likely to have been of similar hydrophobicity.

When ~ 200 pmoles of the putative Lys-B tryptic-peptide was subjected to sequence analysis, 23 amino acid residues of sequence were determined (as indicated in figure 5.6) which exactly matched the inferred yeast sequence, starting at residue gly^{446}

(for complete sequence of yeast PC see figure 3.10). This residue is preceded in the PC sequence by a lysine residue as would be expected of a trypsin susceptible bond. Given that cleavage by trypsin cannot proceed at lysine residues with a modified ε-amino group, the next nearest C-terminal lysine residue (lys^{881}) is likely to represent Lys-B. The C-terminus of this peptide would therefore be expected to be arg^{882}. Although the sequence analysis of this peptide did not proceed sufficiently far, to allow positive identification of a modified lysine residue, as would be expected in cycle 36 of the analysis, supporting evidence comes from the inability to sequence putative Lys-B peptides after fragmentation of PC by endoproteinase Glu-C. If Lys-B is Lys^{881} as suggested, the N-terminus of the peptide produced by cleavage C-terminal of glu^{875} would be gln^{876} which, under acid conditions, may form an N-terminal N-acetyl-glutamine residue (Allen, 1981), thereby blocking the N-terminus to Edman degradation.

The inability to sequence the putative Lys-A tryptic peptide, prompted the use of other methods of fragmentation of the modified enzyme. Table 5.1 shows a summary of the results of a number of peptide isolations, indicating the amount of starting material for each experiment, protease treatment, yield upon purification and result of sequence analysis. Of the candidate TNP-peptides isolated, under various regimes, only one sample provided N-terminal sequence data, implicating Lys^{881} as the likely identity of Lys-B.

Sequence analysis of candidate Lys-A peptides was undertaken using Edman degradation with little success. Although up to 100pmoles of purified peptides, isolated after trypsin or endoproteinase Glu-C treatment were analysed, no amino acid sequence could be determined. The most likely reasons for this are the known "wash out" from the reaction vessel of very short peptides, and N-terminal blockage. It is also possible that the sensitivity of this peptide sequencer is a limiting factor in identification of the sequence of these peptides. Although fast atom bombardment (FAB) mass spectrometry provides an alternative method of sequencing peptides that are N-terminally blocked, the sensitivity limit of the available spectrometer (≥ 200 pmoles) effectively precluded it from this study.

Due to the low yields encountered in purification experiments such as this, and the apparent inability to analyse candidate peptides by the available means, it is clear that other methods for identifying Lys-A are required. Western analysis of TNBS-modified sheep PC (Appendix III) indicated that Lys-A is contained within the C-terminal Mr 75,000 chymotryptic fragment, and most likely in the C-terminal Mr 28,500 of the protein. Given that Lys-A is likely to be conserved between the various species of PC so far tested, a

sequence comparison, highlighting the conservation of lysine residues in rat and yeast PC is shown in figure 5.7. Thirteen lysine residues are conserved between the two species in this C-terminal 750 residues, limiting the number of likely Lys-A candidates. Five further lysines from yeast PC can be excluded from this list, lys^{1135} - the biotin attachment site, lys^{881} - proposed Lys-B residue and lys^{845}, lys^{734} & lys^{741} - known to be susceptible to trypsin cleavage after TNBS modification.

Of the remaining eight candidate Lys-A lysine residues, only five are located in the C-terminal Mr 28,500 of the enzyme, lys^{1046}, lys^{1051}, lys^{1094}, lys^{1115} and lys^{1124}. On this basis, the most suitable method of determining the identity of Lys-A and hence the location of the acetyl-CoA binding site is with site-directed mutagenesis, targeting these five lysines in the first instance, looking for the effect substitution of each residue has on the acetyl-CoA dependent activity and acetyl-CoA binding. Such analyses are most satisfactorily performed by expressing plasmid-borne mutant constructs in PC deficient yeast strains. Appropriate strains are under construction in this laboratory, but are complicated by the identification of a second, differentially expressed, PC gene in *S. cerevisiae* (M. Walker, *pers. comm.*).

5.2.4 Computer Aided Search for Acetyl-CoA Binding Site

Due to the absence of unambiguous sequence localisation of Lys-A, and hence the acetyl-CoA binding site, sequence similarities between proteins known to bind acyl-CoA molecules were sought. Such an analysis may fail to identify the acetyl-CoA binding site of PC for two reasons,

1) the acetyl-CoA site of PC is an allosteric site and may be formed therefore from different residues and tertiary structures to catalytic sites,

2) the structures of the acetyl-CoA binding sites of citrate synthase (Remington *et al.*, 1982) and chloramphenicol acetyltransferase (Leslie *et al.*, 1988; Day *et al.*, 1988), which share no detectable sequence conservation, have been studied by X-ray diffraction analysis, indicating the involvement of important hydrogen-bonding interactions between the adenosine portion of coenzyme-A and main chain amido and carbonyl groups. These interactions

between the "adenosine recognition loop" and the ligand are therefore not necessarily dependent on a particular sequence of amino acid residues.

Lopez-Casillas *et al.* (1988), speculated that a region of rat ACC (residues 1958 - 1987) constituted an acetyl-CoA binding site, based on a low level of similarity (10%) to the sequence of the adenosine recognition loop of pig citrate synthase (Remington *et al.*, 1982). Significant sequence conservation was however, observed by these authors between this region of rat ACC and the β-subunit of rat and human PCC (~ 45%). While both ACC and the β-subunit of PCC bind acyl-CoAs, for the reasons mentioned in point 2 above, the conclusion that this region of the carboxylases constitutes the adenosine-binding portion of the coenzyme-A binding site seems somewhat premature.

As shown in figure 5.8, the sequence similarities discussed above (Lopez-Casillas *et al.*, 1988) are in fact only part of a much larger region of sequence conservation. When rat ACC (residues 1711 - 2019), rat PCC β-subunit (residues 84 - 391) and *P. shermanii* TC 12S subunit (residues 63 - 363) where compared to each other using the programs SEQHP and BESTFIT, statistically significant similarities were observed (figure 5.8). In addition, the rat PCC β-subunit and *P. shermanii* TC 12S subunit sequences are very highly conserved over nearly their entire lengths (Samols *et al.*, 1988).

This level of sequence conservation is further evidence to support Lynen's hypothesis of the evolution of the biotin carboxylases, in that as indicated in Table 1.1, all three enzymes must be able to bind coenzyme-A. In fact, the finding that the TC 12S subunit and the PCC β-subunit are more closely related in sequence than ACC is not surprising since they catalyse essentially the same reaction (interconversion of methylmalonyl-CoA and propionyl-CoA). It is therefore very likely that the region identified in figure 5.8 represents the coenzyme-A binding domain of the respective enzymes as shown schematically in figure 3.12. It is interesting that the region of rat ACC speculated to be the adenosine binding region by Lopez-Casillas *et al.* (1988) is contained within this proposed domain.

While this sequence comparison localises the likely coenzyme-A binding regions of ACC, PCC and TC, no similarities could be detected with the sequences of rat or yeast PC. This is perhaps not surprising however, due to the different nature of the binding sites (allosteric versus catalytic, as noted in point 1 above).

A number of interesting anomalies are highlighted by the apparent conservation of sequence indicated in figure 3.12. The structure of intact TC (26S species) has been well studied (for review see Wood and Zwolinski, 1976) and comprises 30 polypeptides , including 12 x 1.3S (biotin-carrier), 6 dimeric 5S subunits (pyruvate/oxaloacetate-binding) and a hexameric 12S subunit (methylmalonyl-CoA/propionyl-CoA-binding). In adddition to carrying the biotin moiety, the 1.3S subunits act to bind the 5S subunits to the hexameric core 12S species. Poto *et al.* (1978) using a photoaffinity analogue of acetyl-CoA (p-azido[^{14}C]benzoyl-CoA) were able to demonstrate that each monomer of the 12S subunit binds two molecules of acetyl-CoA.

The active subunit structure of PCC has been shown to be represented by an $\alpha_6\beta_6$ configuration (Haase *et al.*, 1982) and has been observed in the electron microscope to have a similar overall structure to that of transcarboxylase (Haase *et al.*, 1984). Based on these results, in addition to the sequence comparisons shown in figure 5.8, it seems clear that each subunit of PCC contains only one substrate binding site. Comparing this then with the very high level of sequence similarity between the 12S monomer and the PCC β subunit shown in figure 3.17 it is difficult to envisage two acyl-CoA binding sites on the 12S monomer. Although it is likely that the true C-terminus of the 12S monomer does not correspond to that reported by Thornton *et al.* (1987) giving a slightly larger polypeptide than indicated here (N.F.B. Phillips, Cleveland, *pers. comm.*), the size of the 12S monomer is ultimately limited by the polycistronic nature of the operon transcript. The genes in the 3.8 kb TC operon are arranged 12S, 5S and 1.3S (5´to 3´) respectively (Thornton *et al.*, 1987), suggesting that the maximum size of the 12S monomer is 640 - 650 amino acids. As no detectable sequence duplication appears to have occured within the 12S monomer, in order to obtain 12 CoA-ester sites within the hexameric 12S subunit, one would have to speculate that one or both of the CoA sites is made up of contributions from adjacent subunits, or a second form of the 12S monomer exists, encoded by a second gene. Evidence for this second possibility is found in the presence of proteolytic peptide fragments from TC that do not correspond to the sequence predicted by the existing gene sequence (N.F.B. Phillips, Cleveland, *pers. comm.*). Although it is unlikely that the CoA-ester binding sites of any putative second 12S subunit polypeptide would have any sequence homology with the acetyl-CoA site of PC, such a finding would have interesting consequences for the structural model of TC.

5.2.5 Localisation of Lys-A Within the Quaternary Structure of Yeast PC

Characterisation of the acetyl-CoA binding site involves delineation of the amino acid residues in and around the binding site, and also the relative position within the overall shape of the molecule. Although the amino acids around the site have yet to be determined, as described in sections 5.2.3 & 5.2.4, it remains important that the binding site be localised within the quaternary structure of the enzyme. This has been achieved using the electron microscope to visualise Fab molecules complexed with TNBS-modified yeast PC. Samples were mounted for electron microscopy and analysed by Professor Frank Mayer, a visiting scientist from the University of Göttingen, FRG.

Fab fragments, from serum raised against TNP conjugated to BSA, were isolated as described in section 2.2.15. Yeast PC was modified with 0.5mM TNBS for 80 minutes at 30°C as described in section 2.2.17. In order to preserve the tetrameric structure of PC during the mounting procedure, TNBS-treated enzyme was cross-linked with 1 mM DTSP for 10 minutes at 30°C. After quenching the reaction by the addition of L-lysine to 10 mM, tetramers of PC were isolated by gel filtration chromatography. The resulting particles were then complexed with anti-TNP Fab molecules. Gel filtration chromatography was used to separate the complexes from the excess of unbound Fab molecules, which would otherwise complicate the analysis. The complexes were mounted and stained with 4% (w/v) uranyl acetate and observed in the electron microscope at a final magnification of 250,000x. The experiment was conducted in three parts, in order to differentiate between Fab fragments bound at Lys-A and Lys-B. That is, PC samples were either TNBS modified in the absence of any additions, modified in the presence of 0.5mM acetyl-CoA or unmodified.

A gallery of tetrameric PC particles bound to one or more Fab molecules is shown in figure 5.9. Group-a of the gallery represents PC particles modified in the presence of 0.5mM acetyl-CoA (mostly Lys-B) and group-b represents PC modified in the absence of acetyl-CoA (Lys-A and Lys-B). Fab fragments can be seen as small white masses in these micrographs, attached to larger more diffuse tetramers of PC. No binding of Fab molecules was detected in control micrographs of samples not treated with TNBS (data not shown).

Diagrammatic interpretations of the micrographs are also shown in the figure, with Fab molecules depicted as small black particles bound to large white PC tetramers. In most instances, the characteristic four mass centres of PC are clearly visible, representing the

individual subunits. In gallery-a where acetyl-CoA was included in the TNBS treatment (ie. modification mostly at Lys-B), complexes are observed where Fab molecules lie toward the edges of the subunits, though still toward the inter-subunit junction. When TNBS modification was conducted in the absence of acetyl-CoA (ie. at Lys-A & Lys-B), two positions of Fab binding are seen (gallery-b), one representing the same view as samples treated in the presence of acetyl-CoA, the other centrally located within the tetramer.

Observation of modified enzyme-Fab complexes clearly indicates that the region of the polypeptide immediately surrounding Lys-A and Lys-B are at the surface of the molecule. This is not surprising in that these two lysine residues must be available to the solvent in order to react with TNBS. The results shown in figure 5.9 localise Lys-A toward the centre of each tetramer, near the active site of the enzyme as shown schematically in figure 5.10. This finding is consistent with the results of a similar electron microscopic studies of TNBS-modified sheep PC (Chapman-Smith, 1981; M. Rohde, unpublished results).

Western analysis experiments, summarised in appendix III, indicate that Lys-A is most likely within the C-terminal Mr 28,500 of each subunit. This is also consistent with the electron microscopy results presented here, given that the domain model (section 4.3.2) predicts that the biotin-carrier domain (C-terminal Mr 20,000) of yeast PC is likely to coincide with the region implicated as containing Lys-A (figure 5.10).

From the micrographs (figure 5.9) the position of Lys-B appears to be more towards the edge of each subunit, compared with Lys-A, yet still near the inter-subunit junction. This again is consistent with the proposed domain model for PC (figure 4.14) and the identity of Lys-B (lys^{881}), which in the primary structure is at the C-terminal end of the pyruvate domain.

5.3 CONCLUSION

Delineation of the mechanism of allosteric activation of PC by acetyl-CoA has been a major aim of much research over the last two and a half decades. While the nature of the residues involved in the acetyl-CoA binding site have only been poorly characterised, the finding here that the allosteric site (represented by Lys-A) is at or near the reaction centre of yeast PC provides a useful clue as to the likely mechanism of activation.

As discussed by Keech and Attwood (1985), acetyl-CoA has been observed to have a number of effects on PC,

(1) it increases the overall reaction rate

(2) it decreases the apparent Km value for both HCO_3^- and pyruvate

(3) it is absolutely required for the ATP:Pi exchange reaction

(4) it stimulates the rate of pyruvate:oxaloacetate exchange reaction

(5) it protects the enzyme from cold and dilution inactivation

(6) it induces the "tight" tetrahedral conformation of PC

(7) it allows transfer of $^{14}CO_2$ onto biotin.

The finding that acetyl-CoA is able to influence events at both partial reaction subsites led Clements (1977) to propose that acetyl-CoA spans the active site, enhancing the reaction rate by shielding the reaction centre to the aqueous environment. Although this role for acetyl-CoA appears to be contrary to the current view of allosterism, whereby binding of the allosteric activator induces a favourable conformational change, it does fit the experimental findings summarised above.

At this point, it is not clear whether the "hydrophobic capping" model of Clements (1977) is effective in enhancing the reaction rate or if the role of acetyl-CoA is simply to hold the two reaction subsites at appropriately close distances, facilitating reaction between the substrates and the biotin moiety. The localisation of the acetyl-CoA binding site to the region surrounding the reaction centre of PC, while confirming the likelihood of these proposals, cannot discriminate between them.

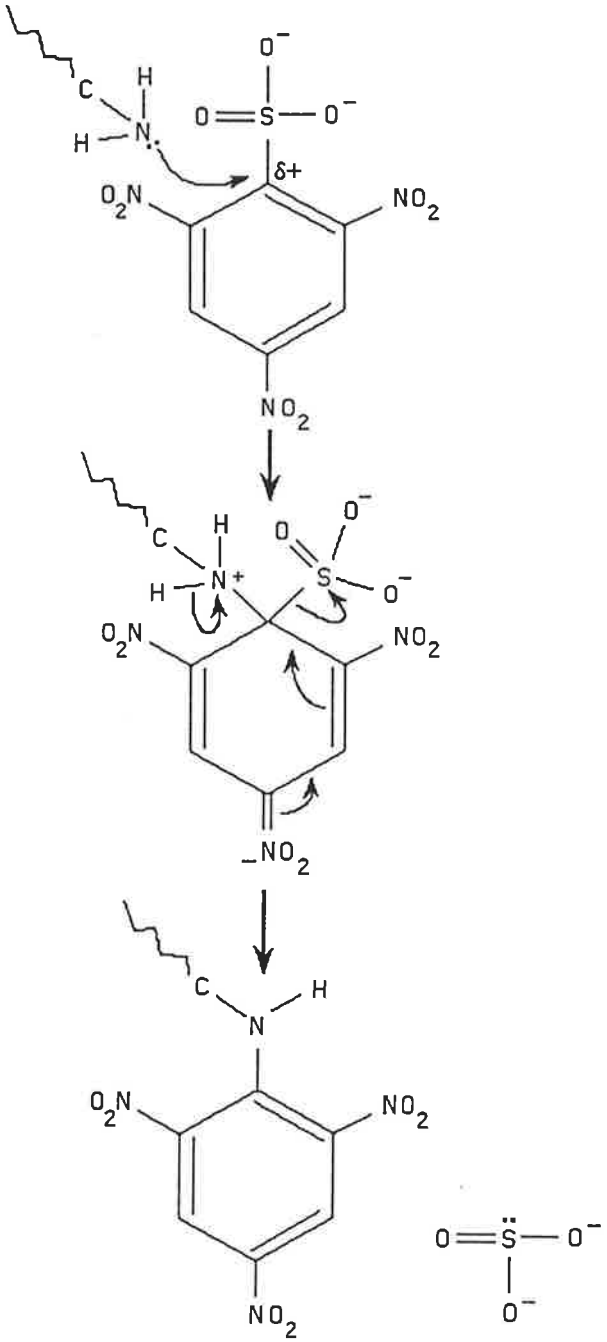


Figure 5.1 The mechanism of the reaction between trinitrobenzene sulphonic acid (TNBS) and a primary amino group.

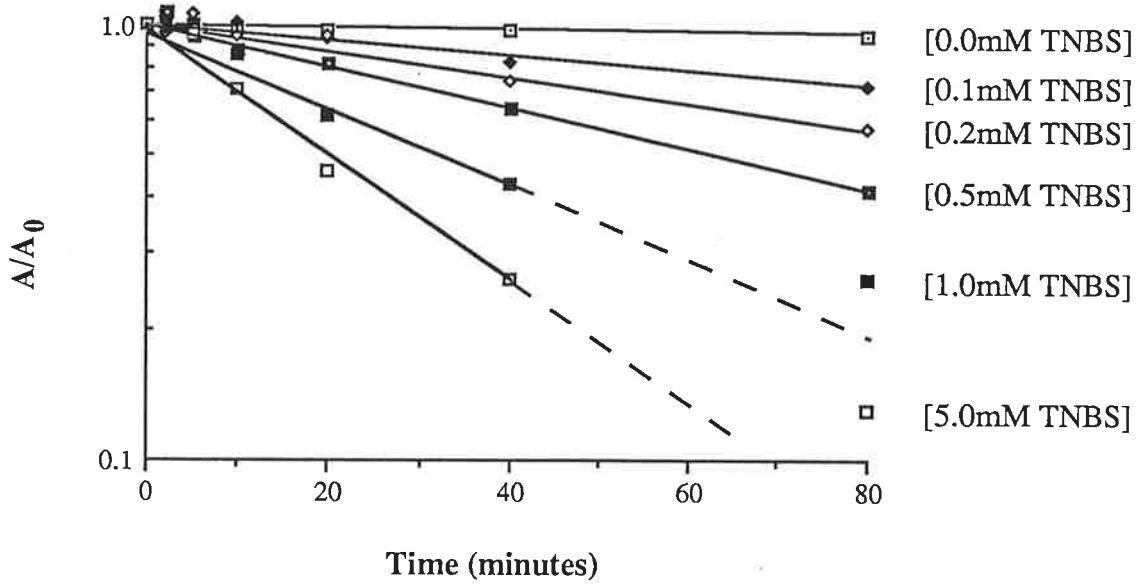
Figure 5.2 The Effect of TNBS-Modification on the Activity of Yeast Pyruvate Carboxylase

a) Yeast pyruvate carboxylase (s.a. 12 U/mg) was treated with TNBS [between 0.0 and 5.0 mM] at 30°C, and at the times indicated samples were removed and assayed for pyruvate carboxylase activity as described by Keech and Farrant (1968). The data is presented as the fraction of initial activity remaining (A/A0) versus time for each concentration of TNBS.

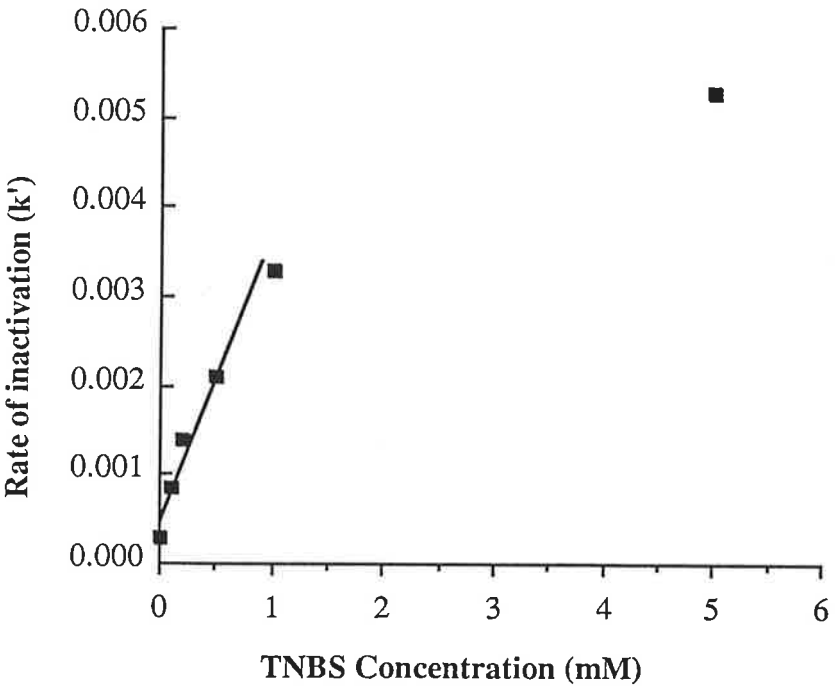
b) The rate of inactivation of yeast pyruvate carboxylase at each concentration of TNBS was calculated from the data in figure 5.2-a (above) as described in appendix I. This data is shown as k´ versus TNBS concentration.

c) A replot of the data from figure 5.2-b, in the form of log k´versus log(concentration of TNBS) is shown. The slope of this line (TNBS concentrations in the range 0.1-1.0 mM) is an indication of the order of the reaction as indicated in appendix I. The slope of this line is 0.7.

a) YPC - Concentration-dependence of TNBS Inhibition



b) YPC - Rate of Inactivation by TNBS



c) YPC - Order of the Reaction

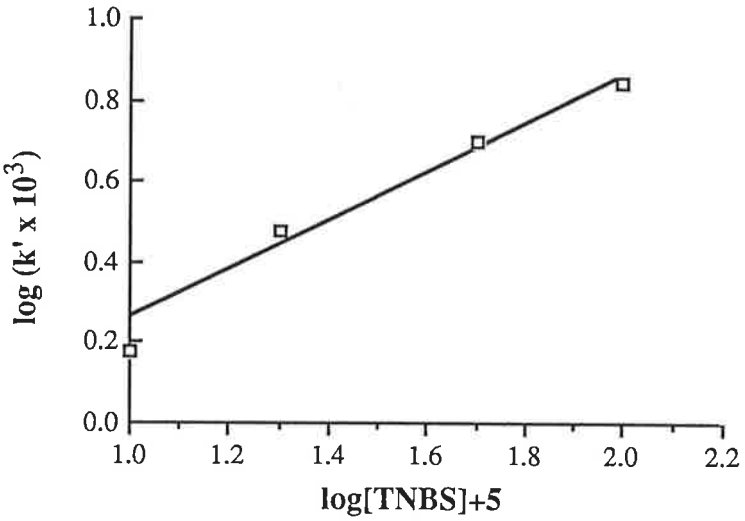


Figure 5.3 The Effect of Acetyl-CoA and 3′5′ADP on the TNBS-Mediated Inactivation of Yeast Pyruvate Carboxylase

Yeast pyruvate carboxylase (1 mg/ml, s.a. 12 U/mg) was treated with TNBS [0.5 mM] at 30°C in 100 mM Tris-Cl pH 7.2 in the presence of various additions. At the times indicated samples were removed and assayed for pyruvate carboxylase activity as described by Keech and Farrant (1968). The data is presented as the fraction of initial activity remaining (A/A_0) versus time.

Additions:

- □ - no TNBS (control)
- ● - 0.5 mM acetyl-CoA
- ■ - 5.0 mM 3′5′ADP
- ○ - 0.5 mM 3′5′ADP
- △ - no additions

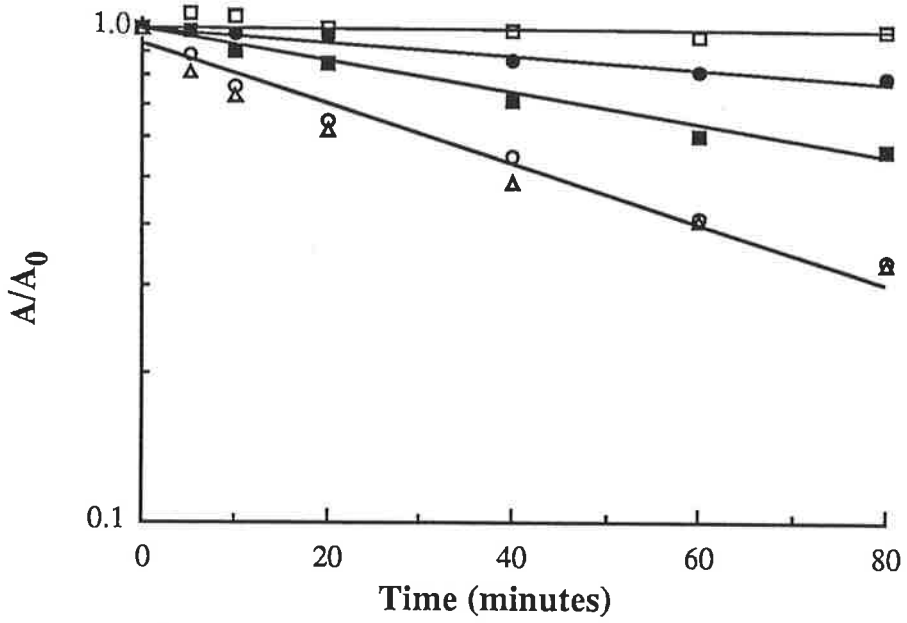

1.0
0.1
A/A0
0
20
40
60
80
Time (minutes)

Figure 5.4 The Effect of TNBS-Modification on the Capacity of Yeast Pyruvate Carboxylase to Bind Acetyl-CoA

The ability of TNBS-treated and control yeast pyruvate carboxylase samples to bind acetyl-CoA was tested as described in section 2.2.13.

Yeast pyruvate carboxylase (3 mg/ml, s.a. 13.3 U/mg) was treated with TNBS [0.5 mM] at 30°C in 100 mM Tris-Cl pH 7.2 until 40% initial activity remained. The amount of [^{14}C]acetyl-CoA in the filtrate is expressed as a percentage of that observed when pyruvate carboxylase or TNBS-treated pyruvate carboxylase was omitted from the binding solution.

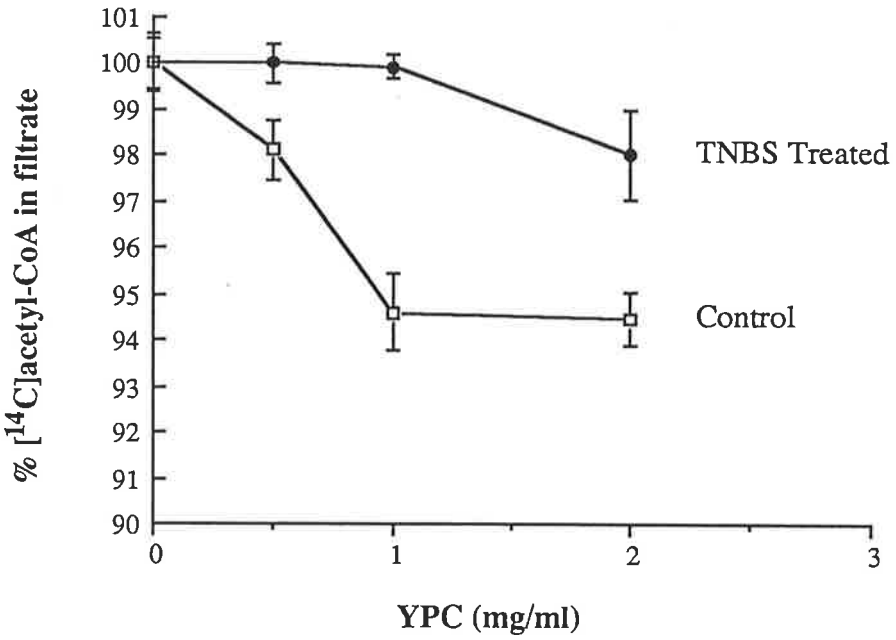

% [14C]acetyl-CoA in filtrate
101
100
99
98
97
96
95
94
93
92
91
90
0
1
2
3
YPC (mg/ml)
TNBS Treated
Control

Figure 5.5 Analysis and Isolation of TNBS-Modified Peptides Resulting from Exhaustive Trypsin Digestion of TNBS-Modified Yeast Pyruvate Carboxylase

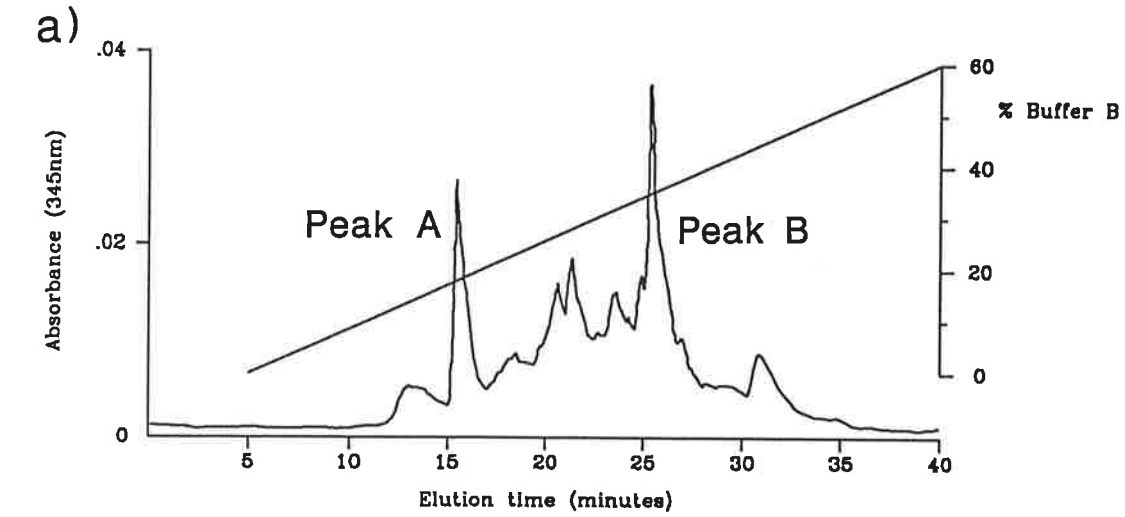
Yeast pyruvate carboxylase (1 mg/ml, s.a. 20 U/mg) in 100 mM Tris-Cl pH 7.2 was treated with TNBS [0.5 mM] at 30°C for 80 minutes. A sample was similarly modified, but in the presence of 0.5 mM acetyl-CoA. The reactions were quenched by the addition of L-lysine to 10 mM. Small molecular weight materials were removed from these samples by rapid gel filtration (section 2.2.5). After exhaustive trypsinolysis in the same buffer containing 4 M urea and 1 mM DTT, samples were subjected to reverse phase high performance liquid chromatography (rp-HPLC) as described in section 2.2.18.

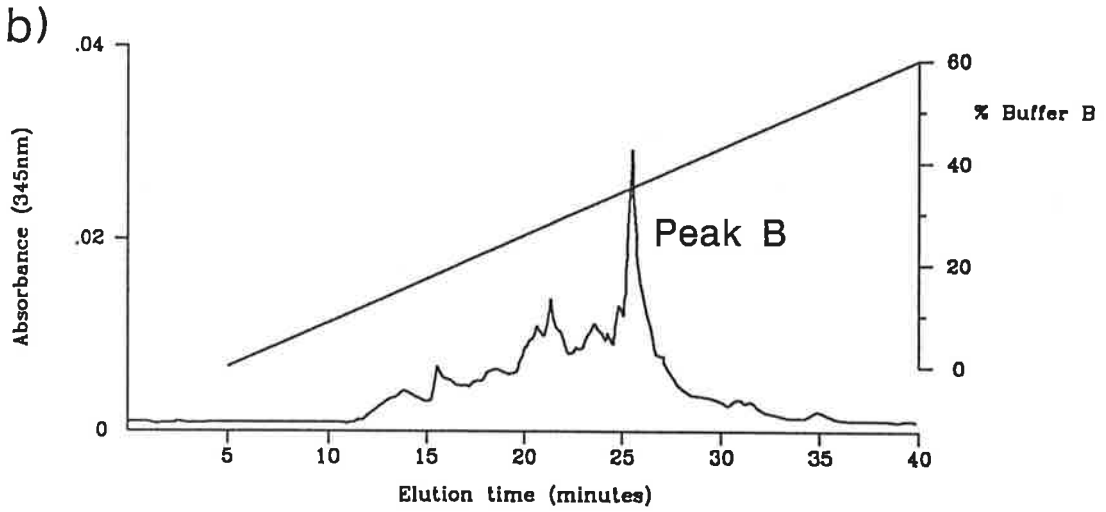
a) 480 µg of trypsinised, TNBS-treated yeast pyruvate carboxylase was applied to a butyl-silica (C4) column (2.1 mm I.D. X 30 mm). The flow rate was 0.2 ml/minute and the initial solvent conditions consisted of 0.1% (v/v) trifluoroacetic acid (TFA). Elution was effected by an increasing concentration of buffer B (0.1% (v/v) TFA/acetonitrile) over 40 minutes as shown. The profile represents the absorbance detected in the sample at a wavelength of 345 nm. Fractions were collected at 30 second (100 µl) intervals across the time period.

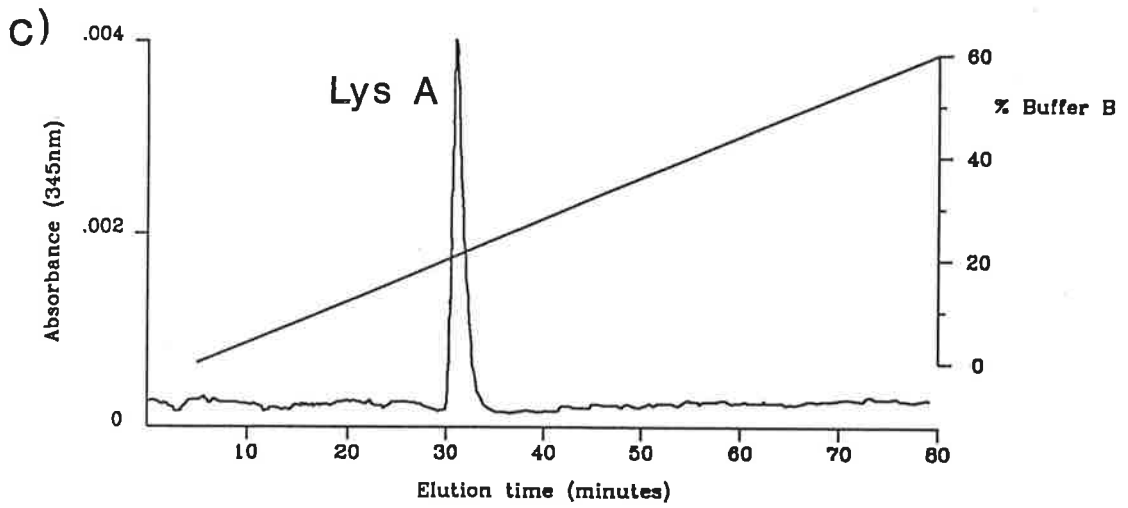
b) 480 µg of trypsinised, TNBS-treated (in the presence of 0.5 mM acetyl-CoA) yeast pyruvate carboxylase was applied to a butyl-silica (C4) column (2.1 mm I.D. X 30 mm). The flow rate was 0.2 ml/minute and the initial solvent conditions consisted of 0.1% (v/v) TFA. Elution was effected by an increasing concentration of buffer B (0.1% (v/v) TFA/acetonitrile) over 40 minutes as shown. The profile represents the absorbance detected in the sample at a wavelength of 345 nm. Fractions were collected at 30 second (100 µl) intervals across the time period .

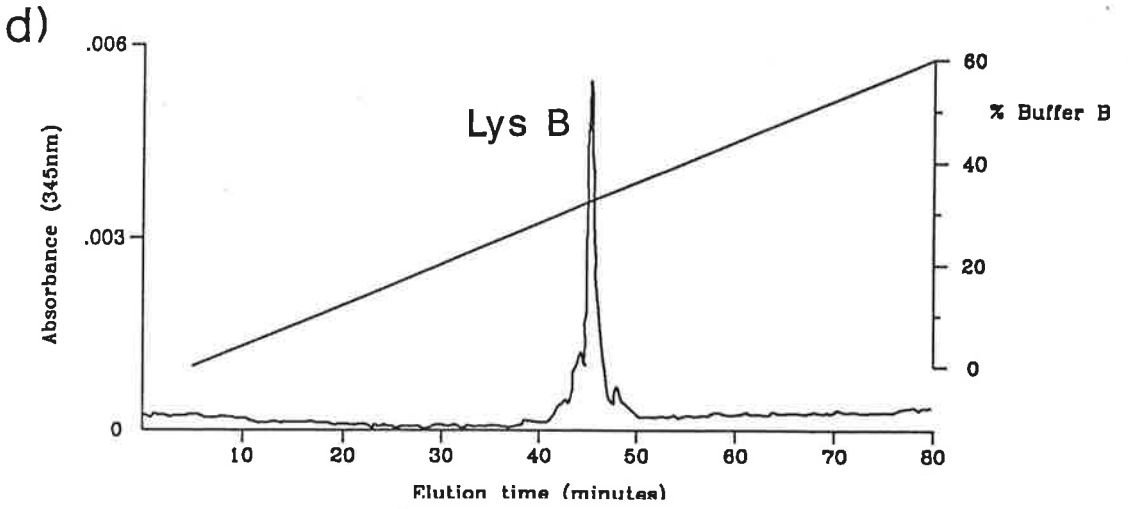
c) Approximately 250 pmoles of Peak A (from figure 5.5-a above) was applied to a butyl-silica (C4) column (2.1 mm I.D. X 30 mm). The flow rate was 0.2 ml/minute and the initial solvent conditions consisted of 0.1% (v/v) heptafluorobutyric acid (HFBA). Elution was effected by an increasing concentration of buffer B(0.1% (v/v) HFBA/propanol) over 80 minutes as shown. The profile represents the absorbance detected in the sample at a wavelength of 345 nm. Fractions were collected at 30 second (100 µl) intervals across the time period.

d) Approximately 350 pmoles of Peak B (from figure 5.5-b above) was applied to a butyl-silica (C4) column (2.1 mm I.D. X 30 mm). The flow rate was 0.2 ml/minute and the initial solvent conditions consisted of 0.1% (v/v) HFBA. Elution was effected by an increasing concentration of buffer B(0.1% (v/v) HFBA/propanol) over 80 minutes as shown. The profile represents the absorbance detected in the sample at a wavelength of 345 nm. Fractions were collected at 30 second (100 µl) intervals across the time period.


a)
Peak A
Peak B
Absorbance (345nm)
.04
.02
0
% Buffer B
60
40
20
0
5
10
15
20
25
30
35
40
Elution time (minutes)


b)
Peak B
Absorbance (345nm)
.04
.02
0
% Buffer B
60
40
20
0
5
10
15
20
25
30
35
40
Elution time (minutes)


c)
Lys A
Absorbance (345nm)
.004
.002
0
% Buffer B
60
40
20
0
10
20
30
40
50
60
70
80
Elution time (minutes)


d)
Lys B
Absorbance (345nm)
.006
.003
0
% Buffer B
60
40
20
0
10
20
30
40
50
60
70
80
Elution time (minutes)

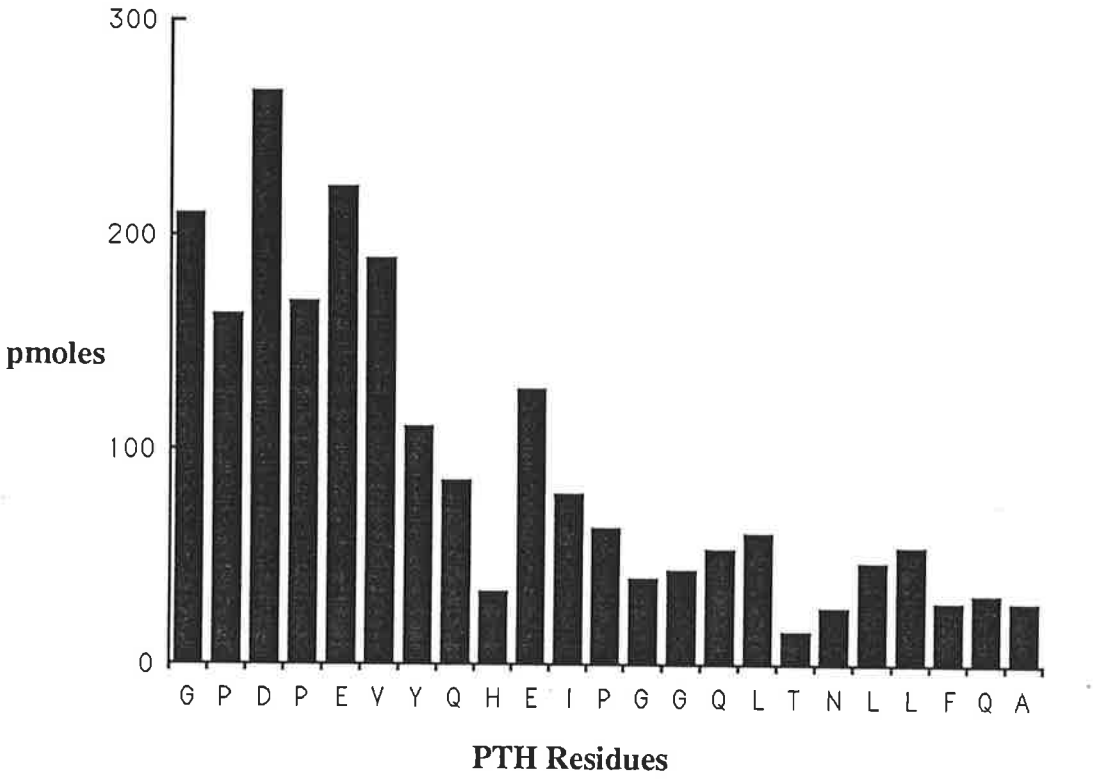


Figure 5.6 Sequence of Lys-B peptide Approximately 200 pmoles of the TNBS-modified tryptic fragment of yeast PC, isolated by rp-HPLC as shown in figure 5.5, was subjected to Edman degradation (Hunkapillar *et al* , 1983). The yield of the PTH-derivative liberated in each cycle is shown, along with the corresponding amino acid sequence in single letter code. The sequence analysis terminated due to mechanical failure after the 23rd cycle.

Amount of starting material	Protease	Yield of purified peptide		Sequence obtained
0.48 mg yPC (3.7 nmoles)	trypsin	Lys-A Lys-B	100 pmoles 200 pmoles	none GPDPEVYQ....
1.94 mg yPC (16 nmoles)	endo. glu-C (C-terminal of glu)	Lys-A Lys-B	50 pmoles 60 pmoles	none none
0.6 mg yPC (4.7 nmoles)	endo. glu-C (C-terminal of asp & glu)	Lys-A Lys-B	10 pmoles —	not submitted —

Table 5.1 Summary of TNBS-Modified Peptide isolations

Figure 5.7 Comparison of the Sequences of the C-terminal 750 Amino Acids Rat and Yeast Pyruvate Carboxylase, Highlighting the Conservation of Lysine Residues

The amino acid sequences of rat pyruvate carboxylase (RPC), yeast pyruvate carboxylase (YPC; Lim *et al.*, 1988) were aligned for maximum homology using the program FASTA, as shown in figure 3.10. The lysine residues conserved between the two sequences are highlighted in the figure by (•) .

```
                                        10        20        30
Ratpc                                DLECLSLRKLHSALLLALQTNIPFLQNVLNNQQFLA

YeastPC AGTIISPHYDSMLVKCSCSGSTYEIVRRKMIRALIEFRIRGVKTNIPFLLTLLTNPVFIE
        400       410       420       430       440       450

          40        50        60        70        80        90
Ratpc   GIVDTQFIDENPELFQLRPAQNRAQKLLHYLGLPALPFSSLLTFSCFPP-CAPALVKDTS
                                    •
YeastPC GTYWGTFIDDTPQLFQMVSSQNRAQKLLHYLADVADNGSSIKGQIGLPKLKSNPSVPHLH
        460       470       480       490       500       510

          100       110       120       130       140       150
Ratpc   WSMALPLQSPSSPPPAGFRDILLREGPEGFARAVRNHQGLLLMDTTFRDAHQSLLATRVR

YeastPC DAQGNVINVTKSAPPSGWRQVLLEKGPAEFARQVRQFNGTLLMDTTWRDAHQSLLATRVR
        520       530       540       550       560       570

          160       170       180       190       200       210
Ratpc   THDLKKIAPYVAHNFNNLFSIENWGGATFDVAMRFLYECPWRRLQELRELIPNIPFQMLL

YeastPC THDLATIAPTTAHALAGAFALECWGGATFDVAMRFLHEDPWQRLRKLRSLVPNIPFQMLL
        580       590       600       610       620       630

          220       230       240       250       260       270
Ratpc   RGANAVGYTNYPDNVVFKFCEVAKENGMDVFRIFDSLNYLPNMLLGMEAAGSAGGVVEAA
                                    •
YeastPC RGANGVAYSSLPDNAIDHFVKQAKDNSVDIFRVFDALNDLEQLKVGVDAVKKAGGVVEAT
        640       650       660       670       680       690

          280       290       300       310       320       330
Ratpc   ISYTGDVADPSRTKYSLEYYMGLAEELVRAGTHILCIKDMAGLLKPAACTMLVSSLRDRF
                    •                             •       •
YeastPC VCFSGDMLQPGK-KYNLDYYLEIAEKIVQMGTHILGIKDMAGTMKPAAAKLLIGSLRAKY
        700       710       720       730       740       750

          340       350       360       370       380       390
Ratpc   PDLPLHIHTHDTSGSGVAAMLACAQAGADVVDVAVDSMSGMTSQPSMGALVACTKGTPLD

YeastPC PDLPIHVHTHDSAGTAVASMTACALAGADVVDVAINSMSGLTSQPSINALLASLEGN-ID
         760       770       780       790       800       810

          400       410       420       430       440       450
Ratpc   TEVPLERVFDYSEYWEGARGLYAAFDCTATMKSGNSDVYENEIPGGQYTNLHFQAHSMGL
                                    •
YeastPC TGINVEHVRELDAYWAEMRLLYSCFE--ADLKGPDPEVYQHEIPGGQLTNLLFQAQQLGL
          820       830       840       850       860       870

          460       470       480       490       500       510
Ratpc   GSKFKEVKKAYVEANQSAGGPHQGDTILQDCGGSGPVHGAERVEPGRGRSSGRAAVLPPL
              •
YeastPC GEQWAQTKRAYREANYLLGDIVKVTPTSKVVGDLAKFMVSNKLTSDDVRRLANSLDFPDS
            880       890       900       910       920       930

          520       530       540       550       560       570
Ratpc   CGGVPAGLHWHSPWGFPEPFRSKVGKVLKDLPRIEGRPGASLPPLNLKELEKDLIDRHGE

YeastPC VMDFFEGLIGQPYGGFPEPFRSDV--LRNKRRKLTCRPGLELEPFDLEKIREDLQNRFG-
            940       950       960       970       980       990

          580       590       600       610       620       630
Ratpc   EVYPEDVLSAAMYPDVFAQFKDFTATFGPLDSLNTRLFLQGPKIAEEFEVELERGKTLHI
                                                                •
YeastPC DVDECDVASYNMYPRVYEDFQKMRETYGDLSVLPTRSFLSPLETDEEIEVVIEQGKTLII
              1000      1010      1020      1030      1040      1050

          640       650       660       670       680       690
Ratpc   KALAVSDLN-RAGQRQVFFELNGQLRSILVKDTQAMKEMHFHPKALKDVKGQIGAPMPGK
        •                                           •
YeastPC KLQAVGDLNKKTGEREVYFDLNGEMRKIRVADRSQKVETVTKSKADMHDPLHIGAPMAGV
              1060      1070      1080      1090      1100      1110

          700       710       720       730       740       750
Ratpc   VIDVKVAAGAKVVKGQPLCVLSAMKMETVVTSPMEGTIRKVHVTKDMTLEGDDLILEIE
            •         •         •
YeastPC IVEVKVHKGSLIKKGQPVAVLSAMKMEMIISSPSDGQVKEVFVSDGENVDSSDLLVLLED
              1120      1130      1140      1150      1160      1170

YeastPC QVPVETKA
```

Figure 5.8 Comparison of the Amino Acid Sequences of the Coenzyme-A-binding Regions of Rat Acetyl-CoA Carboxylase, Rat Propionyl-CoA Carboxylase β-subunit and Transcarboxylase 12S-monomer

Regions of the amino acid sequences of rat acetyl-CoA carboxylase (Rat ACC; Lopez-Casillas *et al.*, 1988), rat propionyl-CoA carboxylase β-subunit (Rat PCCβ; Kraus *et al.*, 1986) and 12S monomer of transcarboxylase from *Propionibacterium shermanii* (TC12S; Thornton *et al.*, 1987) were aligned for maximum homology using the program SEQHP. Residues which are conserved between the proteins are highlighted. All three polypeptides are known to bind acyl-CoA. Gaps introduced into the sequences to maximise the alignment are indicated by dashes. The statistical significance of each alignment was tested using the program SEQDP, in each case the probability of the comparisons being the result of a chance match were calculated to be less than 2×10^{-16}.

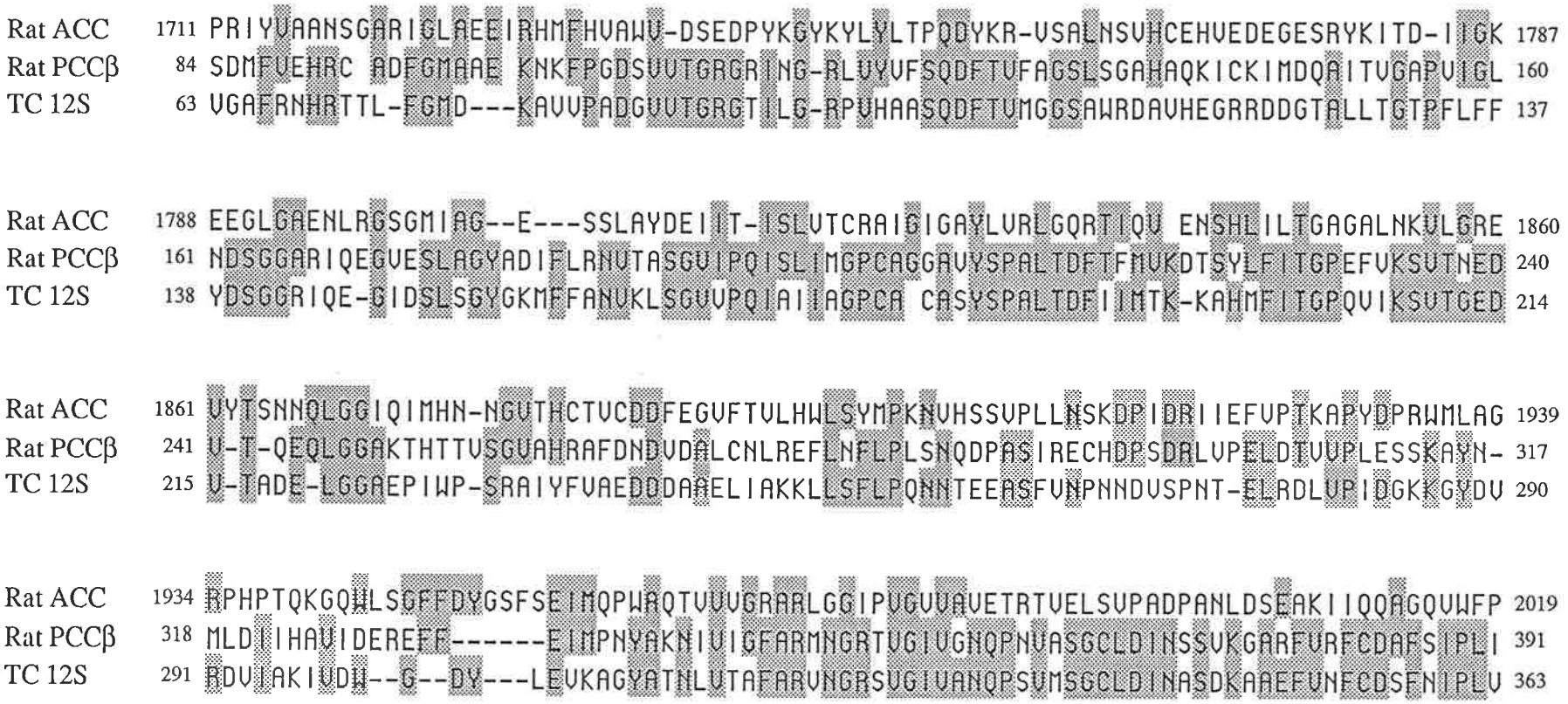

Rat ACC 1711 PRIYVAANSGARIGLAEEIRHMFHVAWV-DSEDPYKGYKYLYLTPQDYKR-VSALNSVHCEHVEDEGESRYKITD-IIGK 1787
Rat PCCβ 84 SDMFVEHRC ADFGMAAE KNKFPGDSVVTGRGRING-RLVYVFSQDFTVFAGSLSGAHAQKICKIMDQAITVGAPVIGL 160
TC 12S 63 VGAFRNHRTTL-FGMD---KAVVPADGVVTGRGTILG-RPVHAASQDFTVMGGSAWRDAVHEGRRDDGTALLTGTPFLFF 137
Rat ACC 1788 EEGLGAENLRGSGMIAG--E---SSLAYDEIIT-ISLVTCRAIGIGAYLVRLGQRTIQV ENSHLILTGAGALNKVLGRE 1860
Rat PCCβ 161 NDSGGARIQEGVESLAGYADIFLRNVTASGVIPQISLIMGPCAGGAVYSPALTDFTFMVKDTSYLFITGPEFVKSVTNED 240
TC 12S 138 YDSGGRIQE-GIDSLSGYGKMFFANVKLSGVVPQIAIIAGPCA CASYSPALTDFIIMTK-KAHMFITGPQVIKSVTGED 214
Rat ACC 1861 VYTSNNQLGGIQIMHN-NGVTHCTVCDDFEGVFTVLHWLSYMPKNVHSSVPLLNSKDPIDRIIEFVPTKAPYDPRWMLAG 1939
Rat PCCβ 241 V-T-QEQLGGAKTHTTVSGVAHRAFDNDVDALCNLREFLNFLPLSNQDPASIRECHDFSDRLVPELDTVVPLESSKAYN- 317
TC 12S 215 V-TADE-LGGAEPIWP-SRAIYFVAEDDDAAELIAKKLLSFLPQNNTEEASFVNPNNDVSPNT-ELRDLVPIDGKKGYDV 290
Rat ACC 1934 RPHPTQKGQWLSGFFDYGSFSEIMQPWAQTVVVGRARLGGIPVGVVAVETRTVELSVPADPANLDSEAKIIQQAGQVWFP 2019
Rat PCCβ 318 MLDIIHAVIDEREFF------EIMPNYAKNIVIGFARMNGRTVGIVGNQPNVASGCLDINSSVKGARFVRFCDAFSIPLI 391
TC 12S 291 RDVIAKIVDH--G--DY---LEVKAGYATNLVTAFARVNGRSVGIVANQPSVMSGCLDINASDKAAEFVNFCDSFNIPLV 363

Figure 5.9 Electron Micrographs of Complexes Formed Between Yeast Pyruvate Carboxylase and Fab Fragments of Anti-TNP Antibodies.

Yeast pyruvate carboxylase (1 mg/ml, s.a. 21 U/mg) was modified with TNBS, crosslinked with DTSP, complexed with anti-TNP Fab fragments and prepared for electron microscopy as described in section 2.2.17.

Shown here are example views from micrographs of complexes between anti-TNP Fab fragments and TNBS-modified yeast pyruvate carboxylase. Also shown are diagramatic interpretations of the micrographs, where the pyruvate carboxylase tetramers are represented by large white molecules and the Fab fragments as smaller, black molecules. The bar is shown for scale and represents 20 nm.

a) Complexes observed in samples where yeast pyruvate carboxylase has been modified with TNBS in the presence of 0.5 mM acetyl-CoA (ie. mostly at Lys-B).

b) Complexes observed in samples where yeast pyruvate carboxylase has been modified with TNBS in the absence of any other additions (ie. at Lys-A and Lys-B).

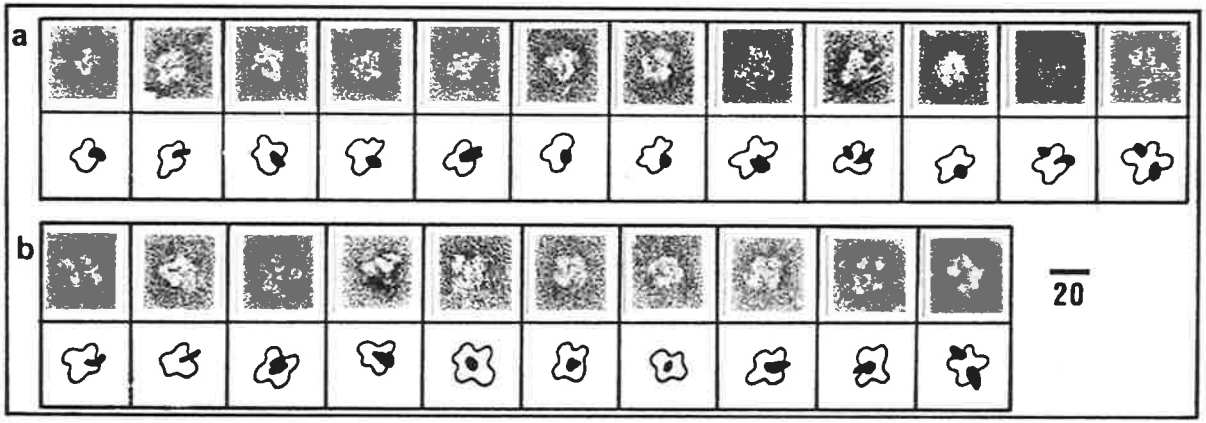
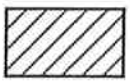
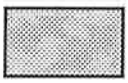

a
b
20

Figure 5.10 A Schematic Representation of The Positions of Lys-A and Lys-B as Interpreted from Electron Micrographs of Complexes between Pyruvate Carboxylase and Fab Fragments of Anti-TNP Antibodies.

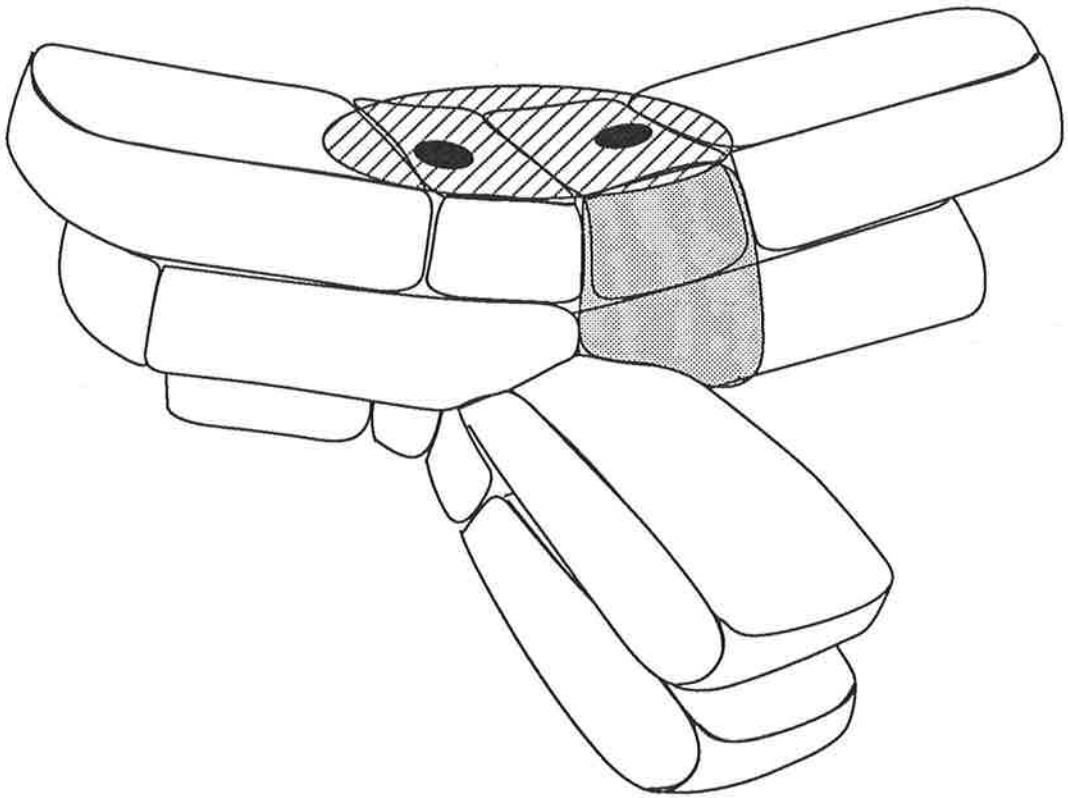
The regions of yeast pyruvate carboxylase implicated by electron microscopy as containing the TNBS modified residues Lys-A and Lys-B, are shown within the context of the proposed domain model.



represents the region of Fab binding at Lys-A



represents the region of Fab binding at Lys-B



CHAPTER 6

DISCUSSION & CONCLUSIONS

The overall aim of the work presented in this thesis, and the on-going aim of this laboratory, is to relate the structure of pyruvate carboxylase to its various catalytic, regulatory and conformational properties. This falls within the wider aim of structural biology, to understand what forces govern the correct folding of large proteins and their interactions with other molecules. Furthermore new enzymes or proteins may then be constructed *de novo* to solve given tasks, on the basis of this understanding. The first step in any such quest is the characterisation of the enzyme's structure, from the primary amino acid sequence to the tertiary and quaternary structure of the polypeptide in 3-dimensions. These results, when paired with studies on functional aspects of the protein, may provide the answers to such important questions as the role of conformational changes and ligand binding in catalysis. Pyruvate carboxylase has therefore, been used as a model system for the study of large complex enzymes.

Described in this chapter is an overview of the results and their implications with regard to the catalytic and allosteric mechanisms, in addition to a discussion of possible future experiments.

6.1 Catalysis

While the chemistry of the reaction catalysed by pyruvate carboxylase has been the subject of intense interest over many years (Knowles, 1989), far less attention has been placed on the particular structures within the enzyme that allow such chemistry to occur. Both the carboxylation of biotin and its subsequent decarboxylation are thought to be base catalysed, yet the identities of such bases remain unknown.

The sequence comparisons shown in chapter 3 have highlighted a number of residues in pyruvate carboxylase that are conserved between proteins of related function. These residues lie within regions predicted to form the 1st and 2nd partial reaction subsites and are therefore potentially important for activity. In particular, the absolutely conserved cysteine residue in the ATP/HCO_3^- binding site (yeast PC- cys^{249}) may be the catalytic base postulated by Tipton and Cleland (1988) as shown in figure 1.2. In addition, a number of lysines residues (yeast PC - lys^{178}, lys^{257} & lys^{284}) are highly conserved, one of which may represent the reactive lysine residue implicated at the ATP-binding site by Easterbrook-Smith *et al.* (1976). Keech and Attwood (1984) and Tipton and Cleland (1988) have suggested that one of the roles of a catalytic base in the 1st partial reaction is to stabilise the enol-form of biotin, in order to enhance its nucleophilicity toward the carboxyphosphate.

The sequence analysis of PC from rat, presented in chapter 3, has allowed further comparison of the sequences associated with the 2nd partial reaction subsites of yeast PC, transcarboxylase and oxaloacetate decarboxylase. A cysteine residue conserved between

the latter three sequences is not observed in the rat PC sequence, suggesting that it is not likely to represent the catalytically important cysteine residue observed by Hudson *et al.* (1975), Bagley (1982) and Bagley *et al.* (1983). The comparison has however highlighted a number of other conserved residues which may be important for the structure or catalytic mechanism of the 2nd partial reaction. It is clear that with the accumulating sequence information and the availability of cloned DNAs for a number of biotin carboxylases, site directed mutagenesis provides a useful way to test these hypotheses.

6.2 Regulation

The regulation of PC activity *in vivo* occurs at three levels; transcription, substrate concentration and allosteric activation by acetyl-CoA. Transcription represents the long term control of PC levels in the cell (Angus *et al.*, 1981), and is governed by such factors as age, nutritional status and the stress state of the organism. The gene for PC from yeast has been isolated (Lim *et al.*, 1988), and this will enable the transcriptional control mechanisms in this species to be characterised. The partial characterisation of the rat PC gene, presented in chapter 3 provides appropriate probes for the future isolation and characterisation of the 5´ regulatory region of that gene. Such information should prove interesting, in that rat PC synthesis is observed to increase rapidly after birth and then gradually decline, reaching a constant level of total PC within 30 days (Yeung *et al.*, 1967; Chang, 1977).

While the mechanism of the allosteric activation of PC by acetyl-CoA is far from clear, the localisation of the acetyl-CoA binding site to a region of the tetrameric protein near the reaction centre does support the proposition of Clements (1977) that acetyl-CoA acts as a bridge between the two partial reaction sites. Although the finding of Nakashima *et al.* (1975) that rat PC is allosterically competent as a protomer suggests that the acetyl-CoA binding site is entirely within one monomer, the localisation of the acetyl-CoA binding site as a part of this study does not rule out contributions from more than one polypeptide. The possibility that acetyl-CoA binds between adjacent monomers does have some merit, in that it readily explains how the ligand is able to influence the stability of the tetrameric molecule, in addition to stimulating activity.

6.3 Evolution of the Biotin Carboxylases

With the advances in sequencing technology, the amino acid sequences of a number of biotin carboxylases have been determined. Comparisons between these sequences has provided much information regarding the catalytic and allosteric mechanisms as discussed above, but in addition, such comparisons provide compelling evidence for the evolutionary hypothesis of Lynen (1975). Clearly, the functional domains of the various

enzymes are related in structure, and most probably resulted from shuffling of genes for small individual components into large multifunction polypeptides by successive gene fusions. It remains to be seen however, whether strong evidence for gene fusion can be found. Such evidence would be expected to take the form of conserved exon/intron boundaries within the genes for eukaryotic enzymes of related function.

The finding that, upon proteolysis, eukaryotic PC can be dissected into at least two discrete portions or domains has interesting implications regarding the structure of PC from *Pseudomonas citronellolis*. This latter enzyme is composed of two separate subunits, a Mr 65,000 (α) biotin-containing subunit and a Mr 54,000 (β) biotin-free subunit (Cohen *et al.*, 1979). As Taylor *et al.* (1975) have shown that synthesis of the α and β subunits of *Ps. citronellolis* PC are coordinated, it seems likely that the protein will either be encoded by a single gene or an operon of two genes. In an analogous manner to the protease treatment reported here, the product of a single gene would then be cleaved either *in vivo* or as a result of the isolation procedure, to form the observed 2 subunit enzyme.

In an attempt to characterise the quaternary structure of *Ps. citronellolis* PC, Goss *et al.* (1981) found that in the intact enzyme the β subunit is apparently hidden from the enviroment by the α subunit (biotin-containing). Given that the overall structure of the enzyme has been observed to be indistinguishable from the yeast and vertebrate PCs (Fuchs *et al.*, 1988), it is hard to envisage a structure where one polypeptide engulfs another of a similar size. As shown in figure 6.1, two hypothetical structures can be drawn, based on the sequence comparisons discussed in chapters 3 & 5. Isolation and sequencing of the gene(s) for *Ps. citronellolis* PC may be expected to provide useful evidence for the overall structure of the enzyme.

6.4 The Direction of Future Work

With the application of molecular genetic techniques to the study of biotin carboxylases a whole new era of experimentation is set to begin. For example, Murtif and Samols (1987) have characterised the residues of transcarboxylsae 1.3S subunit that are important for recognition by the biotinylating enzyme holoenzyme synthetase\, by expression of various mutants in *E. coli*. In addition, expression of deletion mutants of TC 12S monomer gene (Wood *et al.*, 1990) has shown that a region between residues 50 and 60 is required for normal activity.

Observing the effect on activity of specific alterations to the amino acid sequence of PC, is expected to provide important information on the enzyme's mechanism. Before such experiments can be designed intelligently however, two further criteria must be met.

Firstly the availablity of a competent expression system (preferably in a homologous host), and secondly further information regarding which residues to target. As the gene for yeast PC has previously been isolated and characterised in this laboratory, the obvious choice for an expression system is a yeast strain lacking endogenous PC. Construction of such yeast strains is underway in this laboratory, but has been complicated by the discovery of a second, differentially expressed, yeast PC gene which must now also be isolated and deleted by homologous recombination (M. Walker, Adelaide, *pers. comm.*).

Information on which residues are likely to play a role in enzymic function, and therefore are worthy of further study, may be derived in two ways;

(1) classical affinity labelling/peptide isolation

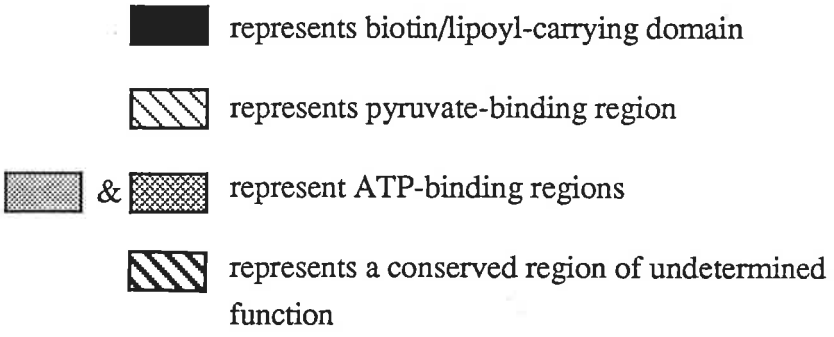
(2) comparison of the sequences of PCs from a wide range of species with sequences from related proteins.

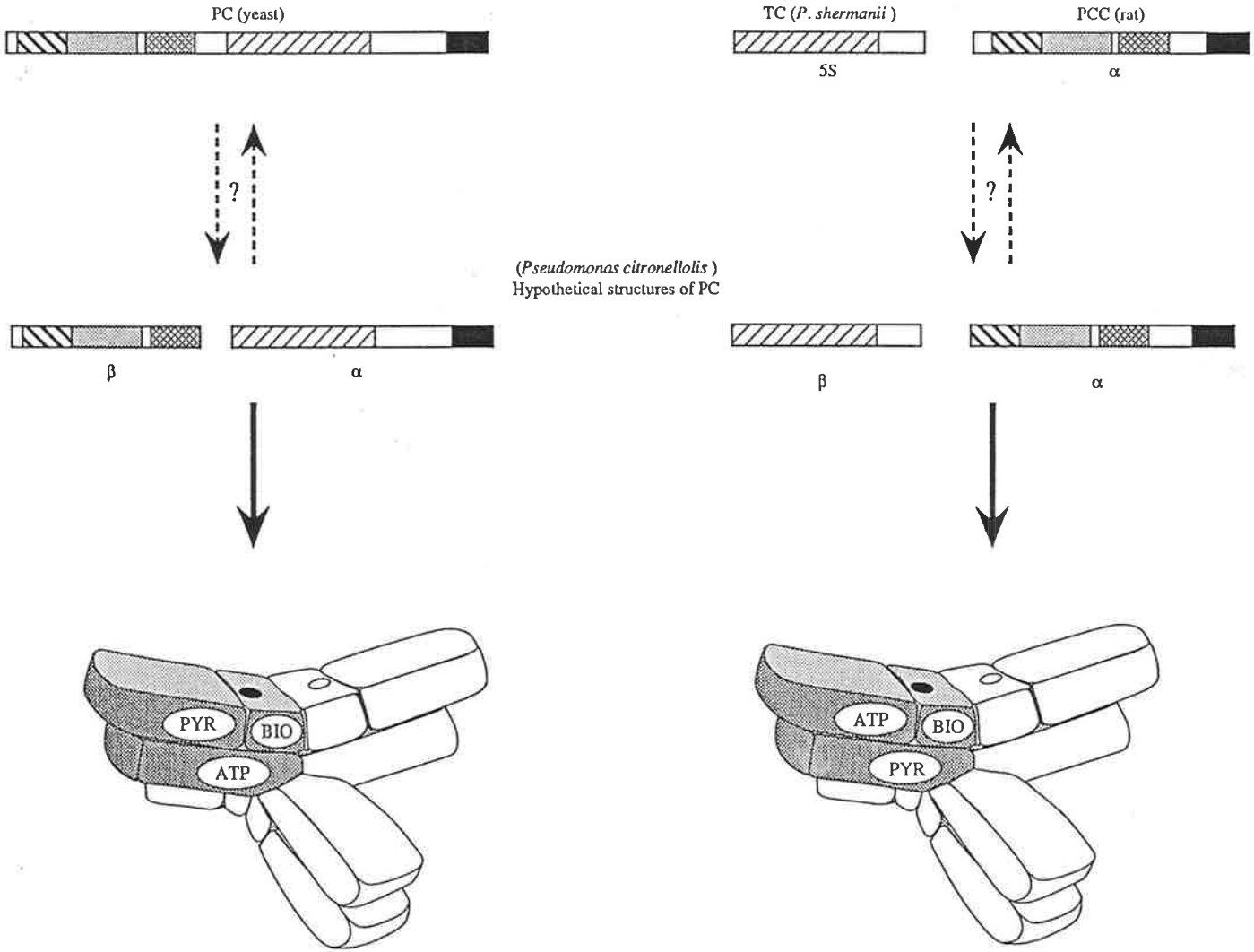
This latter option will be facilitated by on-going studies in this laboratory aimed at characterising the complete amino acid sequences of PCs from human, rat, *Pseudomonas* and yeast (gene II).

As the size of PC appears to limit its suitability for detailed structural analysis, a future line of investigation would be to attempt expression of fragments of the protein, such as individual domains, in sufficient quantities for analysis by 2-dimensional NMR or crystallography. In addition, the functional significance of regions of the protein such as the 1st and 2nd partial subsites could then be tested in isolation, by expression of those particular regions. The likelihood of successful expression of catalytically competent fragments of PC is enhanced by the observation that *E. coli* are able to express TC 1.3S and 12S subunits in an active form (Murtif and Samols, 1987; Wood *et al.*, 1990). The structures of the component domains of PC, once determined could be assembled to provide an overall structure, which combined with the mutagenesis experiments discussed above, may finally allow the relationship between the structure and function of PC to be solved.

Figure 6.1 A Hypothetical Scheme for the Evolution of Pyruvate Carboxylase from *Pseudomonas citrnellolis* and the Possible Subunit Structure of the Enzyme

Two hypothetical evolutionary lines for pyruvate carboxylase from *Pseudomonas citronellolis* are shown, along with the predicted quaternary structures of the proteins. The proposed functional domains of yeast pyruvate carboxylase, transcarboxylase 5S-subunit (*Propionibacterium shermanii*) and propionyl-CoA carboxylase α- subunit (rat) are also indicated (from figure 3.12). In each case the α subunit of *Pseudomonas citronellolis* pyruvate carboxylase is shown on the outer surface of the tetrahedral model (Goss *et al.*, 1981, Fuchs *et al.*, 1988).




PC (yeast)
TC (P. shermanii)
PCC (rat)
5S
α
?
?
(Pseudomonas citronellolis)
Hypothetical structures of PC
β
α
β
α
PYR
BIO
ATP
ATP
BIO
PYR

BIBLIOGRAPHY

Aebersold, R.H., Teplow, D.B., Hood, L.E. and Kent, S.B.H. (1986) *J. Biol. Chem.* **261**, 4229-4238, Electroblotting onto activated glass: High efficiency preparation of proteins from analytical sodium dodecyl sulphate-polyacrylamide gels for direct sequence analysis.

Ahmed, A. (1985) *Gene* **39**, 305-310, A rapid procedure for DNA sequencing using transposon-promoted deletions in *Escherichia coli.*

Allen, G. (1981) *Sequencing of proteins and peptides*, North-Holland Publishing Company Amsterdam.

Ambrose, M.C. and Perham, R.N. (1976) *Biochem. J.* **155**, 429-432, Spin-label study of the mobility of enzyme-bound lipoic acid in the pyruvate dehydrogenase multienzyme complex of *Eschericia coli.*

Anfinsen, C.B., Haber, E., Sela, M. and White, F.H. Jr. (1961) *Proc. Natl. Acad. Sci. USA.* **47**, 1309-1314, The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain.

Angus, C.W. and Lane, M.D. and Rosenfeld, P.J. and Kelly, T.J. (1981) *Biochem. Biophys. Res. Comm.* **103**, 1216-1222, Increase in translatable mRNA for mitochondrial pyruvate carboxylase during differentiation of 3T3 preadipocytes.

Ashman, L.K. and Keech, D.B. (1975) *J. Biol. Chem.* **250**, 14-21, Sheep kidney pyruvate carboxylase: Studies on the coupling of adenosine triphosphate hydrolysis and CO_2 fixation.

Ashman, L.K., Keech, D.B., Wallace, J.C. and Nielsen, J. (1972) *J. Biol. Chem.* **247**, 5818-5824, Sheep kidney pyruvate carboxylase: Studies on its activation by acetyl coenzyme A and characteristics of its acetyl coenzyme A independent activity.

Ashman, L.K., Wallace, J.C. and Keech, D.B. (1973) *Biochem. Biophys. Res. Comm.* **51**, 924-931, Desensitisation of pyruvate carboxylase against acetyl-CoA stimulation by chemical modification.

Attwood, P.V. and Cleland, W.W. (1986) *Biochemistry* **25**, 8191-8196, Decarboxylation of oxalacetate by pyruvate carboxylase.

Attwood, P.V. and Keech, D.B. (1984) *Curr. Top. Cell Regul.* **23**, 1-55, Pyruvate carboxylase.

Attwood, P.V., Wallace, J.C. and Keech, D.B. (1984) *Biochem. J.* **219**, 243-251, The carboxybiotin complex of pyruvate carboxylase.

Attwood, P.V., Mayer, F. and Wallace, J.C. (1986a) *FEBS Lett.* **203**, 191-196, Avidin as a probe of the conformational changes induced in pyruvate carboxylase by acetyl-CoA and pyruvate.

Attwood, P.V., Tipton, P.A. and Cleland, W.W. (1986b) *Biochemistry* **25**, 8197-8205, Carbon-13 and deuterium isotope effects on oxalacetate decarboxylation by pyruvate carboxylase.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1989) *Current protocols in molecular biology*. Wiley Interscience, New York, USA.

Bagley, C.J. (1982) *Reactive cysteines of pyruvate carboxylase*, BSc. Honours thesis, University of Adelaide, Adelaide, South Australia.

Bagley, C.J., Goss, N.H., Keech, D.B. and Wallace, J.C. (1983) *Proc. Aust. Biochem. Soc.* **15**, 1, Catalytic activity of sheep liver pyruvate carboxylase depends on one of two conformationally vicinal sulphydryl groups.

Bai, D.H., Pape, M.E., Lopez-Casillas, F., Luo, X.C.L., Dixon, J.E. and Kim, K.-H. (1986) *J. Biol. Chem.* **261**, 12395-12399, Molecular cloning of cDNA for acetyl-coenzyme A carboxylase.

Bai, D.H., Moon, T.-W., Lopez-Casillas, F., Andrews, P.C. and Kim, K.-H. (1986) *Eur. J. Biochem.* **182**, 239-245, Analysis of the biotin-binding site on acetyl-coenzyme A carboxylase from rat.

Barritt, G.J. (1985) Regulation of enzymic activity, In Keech, D.B. and Wallace, J.C., editors, *Pyruvate Carboxylase*, pages 141-178, CRC Series in Enzyme Biology, CRC Press, Boca Raton, USA.

Barritt, G.J. (1985) Disease states in man and other vertebrates. In Keech, D.B. and Wallace, J.C., editors, *Pyruvate Carboxylase*, pages 179-212, CRC Series in Enzyme Biology, CRC Press, Boca Raton, USA.

Bax, A. (1989) *Annu. Rev. Biochem.* **58**, 223-256, Two-dimensional NMR and protein structure.

Bello, J., Harker, D. and De Jarnette, E (1961) *J. Biol. Chem.* **236**, 1358-1359, Note on an X-ray diffraction investigation of reduced-reoxidised ribonuclease.

Birnboim, H.C. and Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513-1523, A rapid alkaline extraction procedure for screening recombinant plasmid DNA.

Blake, C.C.F., Koenig, D.F., Mair, G.A., North, A.C.T., Phillips, D.C. and Sarma, V.R. (1965) *Nature* **206**, 757-761, Structure of hen egg-white lysozyme: A three-dimensional fourier synthesis at 2Å resolution.

Bloxham, D.P., Ericsson, L.H., Titani, K., Walsh, K.A. and Neurath, H. (1980) *Biochemistry* **19**, 3979-3985, Limited proteolysis of pig heart citrate synthase by subtilisin, chymotrypsin and trypsin.

Booker, G.W., Oliver, K.J., Rohde, M. and Wallace, J.C. (1986) *Proc. Aust. Biochem. Soc.* **18**, 16, Localisation of reactive lysyl residues in sheep liver pyruvate carboxylase.

Booker, G.W., Khew-Goodall, Y.-S., Goodall, G.J. and Wallace, J.C. (1986) *Proc. Aust. Biochem. Soc.* **20**, SP26, Pyruvate carboxylase: Allosteric activation by acetyl-CoA (abstract).

Bradford, M.M. (1976) *Anal. Biochem.* **72**, 248-254, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.

Browner, M.F., Taroni, F., Sztul, E. and Rosenberg, L.E. (1989) *J. Biol. Chem.* **264**, 12680-12685, Sequence analysis, biogenesis and mitochondrial import of the a-subunit of rat liver propionyl-CoA carboxylase.

Buckel, W. and Semmler, R. (1982) *FEBS Lett.* **148**, 35-38, A biotin-dependent sodium pump: Glutaconyl-CoA decarboxylase from *Acidaminococcus fermentans*.

Carey, K.J. (1988) *Pyruvate Carboxylase: An Immunological Study*, PhD thesis, University of Adelaide, Adelaide, South Australia.

Carver, J.A., Baldwin, G.S., Keech, D.B., Bais, R. and Wallace, J.C. (1988) *Biochem. J.* **252**, 501-507, Inactivation of chicken liver pyruvate carboxylase by 1,10-phenanthroline.

Cassady, A.I. (1987) *Pyruvate Carboxylase: A Molecular Biological Study*, PhD thesis, University of Adelaide, Adelaide, South Australia.

Cech, T.R. (1983) *Cell* **34**, 713-716, RNA splicing: three themes with variations.

Cech, T.R., Tanner, N.K., Tinoco, I. Jr., Weir, B.R., Zucker, M. and Perlman, P.S. (1983) *Proc. Natl. Acad. Sci.* **34**, 3903-3907, Secondary structure of the *Tetrahymena* ribosomal RNA intervening sequence: Structural homology with fungal mitochondrial intervening sequences.

Chang, L.E. (1977) *Pediat. Res.* **11**, 6-8, The development of pyruvate carboxylase in rat mitochondria.

Chapman-Smith, A. (1981) *Studies on the Interaction of Sheep Liver Pyruvate Carboxylase with the Allosteric Effector, Acetyl-CoA.*, PhD thesis, University of Adelaide, Adelaide, South Australia.

Clements, P.R. (1977) *Pyruvate Carboxylase: The Mode of Action of Acetyl-CoA.*, PhD thesis, University of Adelaide, Adelaide, South Australia.

Coffee, C.J., Bradshaw, R.A., Goldin, B.R. and Frieden, C. (1971) *Biochemistry* **10**, 3516-3526, Identification of the sites of modification of bovine liver glutamate dehydrogenase reacted with trinitrobenzene sulphonate.

Cohen, N.D., Duc, J.A., Beegen, H. and Utter, M.F. (1979) *J. Biol. Chem.* **254**, 9262-9269, Quaternary structure of pyruvate carboxylase from *Pseudomonas citronellolis*.

Creighton, T.E. (1977) *J. Mol. Biol.* **113**, 275-293, Conformational restrictions on the pathway of folding and unfolding of the pancreatic trypsin inhibitor.

Dawson, M.R.C., Elliott, D.C., Elliott, W.H. and Jones, K.M. (1969) *Data for Biochemical Research.* Oxford University Press, 2nd edition, p192.

Day, P.J., Lewendon, A. and Shaw, W.V. (1988) *Biochem. Soc. Trans.* **16**, 715-716, Coenzyme-A-binding site of chloramphenicol acetyltransferase.

Dayhoff, M.O., Schwartz, R.M. and Orcutt, B.C. (1978) *Atlas of Protein Sequence and Structure,* National Biomedical Research Foundation, Washington, D.C., USA 5th Edition.

Devereux, J., Haeberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387-395, A comprehensive set of sequence analysis programs for the VAX.

DeVivio, D.C., Haymond, M.W., Leckie, M.P., Bussman, Y.L., McDougal, D.B. and Pagliari, A.S. (1977) *J. Clin. Endocrinol. Metab.* **45**, 1281-1296, The clinical and biochemical implications of pyruvate carboxylase deficiency.

Dimroth, P. (1980) *FEBS Lett.* **122**, 234-236, A new sodium-transport system energised by the decarboxylation of oxaloacetate.

Dimroth, P. (1985) *Ann. N. Y. Acad. Sci.* **447**, 72-85, Biotin-dependent decarboxylases as energy transducing systems.

Duggleby, R.G., Kinns, H. and Rood, J.I. (1981) *Anal. Biochem.* **110**, 49-55, A computer program for determining the size of DNA restriction fragments.

Easterbrook,-Smith, S.B., Campbell, A.J., Keech, D.B. and Wallace, J.C. (1979) *Biochem. J.* **179**, 497-502, The atypical velocity response by pyruvate carboxylase to increasing concentrations of acetyl-Coenzyme A.

Easterbrook-Smith, S.B., Hudson, P.J., Goss, N.H., Keech, D.B. and Wallace, J.C. (1976) *Arch. Biochem. Biophys.* **176**, 709-720, Pyruvate carboxylase: Mechanisms of the second partial reaction.

Easterbrook-Smith, S.B., Wallace, J.C. and Keech, D.B. (1976b) *Eur. J. Biochem.* **62**, 125-130, Pyruvate carboxylase: Affinity labelling of the magnesium adenosine triphosphate binding site.

Eklund, H., Nordström, B., Zeppezauer, E., Söderlund, G., Ohlsson, I., Boiwe, T., Söderberg, B.-O., Tapia, O. and Brändén, C.-I. (1976) *J. Mol. Biol.* **102**, 27-59, Three-dimensional structure of horse liver alcohol dehydrogenase at 2.4 Å resolution.

Eriksson, A.E., Jones, T.W. abd Liljas, A. (1988) Proteins: Struc. Func. Genet. 4, 274-282, Refined structure of human carbonic anhydrase II at 2.0 Å resolution.

Fersht, A. (1985) *Enzyme structure and mechanism*, W.H. Freeman and company, New York 2nd edition.

Frey, W.H. and Utter, M.F. (1977) *J. Biol. Chem.* **252**, 51-56, Binding of acetyl CoA to chicken liver pyruvate carboxylase.

Freytag, S.O. and Collier, K.J. (1984) *J. Biol. Chem.* **259**, 12831-12837, Molecular cloning of a cDNA for human pyruvate carboxylase: Structural relationship to other biotin-containing carboxylases and regulation of mRNA content in differentiating preadipocytes.

Fuchs, J., Johannssen, W., Rohde, M. and Mayer, F. (1988) *FEBS Lett.* **231**, 102-106, Pyruvate carboxylase from *Pseudomonas citronellolis*: Shape of the enzyme, and localisation of its prosthetic biotin group by electron microscopic affinity labelling.

Fung, C.H. (1971) *The Regulatory Properties and Reaction Mechanism of Chicken Liver Pyruvate Carboxylase*, PhD thesis, Case Western Reserve University, Clevland, USA.

Fung, C.H., Mildvan, A.S. and Leigh, J.S. Jr. (1974) *Biochemistry* **3**, 1160-1169, Electron and nuclear magnetic studies of the interaction of pyruvate with transcarboxylase.

Fung, C.H., Mildvan, A.S., Allerhand, A., Komoroski, R. and Scrutton, M.C. (1973) *Biochemistry* **12**, 620-629, Interaction of pyruvate with pyruvate carboxylase and pyruvate kinase as studied by paramagnetic effects on ^{13}C relaxation times.

Garnier, J., Osguthorrpe, D.J. and Robson, B. (1978) *J. Mol. Biol.* **120**, 97-120, Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins.

Gatley, S.J. (1977) *Fed. Proc. Am. Soc. Exp. Biol.* **36**, 672, Inhibition by benzoic-acid of ureogenesis and gluconeogenesis in isolated rat hepatocytes (abstract).

Gitlin, G., Bayer, E.A. and Wilchek, M. (1988) *Biochem. J.* **250**, 291-294, Studies on the biotin-binding site of avidin:Tryptophan residues involved in the active site.

Goad, W.B. and Kanehisa, M. (1982) *Nucleic Acids Res.* **10**, 247-263, pattern recognition in nucleic acid sequences: A general method for finding local homologies and symmetries.

Goodall, G.J., Baldwin, G.S., Wallace, J.C. and Keech, D.B. (1981) *Biochem. J.* **199**, 603-609, Factors that influence the translocation of the N-carboxybiotin moiety between the two sub-sites of pyruvate carboxylase.

Goodall, G.J., Prager, R., Wallace, J.C. and Keech, D.B. (1983) *FEBS Lett.* **163**, 6-9, A mechanism for the transfer of the carboxyl group from 1´-N-carboxybiotin to acceptor substrates by biotin-containing enzymes.

Goss, J.A., Cohen, N.D. and Utter, M.F. (1981) *J. Biol. Chem.* **256**, 11819-11825, Characterisation of the subunit structure of pyruvate carboxylase from *Pseudomonas citronellolis*.

Goss, N.H., Dyer, P.Y., Keech, D.B. and Wallace, J.C. (1979) *J. Biol. Chem.* **254**, 1734-1739, An electron microscopic study of pyruvate carboxylase.

Grande, H.J., Van Telgen, H.J. and Veeger, C. (1976) *Eur. J. Biochem.* **71**, 509-518, Symmetry and asymmetry of the pyruvate dehydrogenase complexes from *Azotobacter vinelandii* and *Eschericia coli* as reflected by fluorescence and spin-label studies.

Green, N.M., (1963) *Biochem. J.* **89**, 585-591, Avidin: The use of [^{14}C] biotin for kinetic studies and for assay.

Green, N.M., Konieczny, L., Toms, E.J. and Valentine, R.C. (1971) *Biochem. J.* **125**, 781-791, The use of bifunctional biotinyl compounds to determine the arrangement of subunits in avidin.

Griffith, A.D., Cyr, D.M., Egan, S.G. and Tremblay, G.C. (1989) *Arch. Biochem. Biophys.* **269**, 201-207, Inhibition of pyruvate carboxylase by sequestration of coenzyme-A with sodium benzoate.

Guest, J.R., Lewis, H.M., Graham, L.D., Packman, L.C. and Perham, R.N. (1985) *J. Mol. Biol.* **185**, 743-754, Genetic reconstruction and functional analysis of the repeating lipoyl domains in pyruvate dehydrogenase multienzyme complex of *Escherichia coli.*

Haase, F.C., Henrikson, K.P., Treble, D.H. and Allen, S.H.G. (1982) *J. Biol. Chem.* **257**, 11994-11999, The subunit structure and function of the propionyl coenzyme A carboxylase of *Mycobacterium smegmatis.*

Haase, F.C., Beegen, H. and Allen, S.H.G. (1984) *Eur. J. Biochem.* **140**, 147-151, Propionyl coenzyme A carboxylase of *Mycobacterium smegmatis* : An electron microscopic study.

Haberland, M.E., Chen, C.-W. and Smith, E.L. (1980) *J. Biol. Chem.* **255**, 7993-8000, NAD-specific glutamate dehydrogenase of *Neurospora crassa*: Limited action of trypsin and the presence of two distinct domains.

Hartl, F-U, Pfanner, N., Nicholson, D.W. and Neupert, W. (1989) *Biochim. Biophys. Acta* **988**, 1-45, Mitochondrial protein import.

Helmerhorst, E. and Stokes, G.B. (1980) *Anal. Biochem.* **104**, 130-135, Microcentrifuge desalting: A rapid quantitative method for desalting small amounts of protein.

Henrickson, K.P., Allen, S.H.G. and Maloy, W.L. (1979) *Anal. Biochem.* **94**, 366-370, An avidin monomer affinity column for the purification of biotin-containing enzymes.

Hill, R.L., Delaney, R., Fellows, R.E. Jr. and Lebovitz, H.E. (1966) *Proc. Natl. Acad. Sci. USA.* **56**, 1762-1769, The evolutionary origins of the immunoglobins.

Hilpert, W. and Dimroth, P. (1982) *Nature* **296**, 584-585, Conversion of the chemical energy of methylmalonyl-CoA decarboxylation into a Na^+ gradient.

Holzman, T.F., Brems, D.N. and Dougherty, J.J. Jr. (1986) *Biochemistry* **25**, 6907-6917, Reoxidation of reduced bovine growth hormone from a stable secondary structure.

Hudson, P.J., Keech, D.B. and Wallace, J.C. (1975) *Biochem. Biophys. Res. Comm.* **65**, 213-219, Pyruvate carboxylase: Affinity labelling of the pyruvate binding site.

Hunkapillar, M.W. and Hewick, R.M. and Dreyer, W.J. and Hood, L.E. (1983) *Methods Enzymol.* **91**, 399-413, High-sensitivity sequencing with a gas-phase sequenator.

Irias, J.J., Olmsted, M.R. and Utter, M.F. (1969) *Biochemistry* **8**, 5136-5148, Pyruvate carboxylase: Reversible inactivation by cold.

Jansen, R., Kalousek, F., Fenton, W.A., Rosenberg, L.E. and Ledley, F.D. (1989) *Genomics* **4**, 198-205, Cloning of full-length methylmalonyl-CoA mutase from a cDNA library using the polymerase chain reaction.

Johannssen, W., Attwood, P.V., Wallace, J.C. and Keech, D.B. (1983) *Eur. J. Biochem.* **133**, 201-206, Localisation of the active site of pyruvate carboxylase by electron microscopic examination of avidin-enzyme complexes.

Keech, D.B. and Attwood, P.V. (1985) The reaction mechanism, In Keech, D.B. and Wallace, J.C., editors, *Pyruvate Carboxylase*, pages 109-140, CRC Series in Enzyme Biology, CRC Press, Boca Raton, USA.

Keech, D.B. and Farrant, R.K. (1968) *Biochim. Biophys. Acta* **151**, 493-503, The reactive lysyl residue at the allosteric site of sheep kidney pyruvate carboxylase.

Keech, D.B. and Utter, M.F. (1963) *J. Biol. Chem.* **238**, 2609-2614, Pyruvate carboxylase: Properties.

Kester, W.R. and Matthews, B. W. (1977) *J. Biol. Chem.* **252**, 7704-7710, Comparison of the structures of carboxypeptidase A and thermolysin.

Khew-Goodall, Y.-S. (1985) *Pyruvate Carboxylase: Its Interactions with Acetyl CoA.*, PhD thesis, University of Adelaide, Adelaide, South Australia.

King, T.P., Li, Y. and Kochoumian, L. (1978) *Biochemistry* **17**, 1499-1506, Preparation of protein conjugates via intermolecular disulfide bond formation.

Knowles, J.R. (1989) *Annu. Rev. Biochem.* **58**, 195-221, The mechanism of biotin-dependent enzymes.

Knudsen, J. and Grunnet, I. (1982) *Biochem. J.* **202**, 139-143, Transacylation as a chain-termination mechanism in fatty acid synthesis by mammalian fatty acid synthetase.

Kornberg, H.L. (1966) Anaplerotic sequences and their role in metabolism, In Campbell, P.N. and Greville, G.D., editors, *Essays in Biochemistry*, page 1, Academic Press, London, 2nd edition.

Kraus, J.P., Firgaira, F., Novotny, J., Kalousek, F., Williams, K.R., Williamson, C., Ohura, T. and Rosenberg, L.E. (1986) *Proc. Natl. Acad. Sci. USA.* **83**, 8049-8053, Coding sequence of the precursor of the beta subunit of rat propionyl-CoA carboxylase.

Krause, K.L., Volz, K.W. and Lipscomb, W.N. (1987) *J. Mol. Biol.* **193**, 527-553, 2.5 Å structure of aspartate carbamoyltransferase complexed with the bisubstrate analog N-(phosphonacetyl) -L-aspartate.

Krebs, H.A. (1954) *Bull. Johns Hopkins Hosp.* **95**, 19, Considerations concerning the pathways of synthesis in living matter.

Kumar, G. K., Beegen, H. and Wood, H.G. (1988a) *Biochemistry* **27**, 5972-5978, Involvement of tryptophans at the catalytic and subunit-binding domains of transcarboxylase.

Kumar, G. K., Haase, F.C., Phillips, N.B.F. and Wood, H.G. (1988b) *Biochemistry* **27**, 5978-5983, Involvement and identification of a tryptophanyl residue at the pyruvate binding site of transcarboxylase.

Laemmli, U.K., 1970, *Nature* **227**, 680-685, Cleavage of structural proteins during the assembly of the head of bacteriophage T4.

Lamhonwah, A.-M., Barankiewicz, T.J., Willard, H.F., Mahuran, D.J., Quan, F. and Gravel, R.A. (1986) *Proc. Natl. Acad. Sci. USA.* **83**, 4864-4868, Isolation of cDNA clones coding for the α and β chains of human propionyl-CoA carboxylase: Chromosomal assignments and DNA polymorphisms associated with PCCA and PCCB genes.

Lamhonwah, A.-M., Mahuran, D. and Gravel, R.A. (1989) *Nucleic Acids Res.* **17**, 4396, Human mitochondrial propionyl-CoA carboxylase: Localisation of the N-terminus of the pro- and mature α chains in the deduced primary sequence of a full-length cDNA.

Lamhonwah, A.-M., Quan, F. and Gravel, R.A. (1987) *Arch. Biochem. Biophys.* **254**, 631-636, Sequence homology around the biotin-binding site of human propionyl-CoA carboxylase and pyruvate carboxylase.

Lane, M.D., Moss, J. and Polakis, S.E. (1974) *Curr. Top. Cell Regul.* **8**, 139-195, Acetyl coenzyme A carboxylase.

Laußermair, E., Schwarz, E., Oesterhelt, D., Reinke, H., Beyreuther, K and Dimroth, P. (1989) *J. Biol. Chem.* **264**, 14710-14715, The sodium ion translocating oxaloacetate decarboxylase of Klebsiella pneumoniae: Sequence of the integral membrane-bound subunits β and γ.

Lau, E.P., Haley, B.E. and Barden, R.E. (1977) *Biochem. Biophys. Res. Comm.* **76**, 843-849, The 8-azidoadenine analog of S-Benzoyl (3'-dephospho) coenzyme A: A photoaffinity label for acyl CoA:glycine N-acyltransferase.

Layne, E. (1957) *Meth. Enzymol.* **3**, 447-454, Spectrophotometric and turbidimetric methods for measuring proteins.

Leslie, A.G.W., Moody, P.C.E. and Shaw, W.V. (1988) *Proc. Natl. Acad. Sci. USA.* **85**, 4133-4137, Structure of chloramphenicol acetyltransferase at 1.75-Å resolution.

Levy, H.M., Leber, P.D. and Ryan, E.M. (1963) *J. Biol. Chem.* **238**, 3654-3659, Inactivation of myosin by 2,4-dinitrophenol and protection by adenosine triphosphate and other phosphate compounds.

Libor, S.M., Sundaram, T.K. and Scrutton, M.C. (1978) *Biochem. J.* **169**, 543-558, Pyruvate carboxylase from a thermophilic Bacillus: Studies on the specificity of activation by acyl derivatives of coenzyme A and on the properties of catalysis in the absence of activator.

Libor, S.M., Sundaram, T.K., Warwick, R., Chapman, J.A. and Grundy, S.M.W. (1979) *Biochemistry* **18**, 3647-3653, Pyruvate carboxylase from a thermophilic Bacillus: Some molecular characteristics.

Lim, F. (1988) *Pyruvate Carboxylase: The Protein & Gene in Yeast*, PhD thesis, University of Adelaide, Adelaide, South Australia.

Lim, F., Morris, C.P., Occhiodoro, F. and Wallace, J.C. (1988) *J. Biol. Chem.* **263**, 11493-11497, Sequence and domain structure of yeast pyruvate carboxylase.

Lopez-Casillas, F., Bai, D.H., Luo, X.C.L., Kong, I.-S., Hermodson, M.A. and Kim, K.-H. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 5784-5788, Structure of the coding sequence and primary amino acid sequence of acetyl-coenzyme A carboxylase.

Lorber, V., Lifsen, N., Wood, H.G., Sakami, W. and Shreeve, W.W. (1950) *J. Biol. Chem.* **183**, 517-529, Conversion of lactate to liver glycogen in the intact rat, studied with isotopic lactate.

Luo, X., Park, K., Lopez-Casillas, F. and Kim, K.-H. (1990) *J. Biol. Chem.* **86**, 4042-4046, Sructural features of the acetyl-CoA carboxylase gene: Mechanisms for the generation of mRNAs with 5′ end heterogeneity.

Lusty, C.J., Widgren, E.E., Broglie, K.E. and Nyunoya, H. (1983) *J. Biol. Chem.* **258**, 14466-14472, Yeast carbamyl phosphate synthetase: Structure of the yeast gene and homology to Eschericia cilo carbamyl phosphate synthetase.

Lynen, F. (1975) Structures of biotin enzymes in *Lipmann Symposium:Energy, regulation & biosynthesis in Molecular Biology*, D. Richter, editor, 671-698.

Maloy, W.L., Bowien, B.U., Zwolinski, G.K., Kumar, K.G., Wood, H.G., Ericsson, L.H. and Walsh, K.A. (1979) *J. Biol. Chem.* **254**, 11615-11622, Amino acid sequence of the biotinyl subunit from transcarboxylase.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning*, Cold Spring Harbor Laboratory,Cold Spring Harbor, NY, USA. (A Laboratory Manual).

Marks, C.B., Naderi, H., Kosen, P.A., Kuntz, I.D. and Anderson, S. (1987) *Science* **235**, 1370-1373, Mutants of bovine pancreatic trypsin inhibitor lacking cysteines 14 and 38 can fold properly.

Mauro, J.M., Lewis, R.V. and Barden, R.E. (1986) *Biochem. J.* **237**, 533-540, Photoaffinity labelling of carnitine acetyltransferase with S-(p-azidophenacyl) thiocarnitine.

Mayer, F., Wallace, J.C. and Keech, D.B. (1980) *Eur. J. Biochem.* **112**, 265-272, Further electron microscopic studies on pyruvate carboxylase.

Mayer, F., Osmani, S.A. and Scrutton, M.C. (1986) *FEBS Lett.* **192**, 2125-219, Pyruvate carboxylase from Rhizopus arrhizus: Analysis of the subunit structure by electron microscopy.

McClure, W.R., Lardy, H.A. and Kneifel, H.P. (1971) *J. Biol. Chem.* **246**, 3569-3578, Rat liver pyruvate carboxylase: Preparation, properties and cation specificity.

Messing, J. and Vieira, J. (1982) *Gene* **19**, 269-276, A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments.

Mildvan, A.S., Scrutton, M.C. and Utter, M.F. (1966) *J. Biol. Chem.* **241**, 3488-3498, Pyruvate carboxylase: A possible role for tightly bound manganese.

Mornet, D. and Ue, K. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 3680-3684, Proteolysis and structure of skeletal muscle actin.

Morris, C.J. and Reeve, J.N. (1988) *J. Bacteriol.* **170**, 3125-3130, Conservation of structure in the human gene encoding argininosuccinate synthase and the argG genes of the archaebacteria *Methanosarcina barkeri* MS and *Methanococcus vannielli.*

Murtif, V.L. and Samols, D. (1987) *J. Biol. Chem.* **262**, 11813-11816, Mutagenesis affecting the carboxyl terminus of the biotinyl subunit of transcarboxylase: Effects on biotination.

Musik, W.D.L. and Rossmann, M.G. (1979) *J. Biol. Chem.* **254**, 7611-7620, The structure of mouse testicular lactate dehydrogenase isoenzyme C_4 at 2.9Å resolution.

Myers, D.E., Tolbert, B. and Utter, M.F. (1983) *Biochemistry* **22**, 5090-5096, Activation of yeast pyruvate carboxylase: Interactions between acyl coenzyme A compounds, aspartate, and substrates of the reaction.

Nakashima, K., Rudolph, F.B., Wakabayashi, T. and Lardy, H.A. (1975) *J. Biol. Chem.* **250**, 331-336, Rat liver pyruvate carboxylase: Reversible dissociation by chloride salts of monovalent cations.

Nyunoya, H. and Lusty, C.J. (1984) *J. Biol. Chem.* **259**, 9790-9798, Sequence of the small subunit of yeast carbamyl phosphate synthetase and identification of its catalytic domain.

Nyunoya, H. and Lusty, C.J. (1983) *Proc. Natl. Acad. Sci. USA.* **80**, 4629-4633, The carB gene of *Eschericia coli*: A duplicated gene coding for the large subunit of carbamoyl-phosphate synthetase.

Nyunoya, H., Broglie, K.E., Widgren, E.E. and Lusty, C.J. (1985) *J. Biol. Chem.* **260**, 9346-9356, Characterisation and derivation of the gene coding for mitochondrial carbamyl phosphate synthetase of rat.

Ogita, T and Knowles, J.R. (1988) *Biochemistry* **27**, 8028-8033, On the intermediacy of carboxyphosphate in biotin-dependent carboxylations.

Oizumi, J., Shaw, K.N.F., Giudici, T.A., Carter, M., Donnell, G.N. and Ng, W.G. (1983) *J. Inherited Metab. Dis.* **6**, 89-93, Pyruvate carboxylase deficiency with renal tubular acidosis and cystinuria.

Osmani, S.A. and Scrutton, M.C. (1981) *FEBS Lett.* **135**, 253-256, Activation of pyruvate carboxylase from Aspergillus nidulans by acetyl coenzyme A.

Osmani, S.A., Mayer, F., Marston, F.A.O., Selmes, I.P. and Scrutton, M.C. (1984) *Eur. J. Biochem.* **139**, 509-518, Pyruvate carboxylase from Aspergillus nidulans: Effects of regulatory modifiers on the structure of the enzyme.

Osmani, S.A., Scrutton, M.C. and Mayer, F. (1985) *Ann. N. Y. Acad. Sci.* **447**, 56-71, The structure and regulation of Fungal pyruvate carboxylase.

Palacian, E. and Neet, K.E. (1972) *Biochem. Biophys. Res. Comm.* **276**, 297-312, Pyruvate carboxylase:Inactivation of sulfhydryl-group reagents in the presence of certain inorganic anions, substrates, and modifiers and the effects on the allosteric properties of the enzyme.

Penberthy, (1970) *Studies on Pyruvate Carboxylase*, BSc. Honours thesis, University of Adelaide, Adelaide, South Australia.

Pfanner, N., Hartl, F-U, and Neupert, W. (1988) *Eur. J. Biochem.* **175**, 205-212, Import of proteins into mitochondria: A multi-step process.

Polakis, S.E., Guchhait, R.B., Zwergel, E.E., Lane, M.D. and Cooper, T.G. (1974) *J. Biol. Chem.* **249**, 6657-6667, Acetyl coenzyme A carboxylase system of *Escherichia coli:* Studies on the mechanisms of the biotin carboxylase- and carboxyltransferase-catalysed reactions.

Porter, R.R. (1959) *Biochem. J.* **73**, 119-127, The hydrolysis of rabbit γ-globulin and antibodies with crystalline papain.

Poto, E.M., Wood, H.G., Barden, R.E. and Lau, E.P. (1978) *J. Biol. Chem.* **253**, 2979-2983, Photoaffinity labeling and stoichiometry of the coenzyme A ester sites of transcarboxylase.

Reed, G.H. and Scrutton, M.C. (1974) *J. Biol. Chem.* **249**, 6156-6162, Pyruvate carboxylase from chicken liver: Magnetic resonance studies of the effect of substrates and inhibitors on the enviroment of the bound manganese.

Reeke, G.N., Hartsuck, J.A., Ludwig, M.L., Quiocho, F.A., Steitz, T.A. and Lipscomb, W.N. (1967) *Proc. Natl. Acad. Sci. USA.* **58**, 2220-2226, The structure of carboxypeptidase A, VI. Some results at 2.0-Å resolution, and the complex with glycyl-tyrosine at 2.8-Å resolution.

Remington, S., Wiegand,G. and Huber, R. (1982) *J. Mol. Biol.* **158**, 111-152, Crystallographic refinement and atomic models of two different forms of citrate synthase at 2.7 and 1.7 Å resolution.

Retey, J. and Lynen, F. (1965) *Biochem. Z.* **342**, 256-271, Zur biochemishen function des biotins. IX. Der sterische verhauf der carboxylierung von propionyl CoA.

Rohde, M., Lim, F. and Wallace, J.C. (1986) *Eur. J. Biochem.* **156**, 15-22, Pyruvate carboxylase from *Saccharomyces cerevisiae*. Quaternary structure, effects of allosteric ligands and binding of avidin.

Rossmann, M.G. and Argos, P. (1981) *Annu. Rev. Biochem.* **50**, 497-532, Protein folding.

Rylatt, D.B. and Parish, C.R. (1982) *Anal. Biochem.* **121**, 213-214, Protein determination on an automatic spectrophotometer.

Rylatt, D.B., Keech, D.B. and Wallace, J.C. (1977) *Arch. Biochem. Biophys.* **183**, 113-122, Pyruvate carboxylase: Isolation of the biotin-containing tryptic peptide and the determination of its primary structure.

Safer, D., Hainfeld, J., Wall, J.S. and Reardon, J.F. (1982) *Science* **218**, 290-291, Biospecific labeling with undecagold:visualization of the biotin-binding site on avidin.

Samols, D., Thornton, C.G., Murtif, V.L., Kumar, G.K., Haase, C. and Wood, H.G. (1988) *J. Biol. Chem.* **263**, 6461-6464, Evolutionary conservation among biotin enzymes.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA.* **74**, 5463-5467, DNA sequencing with chain terminating inhibitors.

Schwarz, E., Oesterhelt, D., Reinke, H., Beyreuther, K. and Dimroth, P. (1988) *J. Biol. Chem.* **263**, 9640-9645, The sodium ion translocating oxalacetate decarboxylase of *Klebsiella pneumoniae*: Sequence of the biotin-containing a-subunit and relationship to other biotin-containing enzymes.

Scrutton, M.C. and Utter, M.F. (1965) *J. Biol. Chem.* **240**, 3714-3723, Pyruvate carboxylase: Interaction of the enzyme with adenosine triphosphate.

Scrutton, M.C. and Utter, M.F. (1967) *J. Biol. Chem.* **242**, 1723-1735, Pyruvate carboxylase: Some properties of the activation by certain acyl derivatives of coenzyme-A.

Scrutton, M.C. and White, M.D. (1973) *J. Biol. Chem.* **248**, 5541-5544, Pyruvate carboxylase: Specific inactivation of acetyl coenzyme A-dependent oxaloacetate synthesis during modification of the enzyme by trinitrobenzene sulphonate.

Scrutton, M.C., Keech, D.B. and Utter, M.F. (1965) *J. Biol. Chem.* **240**, 574-581, Pyruvate carboxylase: Partial reactions and the locus of activation by acetyl-CoA.

Scrutton, M.C., Pearce, P.H. and Fatebene, F. (1977) *Eur. J. Biochem.* **76**, 219-231, Pyruvate carboxylase: Effect of reaction components and analogues of acetyl-Coenzyme A on the rate of inactivation in the presence and absence of trinitrobenzene sulphonate.

Simon E.J. and Shemin, D. (1953) *J. Amer. Chem. Soc.* **75**, 2520-2523.

Singer, S.J. (1967) *Adv. Prot. Chem.* **22**, 1-54, Covalent labeling of active sites.

Solomon, A.K., Vennesland, B., Klemperer, F.W., Buchanan, J.M. and Hastings, A.B. (1941) *J. Biol. Chem.* **140**, 171-182, The participation of carbon dioxide in the carbohydrate cycle.

Srivastava, G. and Borthwick, I.A. and Brooker, J.D. and Wallace, J.C. and May, B.K. and Elliott, W.H. (1983) *Biochem. Biophys. Res. Comm.* **117**, 344-349, Hemin inhibits transfer of pre-δ-aminolevulinate synthase into chick embryo liver mitochondria.

Staden, R. (1982) *Nucleic Acids Res.* **10**, 2951-2961, An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences.

Sutton, M.R., Fall, R.R., Nervi, A.M., Alberts, A.W., Vagelos, P.R. and Bradshaw, R.A. (1977) *J. Biol. Chem.* **252**, 3934-3940, Amino acid sequence of *Escherichia coli* biotin carboxyl carrier protein.

Svoboda, M., Meuris, S., Robyn, C. and Christophe, J. (1985) *Anal. Biochem.* **151**, 16-23, Rapid electrotransfer of proteins from polyacrylamide gels to nitrocellulose membrane using surface-conductive glass as anode.

Tahara, T., Kraus, J.P. and Rosenberg, L.E. (1990) *Proc. Natl. Acad. Sci. USA.* **87**, 1372-1376, An unusual insertion/deletion in the gene encoding the β-subunit of propionyl-CoA carboxylase is a frequent mutation in caucasian propionic acidemia.

Takai, T. and Yokoyama, C. and Wada, K. and Tanabe, T. (1988) *J. Biol. Chem.* **263**, 2651-2657, Primary structure of chicken liver acetyl-CoA carboxylase deduced from cDNA sequence.

Taniuchi, H. (1970) *J. Biol. Chem.* **245**, 5459-5468, Formation of randomly paired disulfide bonds in des-(121-124) -ribonuclease after reduction and reoxidation.

Taylor, B.L and Frey, W.H. and Barden, R.E. and Scrutton, M.C. and Utter, M.F. (1978) *J. Biol. Chem.* **253**, 3062-3069, The use of the ultracentrifuge to determine the catalytically competent forms of enzymes with more than one oligomeric structure.

Thampy, K.G., Huang, W.-Y. and Wakil, S.J. (1988) *Arch. Biochem. Biophys.* **266**, 270-276, A rapid purification method for rat liver pyruvate carboxylase and amino acid sequence analysis of NH_2-terminal and biotin peptide.

Thornton, C., Samols, D., Kumar, G.K., Haase, C. and Wood, H.G. (1987) *Fed. Proc.* **46**, 2144, Sequence and expression of the *Propionibacterium shermanii* transcarboxylase operon (abstract).

Tipton, P.A. and Cleland, W.W. (1988) *Biochemistry* **27**, 4317-43235, Catalytic mechanism of biotin carboxylase: Steady-state kinetic investigations.

Topper, Y.J. and Hastings, A.B. (1949) *J. Biol. Chem.* **179**, 1255-1264, A study of the chemical origins of glycogen by use of ^{14}C-labeled carbon dioxide, acetate and pyruvate.

Traut, T.W. (1988) *Proc. Natl. Acad. Sci. USA.* **85**, 2944-2948, Do exons code for structural or functional units in proteins?.

Utter, M.F. (1959) *Ann. N.Y. Acad. Sci.* **72**, 451-461, The role of CO_2 fixation in carbohydrate utilization and synthesis.

Utter, M.F. and Keech, D.B. (1960) *J. Biol. Chem.* **235**, PC17-PC18, Formation of oxaloacetate from pyruvate and CO_2.

Utter, M.F. and Keech, D.B. (1963) *J. Biol. Chem.* **238**, 2603-2608, Pyruvate carboxylase: Nature of the reaction.

Utter, M.F. and Kurahashi, K. (1954a) *J. Biol. Chem.* **207**, 787-802, Purification of oxaloacetate carboxylase from chicken liver.

Utter, M.F. and Kurahashi, K. (1954b) *J. Biol. Chem.* **207**, 821-841, Mechanism of action of oxaloacetate carboxylase.

Utter, M.F., Barden, R.E. and Taylor, B.L. (1975) *Adv. Enzymol.* **42**, 1-72, Pyruvate carboxylase: an evaluation of the relationship between structure and mechanism and between structure and catalytic activity.

Valentine, R.C. and Wrigley, N.G. and Scrutton, M.C. and Irias, J.J. and Utter, M.F. (1966) *Biochemistry* **5**, 3111-3116, Pyruvate carboxylase. VIII. The subunit structure as examined by electron microscopy.

Wallace, J.C. (1985) Distribution and biological functions of pyruvate carboxylase in nature, In Keech, D.B. and Wallace, J.C., editors, *Pyruvate Carboxylase*, pages 5-64, CRC Series in Enzyme Biology, CRC Press, Boca Raton, USA.

Wallace, J.C. and Easterbrook-Smith, S.B. (1985) The structure of pyruvate carboxylase, In Keech, D.B. and Wallace, J.C., editors, *Pyruvate Carboxylase*, pages 65-108, CRC Series in Enzyme Biology, CRC Press, Boca Raton, USA.

Wallace, J.C., Phillips, N.B., Snoswell, M.A., Goodall, G.J., Attwood, P.V. and Keech, D.B. (1985) *Ann. N. Y. Acad. Sci.* **447**, 169-188, Pyruvate carboxylase: Mechanisms of the partial reactions.

White, F.H. Jr. (1961) *J. Biol. Chem.* **236**, 1353-1360, Regeneration of native secondary and tertiary structures by air oxidation of reduced ribonuclease.

White, F.H. Jr. (1960) *J. Biol. Chem.* **235**, 383-389, Regeneration of enzymic activity by air-oxidation of reduced ribonuclease with observation on thiolation during reduction with thioglycolate.

Woo, S., Wood, H.G., Kumar, G.K., Magner, W., Beegen, H. and Samols, D. (1990) *FASEB J.* **4**, A2318, An essential region for activity of the 12S subunit of transcarboxylase (abstract).

Wood, H.G., Lifson, N. and Lorber, V. (1945) *J. Biol. Chem.* **159**, 475-489, The position of fixed carbon in glucose from rat liver glycogen.

Wood, H.G. and Barden, R.E. (1977) *Annu. Rev. Biochem.* **46**, 385-413, Biotin enzymes.

Wood, H.G. and Kumar, G.K. (1985) *Ann. N. Y. Acad. Sci.* **447**, 1-22, Transcarboxylase: Its quaternary structure and the role of the biotinyl subunit in the assembly of the enzyme and in catalysis.

Wood, H.G. and Zwolinski, G.K. (1976) *Crit. Rev. Biochem.* **4**, 47-122, Transcarboxylase: Role of biotin, metals and subunits in the reaction and its quaternary structure.

Yeung, D. and Oliver, I.T. (1967) *Biochem. J.* **105**, 1229-1233, Development of gluconeogenesis in neonatal liver: Effect of premature delivery.

Zucker, M. and Stiegler, P. (1981) *Nucleic Acids Res.* **9**, 133-148, Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information.

Appendix I

The Kinetic approach to determining the number of residues modified

The method used by Keech and Farrant (1968) to show that inactivation of sheep pyruvate carboxylase was due to reaction of FDNB with a single amino acid residue per active site was based on that of Levy *et al.* (1963) and Scrutton and Utter (1965).

For the bimolecular reaction between the enzyme and an irreversible inhibitor,

$$E + nI \xrightarrow{k_2} EI_n$$

where E, I, EI_n , n and k_2 are free enzyme, inhibitor, enzyme-inhibitor complex, the number of molecules of inhibitor reacting per active site and the second-order rate constant for the reaction, respectively. The rate of loss of native enzyme can be written as

$$-\frac{dE}{dt} = k_2 (E) (I)^n$$

Now, if I is present in excess so that the concentration of I remains essentially unchanged, the reaction is pseudo first-order and

$$-\frac{dE}{dt} = k' (E)$$

where $k' = k_2.(I)^n$.

Then the proportion of unmodified enzyme (E/E_0) remaining at any time t, is given by

$$\frac{E}{E_0} = e^{-k't}$$

In practice A/A_0 is measured rather than E/E_0 (where A is the catalytic activity at time t, A_0 is the catalytic activity at time 0). As the modified enzyme has no residual activity under the assay conditions used, then

$$\frac{A}{A_0} = \frac{E}{E_0} = e^{-k't}$$

ie., $\ln (A/A_0) = -k't$ or $\log (A/A_0) = -1/2.303\ k't$

Thus, a plot of $\log (A/A_0)$ vs. t gives a straight line of slope $-k'/2.303$.

Now, $k' = k_2 (I)^n$, so $\log k' = \log k_2 + n.\log (I)$

Thus if k´is determined at a number of different concentrations of I, n can be obtained from the slope of a plot of log k´ vs. log (I). Alternatively, $1/t_{0.5}$ (where $t_{0.5}$ is the half time for the pseudo-first order reaction) may be used in place of k´. This changes the intercept of the plot, but has no effect on the slope.

Appendix II

The following peptide sequence data was obtained in the course of attempts to isolate the TNBS-modified peptides of yeast PC (section 5.2.3). The sequences correctly match the amino acid sequence of yeast PC, inferred from genomic clones (Lim *et al.*, 1988).

Trypsin digestion

NH_2 -Ala-Leu-Ile-Glu-Phe-Arg-COOH

NH_2 -Asp-Met-Ala-Gly-Phe-Asn-Lys-COOH

Endoproteinase glu-C digestion

NH_2 -Val-Tyr-Phe-Asp-Leu-Asn-Gly-Glu-COOH

NH_2 -Tyr-Gly-Tyr-Pro-Val-xxx-xxx-Lys-Ala-COOH

Appendix III

The western analysis results presented in poster form (Booker *et al.*, 1986) are summarised here in a schematic drawing. Lys-A was localised to the C-terminal Mr 28,500 of sheep pyruvate carboxylase. An additional modified residue (Lys-B) was localised to an internal region, bounded N-terminally by the cleavage site of chymotrypsin.