The Yeast Pyruvate Carboxylase 2 Gene (*PYC2*) and Structure-function Studies on Yeast Pyc Isozymes

by

Dale Lloyd Henry Val, Ba.Ed."Sc.", B.Sc. (Hons)

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

April, 1995
Department of Biochemistry,
University of Adelaide,
South Australia.

TABLE OF CONTENTS

Summary

Pyruvate carboxylase (Pyc) (EC 6.4.1.1) catalyses the ATP dependent carboxylation of pyruvate to produce oxaloacetate, a reaction of importance both for gluconeogenesis and the replenishment of TCA cycle intermediates utilised in other biosynthetic pathways such as lipogenesis, porphyrinogenesis and, depending on the species or tissue, also in glutamine and acetyl-choline synthesis. This enzyme belongs to a class of biotin enzymes (class I; Samols *et al.*, 1988) which all catalyse very similar two step reactions involving a biotin carboxylation step, followed by a transcarboxylation step. The Pyc protein itself is a tetramer composed of 4 identical subunits arranged in the shape of a splayed tetrahedron, and each subunit (~130 kDa) is divided into at least 3 structural domains; an N-terminal ATP binding domain, a central pyruvate binding/transcarboxylation domain, and a C-terminal biotin-carrier domain (Lim *et al.*, 1988). Sequence comparisons have revealed that each of these domains has considerable homology to the corresponding domains of the other class I biotin enzymes, and other functionally related enzymes.

One of the primary research goals of our laboratory over the years has been to characterise the structure-function relationships of pyruvate carboxylase. Prior to my research, F. Lim had sequenced the Pyc gene for *Saccharomyces cerevisiae* in our laboratory (Lim *et al.*, 1988), and Dr. M. E. Walker, upon disrupting this *PYC* gene, found evidence suggesting that there was a second isozyme for pyruvate carboxylase in yeast (Walker *et al.*, 1991).

Hence the primary research objectives of my project were to:

a) clone and characterise the *PYC2* gene from *Saccharomyces cerevisiae*

b) disrupt the *PYC2* gene in a *pyc1* null so as to construct a strain with no Pyc activity suitable as a host for the expression of mutant Pyc molecules

c) use site-directed mutagenesis to investigate the role of various amino acid residues or motifs which appear to be important on the basis of resultant sequence comparisons, and/or previous biochemical studies.

The major findings reported in my thesis are presented below.

1.) The existence of two *PYC* genes in yeast was confirmed by the presence of two hybridising bands observed on Southern blots of yeast chromosomes separated by pulse field gel electrophoresis, and probed with *PYC* specific probes from the sequenced yeast gene (*PYC1*). The *PYC1* gene was localised to chromosome VII, and *PYC2* was localised to chromosome II. These genes were each further localised on the physical map of *Saccharomyces cerevisiae* by hybridisation to a set of prime λ-clone grid filters obtained from Professor M. V. Olson (University of Washington School of Medicine, Seattle).

2.) A 3.54 kb *PYC2* clone containing 1.87 kb of the 5' non-coding region, and 2866 bp of the 5' coding region of *PYC2* was isolated from a yeast genomic library. The 3' end of this gene was obtained from a separate genomic library, and both these clones were completely sequenced in both directions. The *PYC2* gene was found to be very similar to *PYC1*, having an identity of 85.4% within the open reading frame (ORF) which encodes for an 1180 amino acid protein (Pyc2) with a 92.6% identity to Pyc1.

Extensive sequence comparisons between the domains of Pyc2 and other related proteins have highlighted a number of motifs which may prove to be important in the function of this enzyme. These sequences and their "proposed" functions respectively, are as follows: conserved cysteine and lysine residues in the ATP domain, ion pair involved in the tautomerisation of biotin; QVEH motif unique to biotin carboxylases, involved in biotin carboxylation; HXHXH motif in the pyruvate domain, Zn^{2+} binding site.

During the course of these investigations the sequence of the yeast *PYC2* gene was reported by Stucka *et al.* (1991). Sequence comparisons between the two *PYC2* sequences revealed that the published sequence differed from my data within the ORF by having 36 sequence differences, which include: 35 transversions resulting in 12 predicted amino acid differences, and one insertion at the C-terminus of the published sequence causing a frame shift changing Q1178 to a P, and extending the ORF by 15 bp (5 amino acids). These differences appeared to be due to yeast strain variation and/or sequencing errors in the published sequence.

3.) A recombinant Pyc2 C-terminal peptide containing the entire biotin domain was expressed from the clone I had isolated, and the calculated molecular weight of the resultant purified peptide (determined by mass spectrometry) agreed precisely with the theoretical value for the predicted amino acid sequence. The sequence reported by Stucka *et al.* (1991) differed in this region by a Q1178P/5 amino acid C-terminal extension, and one conservative amino acid difference (R instead of K). Hence, the corresponding recombinant biotin domain peptides containing these different sequences were also expressed, and the effect of these differences on *in vivo* biotinylation was determined.

4.) Homologous recombination was used to disrupt the *PYC2* gene using either the *TRP1* or *HIS3* selectable marker genes. Using this technique a double null mutant strain (DM18) in which both *PYC* genes had been knocked out was constructed to use as a host strain for the expression of mutant Pyc proteins. The absence of Pyc in DM18 was confirmed by Western blotting, enzyme assays and aspartate auxotrophy. In addition, studies with the single null mutant strains indicated that Pyc1 is the more abundant isozyme.

5.) There have been a number of chemical modification studies which have suggested that Pyc enzymes contain an essential cysteine residue in the pyruvate binding/transcarboxylation site. Site-directed mutagenesis was used to change individually each of the 4 cysteines in the pyruvate domain of yeast Pyc1 (to serine) so as to determine whether any of these residues are essential for Pyc activity. None of these mutations produced any major changes in enzyme activity or aspartate dependence of the host strain containing these constructs, indicating that pyruvate carboxylase does not contain an essential cysteine residue involved in base catalysis of the transcarboxylation reaction.

Statement

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University, and 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.

Dale L. H. Val

Acknowledgments

I wish to thank Professor G. E. Rogers and Drs. J. B. Egan and B. K. May for their permission to conduct research in the Department of Biochemistry, University of Adelaide. I am also very grateful to my supervisor, Dr. John C. Wallace for his support and helpful advise during the course of my work. In addition, I acknowledge the financial support of a Commonwealth Postgraduate Research Award.

I would like to acknowledge my appreciation to Dr. J. E. Cronan Jr. for the independent advise and constructive criticism he provided throughout his sabbatical in our laboratory, and thereafter, and for organising the mass spectrometry (reported in Chapter 5) which was performed at the University of Illinois, Urbana-Champaign, USA. Also, I would like to sincerely thank Dr. Michelle Walker for her guidance in the early stages of my project, for providing the *PYC1* clones and null mutant strains, and for her helpful advise throughout. I am very grateful to Dr. Anne Chapman-Smith for her technical assistance with the protein purification procedures, and for her comments and criticisms during the preparation of this thesis. Drs. Briony Forbes and Jill Carr are kindly acknowledged for their assistance in computing matters, and I am grateful to Denise Turner for performing the protein sequencing.

In addition, I would like to convey my appreciation and thanks to the past and present members of the Wallace group including Sally Aplin, Vicky Avery, Grant Booker, Neil Brewster, Kathy Bryant, Jill Carr, Anne Chapman-Smith, Briony Forbes, Goran Forsberg, Graham Hobba, Sam Hodge, Sarawut Jirapakdee, Nick Kallincos, Barbara Magee, Steven Polyak, Gary Shooter, Zee Upton, and Michelle Walker for their friendship, useful feedback on my project, and for generally being an excellent bunch of people to work with.

Finally, I am especially grateful to my parents and brother for their patience, unwavering support and encouragement, and to cousin Dean for allowing me to crash at his place when in Adelaide during the latter stages of my thesis preparation.

Abbreviations

Throughout the text yeast genes are designated by upper-case italics eg. *PYC2*, genes bearing a null mutation are shown in lower-case italics eg. *pyc2*, and the marker gene used in gene disruption to create the null mutants are shown following the abbreviation for the null gene with a space of two colons, eg *pyc2::HIS3*. Enzymes are designated by non-italicised abbreviations with the first letter shown in upper-case. In addition to those accepted for use in the *Journal of Biological Chemistry,* and those defined in the Materials and Methods Chapter, the following abbreviations are used in this thesis:

Acc	acetyl-CoA carboxylase
ACP	acyl carrier protein
oATP	2',3'-dialdehyde derivative of ATP
Bc	biotin carboxylase
BC	biotin carboxylation
BCCP	biotin-carboxyl carrier protein
BP	bromopyruvate
CLPyc	chicken liver pyruvate carboxylase
pCMB	*p*-chloromercuribenzoate
C-P	carboxy-phosphate intermediate
Cps	carbamoyl phosphate synthetase
CT	carboxyl transferase
Da	Dalton
DACM	*N*-(7-dimethylamino-4-methyl-3-coumarinyl) maleimide
DNG	did not grow
DNPS-Cl	2,4-dinitrophenylsulfenyl chloride
DTNB	5,5'-dithiobis(2-nitrobenzoate)
IOB	*o*-iodosobenzoic acid
MOPS	3-[N-Morpholino]propane-sulfonic acid

NEM	N-ethylmaleimide
OD_{600}	optical density at 600 nm
Odc	oxaloacetate decarboxylase
ORF	open reading frame
Pcc	propionyl-CoA carboxylase
PCR	polymerase chain reaction
Pdh	pyruvate dehydrogenase
PEP	phosphoenolpyruvate
Pepck	phosphoenolpyruvate-carboxykinase
PFGE	pulse field gel electrophoresis
PHMB	*p*-hydroxymercuribenzoate
Pk	pyruvate kinase
PKCI-1	protein kinase C inhibitor-1
Pyc	pyruvate carboxylase
SDS	sodium dodecyl sulphate
SH	sulfhydryl group
SKPyc	sheep kidney pyruvate carboxylase
Tc	transcarboxylase
TCA	tricarboxylic acid cycle
Ual	urea amidolyase

ERRATA

p..27, line.2 : read "Schuell" for "Schell"

p.50, line 11 : read "*PYC 2*" for "*PYC 1*"

p.56, line 1 : read "patterns" for "paturns"

p.79, line 5 : read "mass spectrometry" for "mass stectrometry"

p.93, line 23 : read "transamination" for "deamination"

CHAPTER 1

INTRODUCTION

AND

LITERATURE REVIEW

1.1 DISCOVERY OF PYRUVATE CARBOXYLASE

In studying the pathway of gluconeogenesis Krebs and Kornberg (1957) pointed out that all but three steps in glycolysis are readily reversible. One of the steps that appeared to be irreversible on thermodynamic grounds was the dephosphorylation of phosphoenolpyruvate (PEP) by pyruvate kinase to form pyruvate and ATP (Krebs, 1954). This observation was in agreement with earlier isotopic distribution studies which suggested that glycogen synthesis in the liver did not proceed by a direct reversal of the pyruvate kinase reaction (Topper and Hastings, 1949; Lorber *et al.*, 1950 a & b). To account for these data Krebs (1954), and Utter and Kurahashi (1954) proposed a dicarboxylic acid shuttle whereby PEP would be formed from pyruvate via malate and oxaloacetate, with the final step being the conversion of oxaloacetate into PEP by the enzyme PEP-carboxykinase.

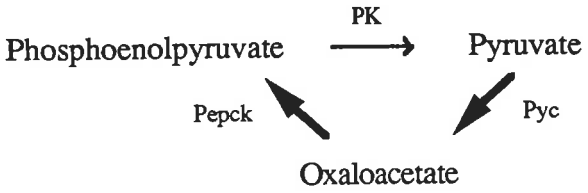
However, chicken liver mitochondria were found to synthesise PEP from pyruvate, despite containing only trace amounts of the malic enzyme believed to be involved in the conversion of pyruvate to malate in the proposed shuttle pathway (Utter, 1959). This suggested that a different enzymatic step must be involved. Indeed, further examination of chicken liver mitochondria revealed the presence of a CO_2-fixing enzyme (pyruvate carboxylase) capable of forming oxalocetate directly from pyruvate and CO_2 (Utter and Keech, 1960). Keech and Utter went on to further characterise pyruvate carboxylase from chicken liver, and the first studies on nature of the reaction, and the properties of pyruvate carboxylase were performed with this enzyme (Utter and Keech, 1963; Keech and Utter, 1963).

Chicken liver pyruvate carboxylase was found to be a biotin-containing enzyme which catalysed the ATP-dependent carboxylation of pyruvate to form oxaloacetate as outlined in equation (1) below. The enzyme activity showed absolute dependence on the presence of Mg^{2+} ions and acetyl-CoA.

$$\text{Pyruvate} + HCO_3^- + \text{ATP} \xrightleftharpoons{\text{acetyl-CoA, } Mg^{2+}} \text{Oxaloacetate} + \text{ADP} + P_i \qquad (1)$$

1.2 DISTRIBUTION AND METABOLIC ROLE

Having determined the nature of the pyruvate carboxylase reaction, Keech and Utter (1963) went on to investigate its apparent gluconeogenic role. They found that in mitochondrial extracts, the combined actions of pyruvate carboxylase (Pyc) and PEP-carboxykinase (Pepck) could indeed lead to the production of PEP. Thus, they proposed that these two enzymes may constitute a major pathway for the reversal of the pyruvate kinase (Pk) reaction during gluconeogenesis (Scheme I below). Numerous later reports have since established the gluconeogenic role of Pyc (reviewed in Wallace, 1985).



Scheme I Reversal of the pyruvate kinase reaction

⟶ glycolytic reactions

⟵ gluconeogenic reactions

In the intervening years since the initial discovery of pyruvate carboxylase in chicken liver, Pyc has been detected in a wide range of species including some groups of bacteria, yeasts, fungi, invertebrates, and vertebrates (reviewed in Wallace, 1985). More recently Pyc has also been detected in plants, in both monocots and dicots (Wurtele and Nikolau, 1990). However, its metabolic role in plants has not yet been determined.

Consistent with the gluconeogenic role of Pyc is the observation that in vertebrates the highest levels of Pyc occur in the liver and kidney (reviewed in Wallace, 1985). Nevertheless, a wider role for Pyc was clearly indicated by the frequent occurance of Pyc in non-gluconeogenic tissues, and in various microbes grown on glucose. By carboxylating pyruvate to form oxaloacetate Pyc also performs an important anaplerotic role, replenishing the TCA cycle in response to the loss of intermediates resulting from the many biosynthetic

reactions which stem from this cycle (reviewed in Utter, 1969). This anaplerotic role of Pyc is consistent with the results of numerous localisation studies showing that in vertebrates Pyc is exclusively mitochondrial (Wallace 1985; Rohde *et al.*, 1991).

On the other hand, in filamentous fungi (Osmani and Scrutton, 1983; 1985) and yeast (Rohde *et al.*, 1991; Walker *et al.*,1991) Pyc is located solely in the cytoplasm. Despite this, there are several lines of evidence suggesting that Pyc also has an important anaplerotic role in these organisms. Firstly, isotope distribution studies (Cazzulo *et al.*, 1968; Oura *et al.*, 1980) and NMR studies (Sumegi *et al.*, 1992) with *S. cerevisiae* have each demonstrated that significant flux through the TCA cycle occurs via Pyc. Secondly, in the case of *Aspergillus niger*, Pyc activity has been shown to be essential for citrate formation during growth on simple sugars (Henderson and Lamonds, 1966). Finally, in the case of both *S. cerevisiae* (Walker *et al.*, 1991; Walker and Wallace, 1991; Stucka *et al.*, 1991; Brewster *et al.*, 1994) and *Aspergillus* (Skinner and Armitt, 1972; McCullough and Roberts, 1980), Pyc mutants grown on hexoses exhibit growth dependence on L-aspartate, or other sources of C_4 TCA cycle intermediates.

1.3 FAMILY OF BIOTIN ENZYMES

Pyruvate carboxylase belongs to a large family of enzymes which all utilise biotin as a mobile carboxyl carrier (reviewed in Moss and Lane, 1972; Obermayer and Lynen, 1976; Wood and Barden, 1977; Samols *et al.*, 1988). These enzymes catalyse their reactions in two discrete steps carried out at separate sites, and the biotin prosthetic group acts as a flexible arm carrying the carboxyl group between the two sites. Depending on the nature of the original donor and final carboxyl acceptor, biotin enzymes can be divided into three classes (Wood and Barden, 1977; Samols *et al.*, 1988): Class I, Carboxylases; Class II, Decarboxylases; and Class III, Transcarboxylases. All eukaryotic enzymes belong to Class I, while prokaryotes contain enzymes from all three classes.

Carboxylases all use bicarbonate in the ATP-dependent carboxylation of biotin, followed by carboxyl-transfer to their specific acceptor molecule, either pyruvate (2-oxo acid), urea (2-oxo amide), or an acyl-CoA (2-oxo thioester). Decarboxylases on the other hand (present in anaerobic bacteria) all catalyse an ATP-independent decarboxylation of a

specific β-keto acid or acyl-CoA, coupled to sodium ion export. Transcarboxylase from *Propionibacterium shermanii* is the only example of a Class III enzyme, and this enzyme catalyses the carboxyl-transfer from methylmalonyl-CoA to pyruvate to form oxaloacetate and propionyl-CoA (reviewed in Wood, 1979).

Although it is evident that the members of the family of biotin enzymes are both functionally and structurally related, at this stage the actual rearrangement processes and ancestral relationships that link the various members of this family of proteins are poorly understood.

Comparing the partial reactions of biotin enzymes (both within and between the three classes) with their quaternary structures led to the recognition of three common functional activities: biotin carboxylation, BC; biotin-carboxyl carrier protein, BCCP; and carboxyl transferase, CT. Furthermore, each activity appeared to be carried out by different structural units contained either on separate subunits, or together on a multi-functional polypeptide chain. This led Lynen (1975) to propose that the multi-functional proteins typically found in eukaryotes were derived by a process of successive gene fusions, with each class of functional domain having originated from a separate ancestral gene.

However, Toh *et al.* (1993) have since put forward a quite different proposal. On the basis of sequence comparisons, these workers have proposed a phylogenetic tree outlining the divergence into three groups from an ancestral molecule, which they suggest contained a BC-BCCP fusion and used an acyl-CoA as the carboxyl acceptor (ie, a Class I molecule). They proposed that the progenitor of eukaryotic acetyl-CoA carboxylases (Acc; sub-group of Class I) probably diverged first, followed later by a divergence that led to mammalian propionyl-CoA carboxylases (Pcc; another sub-group of Class I) on one branch, and the pyruvate carboxylases, *E. coli* Acc, and Class II and III enzymes on the other branch. Furthermore, as some of the prokaryotic enzymes in the final branch have their functional units on separate subunits (*E. coli* Acc, and transcarboxylase from *P. shermanii*), the model of Toh *et al.* (1993) implies that these enzymes must have arrived by rearrangements other than gene fusion, as originally proposed by Lynen (1975).

1.4 REACTION MECHANISM

As previously mentioned, biotin enzymes share a number of common features. They all catalyse two step reactions involving at least one of the two common catalytic activities (biotin carboxylation, BC; and carboxyl transferase, CT), and the partial reactions are carried out at separate sites linked by the biotin mediated transfer of a carboxyl group. Furthermore, biotin enzymes which catalyse the same type of partial reaction (BC or CT) exhibit marked similarities in terms of amino acid sequence (see section. 1.5.1), and reaction kinetics. Hence its believed that all biotin enzymes share essentially the same basic mechanistic path for the BC or CT reactions (Knowles, 1989).

The overall reaction catalysed by Pyc, as is the case for all Class I biotin enzymes (Samols *et al.*, 1988), involves a biotin carboxylation step followed by a carboxyl transferase step. This separation into two partial reactions has proven to be extremely useful, enabling each half reaction to be studied in isolation by measuring isotope exchange (Utter and Keech, 1963; Keech and Utter, 1963; Scrutton and Utter, 1965; Northrop and Wood, 1969; Scrutton and Young, 1972).

As there have been excellent recent reviews detailing the evidence for and against each of the proposed mechanisms for the biotin carboxylation and carboxyl-transfer reactions (Knowles, 1989; Attwood, 1995), the following sections will be confined to only the most likely mechanism(s) for each partial reaction.

1.4.1 1st partial reaction

1.4.1.1 Binding of bicarbonate and ATP

The initial characterisation of chicken liver pyruvate carboxylase by Keech and Utter (1963) showed that the bicarbonate ion (or CO_2) was the source of the carboxyl group, and the reaction was dependent on ATP (see equ. 2).

$$MgATP + HCO_3^- + \text{ENZ-Bio} \xrightleftharpoons{\text{Mg, Acetyl-CoA}} MgADP + P_i\ \text{ENZ-Bio-COO}^- \qquad (2)$$

Tipton and Cleland (1988a) reported kinetic data suggesting a degree of synergism in the binding of ATP and HCO_3^- to the biotin carboxylase subunit of *E. coli* Acc. However, there is still some discrepancy in the literature as to whether ATP and HCO_3^- add to the Pyc enzyme in random order (McClure *et al.*, 1971; Barden *et al.*, 1972), as is believed to be the case for the Bc subunit of *E. coli* Acc enzyme (Tipton and Cleland, 1988a), or sequentially, with MgATP binding first (Warren and Tipton, 1974; Attwood and Graneri, 1992).

1.4.1.2 The carboxy-phosphate intermediate

Prior to the initial characterisation of chicken liver Pyc (Keech and Utter, 1963), ^{18}O exchange studies with propionyl-CoA carboxylase revealed that biotin carboxylation reactions involve a direct reaction between ATP and HCO_3^- (Kaziro *et al.*, 1962). The bicarbonate ion is believed to act as a nucleophile removing the γ-phosphate group from ATP to form a carboxy-phosphate (C-P) intermediate (equ. 3).

(3)

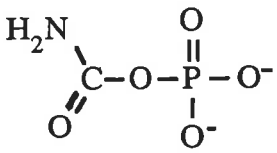
Evidence for this view comes primarily from studies demonstrating that pyruvate carboxylase and the biotin carboxylase subunit of *E. coli* Acc can each utilise carbamoyl-phosphate, a structural analogue of C-P (see Fig. 1.1), to carboxylate biotin (Ashman and Keech, 1975; Attwood and Graneri, 1991; Polakis *et al.*, 1972, 1974). Once formed the carboxy-phosphate intermediate then carboxylates the N-1 of biotin (Polakis *et al.*, 1972; Guchhait *et al.*, 1974), either directly, or after decomposing to CO_2.

1.4.1.3 Carboxylation of biotin

Upon recognising that CO_2 is a better electrophile than the C-P intermediate, Sauers *et al.* (1975) proposed that biotin was probably biotinylated by the CO_2 generated from decomposition of the C-P intermediate, rather than C-P itself. Furthermore, as the

concentration of bicarbonate is at least 10 times the concentration of dissolved CO_2 at physiological pH and temperature, Sauers *et al.* (1975) suggested that the role of ATP may be to collect HCO_3^- ions from solution in order to deliver CO_2 to biotin at the active site. However, although this appears to be an attractive scenario there is to date no experimental evidence to favour the biotin carboxylation via CO_2 rather than the C-P intermediate.

Carbamoyl-phosphate



Carboxy-phosphate

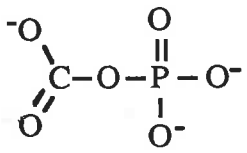
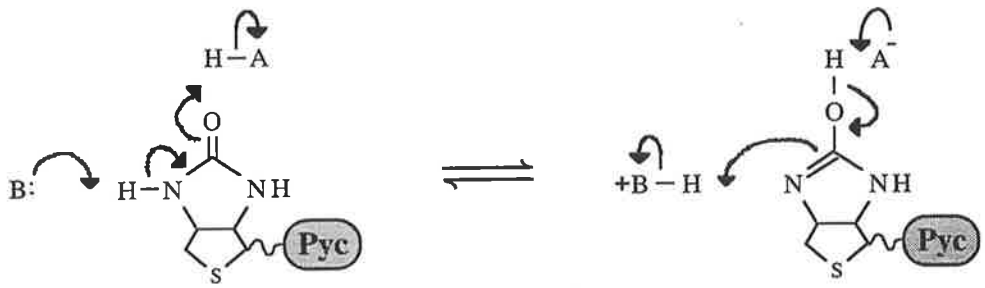


Figure 1.1 Comparison between the structures of carbamoyl-phosphate and carboxyphosphate

It is believed that in order for the N-1 atom of biotin to be sufficiently nucleophilic to react with the CO_2 or the C-P intermediate, a base-catalysed tautomerisation from the keto to the enol form must occur (see equ. 4), as the enol form is 10^{10} times more nucleophilic (Hegarty *et al.*, 1969).



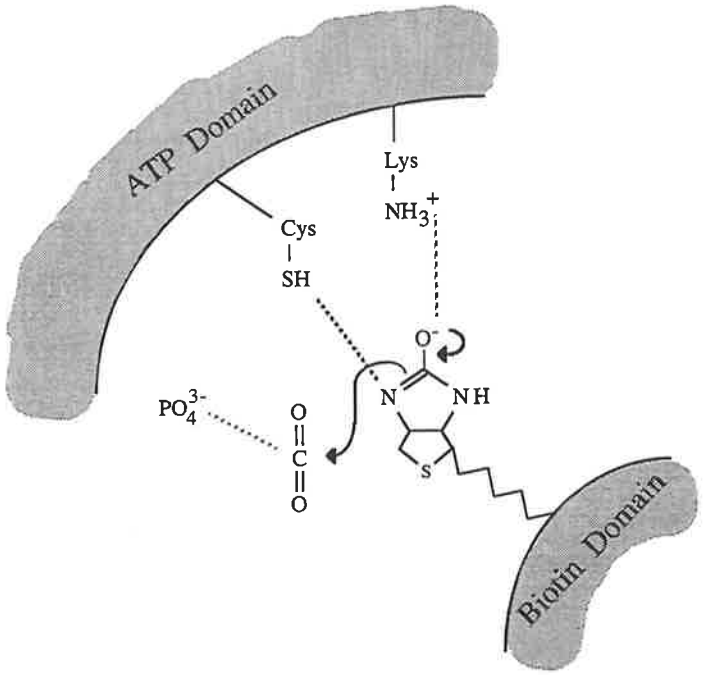
Keto form **Enol form**

(4)

To this end, pH reaction profiles obtained for Pyc (Attwood and Cleland, 1986), and pH profiles and isotope effects reported for the Bc subunit of *E. coli* Acc (Tipton and Cleland, 1988 a & b) have all implicated a number of ionisable groups in the biotin carboxylation reaction. More specifically, the findings of Tipton and Cleland (1988 a & b) led them to

propose the involvement of a base-sulfhydryl ion pair in the enolisation of biotin. They reasoned that the base may deprotonate the sulfhydryl group, which in turn deprotonates the N-1 of biotin during tautomerisation to the enol form (see Scheme II below). The protonated base then functions to stabilise the negative charge on the ureido oxygen of the enolised biotin.

In 1976 Easterbrook-Smith and co-workers reported that a competitive inhibitor oATP (2',3'-dialdehyde derivative of ATP) could be cross-linked to a lysine residue in the 1st partial reaction site of sheep liver Pyc using sodium borohydride. More recently DACM (*N*-(7-dimethylamino-4-methyl-3-coumarinyl) maleimide) modification studies have implicated the involvement of a sulfhydryl group in the 1st partial reaction (Wernberg and Ash, 1993). Wernberg and Ash (1993) went on to show that a lysine-cysteine ion pair separated by less than ~3 angstroms is located at or near the 1st partial reaction site, as lys-cys cross-linking with o-phthalaldehyde caused inactivation of the 1st partial reaction. Taken together these data imply that base involved in the above mentioned sulfhydryl deprotonation and enol stabilisation, is most likely a lysine residue (see Scheme II)



Scheme II **Carboxylation of biotin by carboxy-phosphate**

In addition to the obvious role of biotin as the carboxyl group acceptor, Attwood and Graneri (1992) recently found that the reaction between HCO_3^- and ATP was dependent on the presence of biotin at the 1st partial reaction site.

This observation is in agreement with an earlier observation made by Scrutton and Utter (1965), who found that the binding of ATP protected chicken liver Pyc against inactivation by avidin, probably due to a conformational change which stabilises the presence of biotin in the active site, thus rendering it inaccessible to avidin. This mandatory presence of biotin in the active site presumably ensures that ATP is not wasted in the synthesis of the very labile C-P intermediate (Sauers *et al., 1975*), unless the biotin moiety is correctly positioned to accept the carboxyl group it provides.

1.4.1.4 Role of magnesium ions

As previously mentioned Utter and Keech (1963) found that chicken liver pyruvate carboxylase has an absolute requirement for magnesium ions. Subsequent studies with the sheep kidney enzyme (Bais and Keech, 1972) and the Bc subunit of *E. coli* Acc (Allen *et al.*, 1984; Tipton and Cleland, 1988a), have shown that this magnesium dependence is a property of the biotin carboxylation reaction.

From kinetic studies with the sheep kidney enzyme, Keech and Barrit (1967) concluded that free Mg^{2+} ions bind to an allosteric site on Pyc resulting in a decrease in the K_m for MgATP. A more extensive study revealed that Mg^{2+} needs to be bound to the enzyme before the hydrolysis of the ATP can occur (Bais and Keech, 1972). This conclusion is supported by studies with the rat liver (McClure *et al.*, 1971) and chicken liver (Attwood and Graneri, 1992) enzymes, showing the addition of Mg^{2+} and MgATP appear to be equilibrium-ordered with Mg^{2+} adding first. Finally, Attwood and Graneri (1992) have suggested that the stimulatory effect of Mg^{2+} ions is probably a result of the increased affinity of the biotin group for the 1st partial reaction site which occurs once an MgATP is bound (see section. 1.4.1.3).

1.4.1.5 Role of acetyl-CoA

The initial studies with the chicken liver enzyme (Utter and Keech, 1963; Keech and Utter, 1963) demonstrated that acetyl-CoA is an activator of pyruvate carboxylase. Furthermore, these studies clearly showed that acetyl-CoA does not participate directly in the reaction, inferring that it acts as an allosteric activator, exerting its effects by conformational changes induced upon binding to a separate site on the enzyme. Consequently, in the intervening years there have been numerous studies investigating the effects of this molecule on pyruvate carboxylase (reviewed in Wallace, 1985; Wallace and Easterbrook-Smith, 1985; Barritt, 1985; Attwood, 1995) (see sections 1.5.6 & 1.5.8).

Studies on the degree of activation of Pyc from different species have revealed that the dependence of a particular Pyc enzyme on acetyl-CoA varies widely between species, for example: avian enzymes show total dependence (Utter and Keech, 1960; Keech and Utter, 1963); rat (McClure *et al.*, 1971), *Bacillus stearothermophilus* (Libor *et al.*, 1978), and sheep (Ashman *et al.*, 1972) enzymes have a very high dependence; the yeast enzyme is stimulated 2-4 fold (Cazzulo and Stoppani, 1968; Ruiz-Amil *et al.*, 1965); while the enzymes from *Pseudomonas citronellolis* (Seubert and Remberger, 1961) and *Aspergillus niger* (Bloom and Johnson, 1962) are unaffected by acetyl-CoA. Hence, it would appear that the acetyl-CoA dependence of a particular Pyc enzyme varies according to the different metabolic needs of that particular cell type or growth condition.

Regarding the effect of acetyl-CoA on the Pyc reaction, it has been found that the primary effect on the reaction mechanism is in stimulating the rate of the biotin carboxylation reaction (Scrutton *et al.*, 1965; Ashman *et al.*, 1972; Attwood, 1993). This involves a considerable decrease in the K_m for HCO_3^- (Cooper and Benedict, 1966; Ashman *et al.*, 1972), and the Ka for Mg^{2+} (Attwood and Graneri, 1991; 1992). In addition to these effects on the 1st partial reaction, acetyl-CoA has also been found to cause a small reduction in the K_m for pyruvate (Ashman *et al.*, 1972).

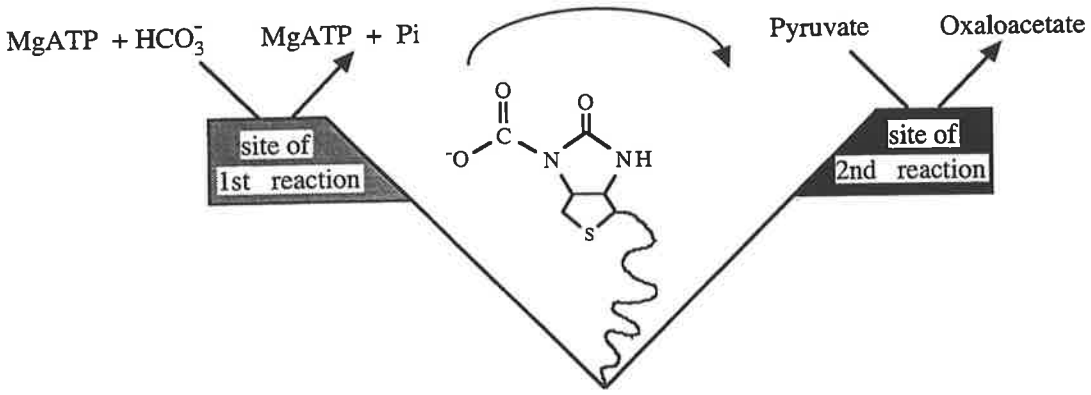
1.4.2 Translocation of the carboxy-biotin complex

1.4.2.1 Stability of the complex

Studies with Pyc isolated from sheep (Easterbrook-Smith *et al.*, 1976; Goodall *et al.*, 1981; Attwood *et al.*, 1984) and chicken (Attwood and Wallace, 1986) have shown that the carboxy-biotin complex is quite stable while it remains bound to the 1st reaction site. The presence of bound Mg^{2+} was found to further stabilise the complex by increasing its affinity for this site (Easterbrook-Smith *et al.*, 1976; Goodall *et al.*, 1981). In fact, upon fitting their kinetic data to the equations derived for alternative dissociation models, Attwood *et al.* (1984) concluded that dissociation of the carboxy-biotin moiety from the 1st partial reaction site requires the prior dissociation of the bound Mg^{2+} ion.

1.4.2.2 Translocation signal

Easterbrook-Smith *et al.* (1976) found that pyruvate binding at the 2nd partial reaction site acts as the signal inducing the dissociation of carboxy-biotin from the 1st reaction site, followed by its translocation to the pyruvate binding site (see Scheme III below).



Scheme III **Translocation of the carboxy-biotin complex, adapted from Attwood (1995)**

They concluded that the complex was unstable in the region of the 2nd reaction site, rapidly transferring the carboxyl group to pyruvate (transcarboxylation). However, if the pyruvate binding site is unoccupied the carboxyl is instead rather wastefully transferred to water, regenerating bicarbonate and short circuiting the reaction. This explained the fact that the stoichiometry of the overall Pyc reaction was found to change with pyruvate concentration.

As pyruvate was known to move in and out of the 2nd reaction site at approximately 100 times the rate of the overall reaction (Mildvan *et al.*, 1966; Mildvan and Scrutton, 1967), Easterbrook-Smith *et al* . (1976) further concluded that the final transcarboxylation step must be rapid, while the translocation of the carboxy-biotin to the 2nd reaction site is actually the rate limiting step in the carboxylation of pyruvate.

Goodall and co-workers (1981) investigated the translocation signal in more detail using a range of pyruvate analogues. They found that unlike the methyl group, the oxo and carboxyl groups of pyruvate were both required, and the rate of translocation varied considerably depending on the particular structure of the analogue. Moreover, upon comparing the kinetics of translocation reported for the analogue 2-oxobutarate (Goodall *et al.*, 1981) with their data for pyruvate, Attwood *et al* (1984) concluded that the slower rate of translocation observed with 2-oxobutarate was due to the normally rapid dissociation of carboxy-biotin from the 1st reaction site becoming rate limiting. This suggested that one way that pyruvate may be signalling the translocation of the carboxy-biotin is by reducing its affinity for the 1st reaction site. Furthermore, this reduced affinity of the carboxy-biotin complex for the 1st reaction site appears to be due to the reduction in affinity of the enzyme for Mg^{2+}, as the dissociation constant for Mg^{2+} is inversely related to the ability of a given pyruvate analogue to act as a translocation signal (Keech and Attwood, 1985).

1.4.2.3 Translocation process

Attwood and Wallace (1986) were able to study the dissociation and translocation processes separately by making use of the fact that the rate limiting step in the carboxyl-transfer changes from being the dissociation step with 2-oxobutarate as the acceptor, to the

biotin translocation step with pyruvate as the acceptor. By determining the effect of temperature on the rates of these two processes they were able to calculate their associated activation energies. This led them to conclude that the translocation process involves large changes in protein conformation, which may explain why this step is normally rate limiting.

Consistent with the above findings is the earlier suggestion that an upstream proline "hinge" may allow the whole biotin-carboxyl carrier protein domain to "flip-flop" between the two active sites (Wood and Zwolinski, 1976), in a similar fashion to the linkers which facilitate the movement of the structurally and functionally similar (see section 1.6.1) lipoyl carrier domains between the catalytic centres of the 2-oxo acid dehydrogenase complexes (Miles *et al.*, 1988). However, it is known from the crystal structure of biotin (De Titta *et al.*, 1976) that the flexible "arm" of the biotinyl-lysine is approximately 14 angstroms in length, while the distance between the active sites of transcarboxylase has been estimated by NMR techniques to be only 7 angstroms (Fung *et al.*, 1973), seemingly negating the need for large conformational changes. Clearly more research is required in order to come to a more complete understanding of the structural changes involved in the translocation process.

1.4.3 2nd partial reaction

The 2nd partial reaction catalysed by pyruvate carboxylase involves a process of carboxyl-transfer. The carboxyl group is transferred from the N-1 of biotin to the acceptor (C3 of pyruvate in the case of Pyc), with a concomitant swap with the proton from the acceptor. As is the case with all biotin enzymes, the carboxyl-transfer (CT) reaction proceeds with retention of configuration about the acceptor carbon (Mildvan *et al.*, 1966; Rose *et al.*, 1976), and proton exchange between the carboxyl acceptor and the solvent only occurs when the product, oxaloacetate, is formed (Mildvan *et al.*, 1966). In addition, by measuring the deuterium and ^{13}C isotope effects on the rate of oxaloacetate decarboxylation, Attwood *et al.* (1986b) clearly showed that the proton and carboxyl-transfers occur in separate steps, with the enolate of pyruvate being the intermediate for the reaction. However, at present there is still some discrepancy in the literature regarding the remaining mechanistic details in the pathway of CT reactions.

Comparing their empirical data with the theoretical isotope effects calculated for three alternative CT mechanisms, Attwood *et al.*. (1986b) concluded that the most likely mechanism involves the carboxyl-transfer step flanked by two enzyme catalysed proton transfers which involve an enzymic base with a low fractionation factor (Scheme IV). They further concluded that the enzymic base must be a thiol group contained in a tight ion pair with a positive group such a lysine, as this was the only grouping known to have a sufficiently low fractionation factor. However, Cleland (1992) has since noted that low fractionation factors can also be caused by low-barrier hydrogen bonds formed between two groups with very similar pKs, such as a hydrogen between two carboxylate oxygens.

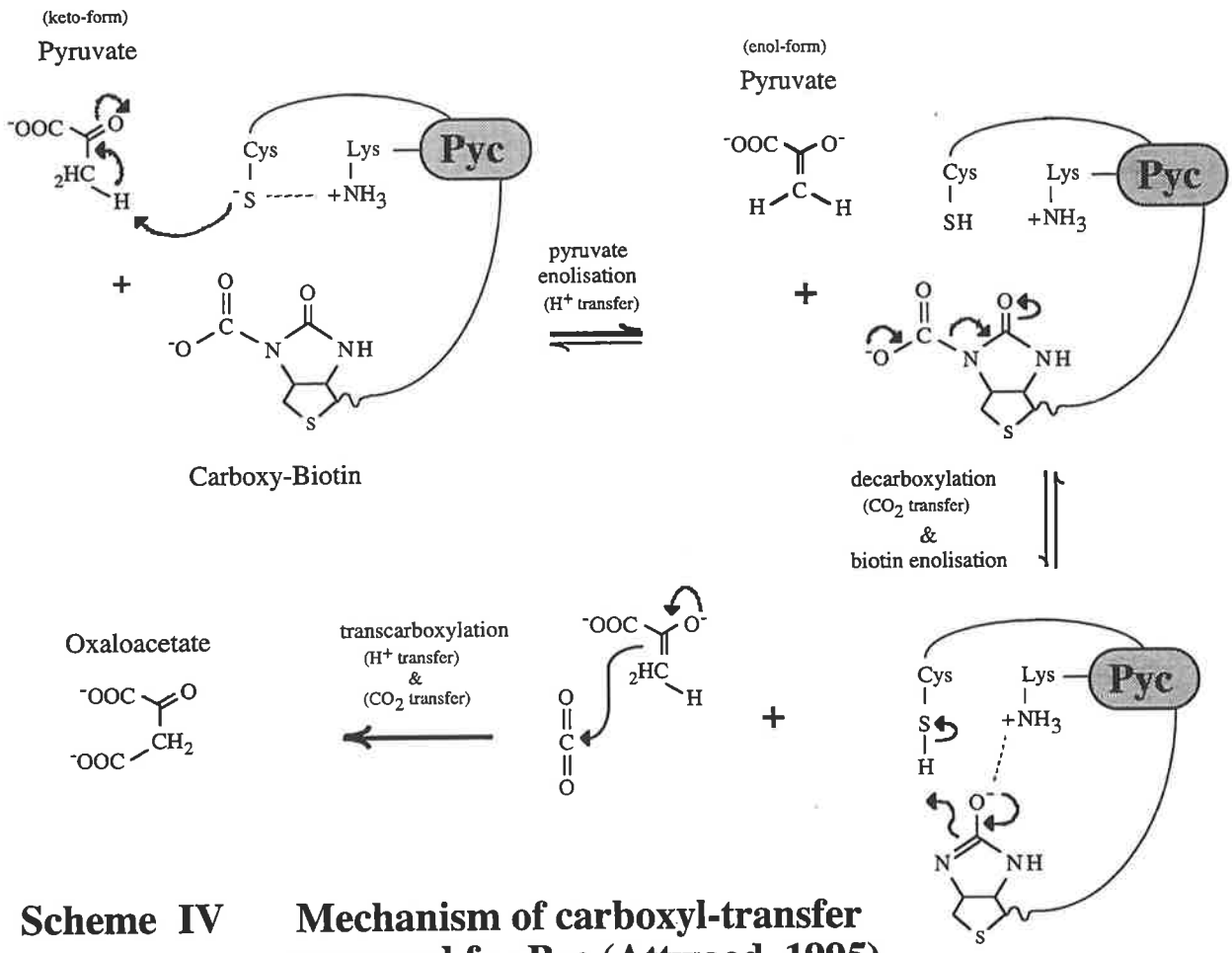
O'Keefe and Knowles (1986) similarly determined the deuterium and ^{13}C isotope effects on the CT reaction catalysed by transcarboxylase. On the basis of their data they favour a different mechanism involving a proton transfer step flanked by two carboxyl-transfer steps (Scheme V below). One additional point in support of this mechanism is that it involves no enzymic acid-base catalysis, which is consistent with pH independence displayed in the kinetic pH profile plots for the reaction catalysed by Pyc (Attwood and Cleland, 1986). However, Attwood *et al.* (1986b) calculated that the theoretical isotope effects for the mechanism favoured by O'Keefe and Knowles did not appear to agree with the isotope effects they measured for the CT reaction catalysed by Pyc.

1.5 STRUCTURE OF PYRUVATE CARBOXYLASE

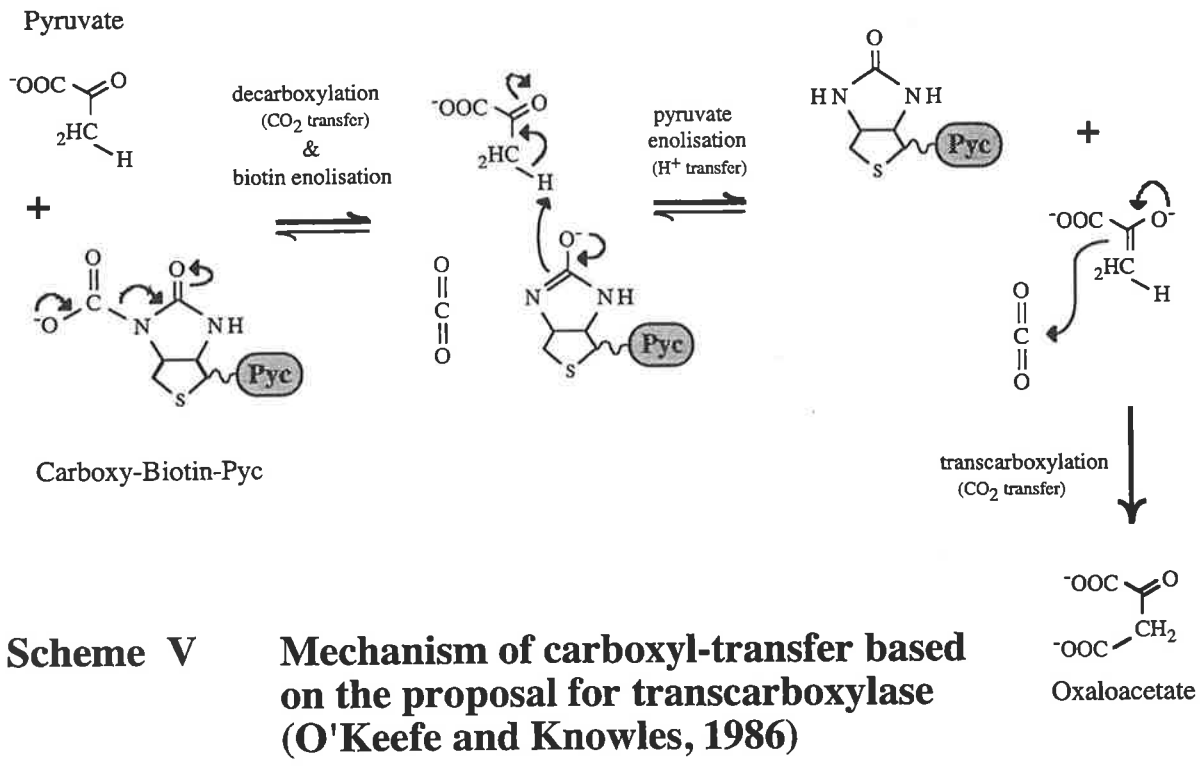
1.5.1 Primary structure

The first study into the primary structure of pyruvate carboxylase involved N-terminal sequencing of the biotinyl tryptic peptides (19-24 amino acids) from chicken, turkey, and sheep liver Pyc (Rylatt *et al.*, 1977). These workers found a high degree of sequence identity between the sequences surrounding the biotin attachment site of the three Pyc enzymes, and the corresponding sequences reported for transcarboxylase and *E. coli* Acc (Wood and Zwolinski, 1976).

With the advent of recombinant techniques, more extensive stretches of primary sequence were determined for the human Pyc enzyme (Freytag and Collier, 1984;



Scheme IV **Mechanism of carboxyl-transfer proposed for Pyc (Attwood, 1995)**



Scheme V **Mechanism of carboxyl-transfer based on the proposal for transcarboxylase (O'Keefe and Knowles, 1986)**

Lamhonwah *et al.*, 1987). Finally, workers from our laboratory determined the first complete sequence of a pyruvate carboxylase enzyme using the gene from *Saccharomyces cerevisiae* (Morris *et al.*, 1987; Lim *et al.*, 1988). In the intervening years since 1988 a number of other partial and complete Pyc sequences have been determined (reviewed in Chapter 4).

1.5.2 Domain structure

Upon determining the complete sequence of the yeast gene, Lim *et al.* (1988) searched the available data bases for proteins showing sequence similarity with Pyc. Three groups of enzymes were found to contain considerable sequence identity with various regions of yeast Pyc. These were biotin carboxylases, lipoamide transferases, and carbamoyl phosphate synthetases (Cps). The homology with Cps enzymes was interesting in the light of the considerable structural similarity between carbamoyl-phosphate, and the proposed carboxy-phosphate intermediate (see Fig. 1.1) in the 1st partial reaction of Pyc (see section. 1.4.1.2). Along with their partial proteolysis results, these data led Lim *et al.* (1988) to conclude that Pyc is composed of 3 separate structural domains with homology to other enzymes catalysing similar mechanistic processes. More specifically, Pyc was found to have the following domain composition: 1) an N-terminal domain with one region of homology to ATP binding biotin enzymes (Pyc, Acc and Pcc), and another more C-terminal region with homology to Cps enzymes; 2) a central domain with homology to pyruvate binding biotin enzymes Pyc and transcarboxylase); and 3), a C-terminal biotin domain with homology to all biotin enzymes and the lipoyl carrier domains of pyruvate dehydrogenases (see Fig. 1.2).

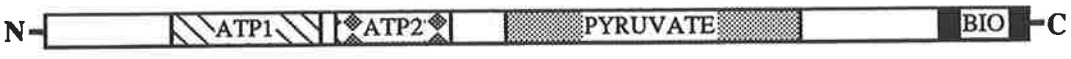


Figure 1.2 Schematic representation of the domain structure of pyruvate carboxylase, Adapted from Lim *et al.* (1988)

1.5.3 Quaternary structure

Pyruvate carboxylases are known to exist as tetramers with a molecular weight of approximately 540 kDa. The ~130 kDa monomers (Lim *et al.*, 1988) each contain a metal binding site, and a covalently attached biotin prosthetic group. Although it remained to be definitively proven at the time, the majority of evidence (reviewed in Wallace and Easterbrook-Smith, 1985) suggests that each of the monomers are identical. The only exception to this rule is the enzyme from *Pseudomonas* which is composed of α and β subunits (65 kDa and 54 kDa respectively; Barden *et al.*, 1975) arranged in an $\alpha_4\beta_4$ structure (Cohen *et al.*, 1979a). Oddly the α subunit, which contains the biotin moiety and both the half reaction sites, seems to entirely surround the β subunit (Goss *et al.*, 1981).

Electron microscopy studies into the quaternary structure of Pyc initially suggested that the yeast enzyme has a rhombic structure (Young *et al.*, 1969; Cohen *et al.*, 1979b), while conflicting reports suggested that vertebrate enzymes were either a rhombic (Cohen *et al.*, 1979c) or a splayed tetrahedron (Goss *et al.*, 1979). Subsequently, Mayer *et al.*, (1980) conducted a more detailed study showing definitively that vertebrate Pyc enzymes (chicken, sheep and rat) actually have a tetrahedron-like structure. They reported the existence of a cleft along the midline of the long axis of the monomers, and the opposite pairs of subunits (in orthogonal planes) were interacting at their convex surfaces. Furthermore, by using electron microscopy to study the long chain-like structures produced by avidin-Pyc complexes, Johannssen *et al.* (1983) located the biotin prosthetic group to the midline cleft on the concave monomer surface within 3 nm of the inter-subunit junction. As some or all of these structural features have since been reported for the *Aspergillus* (Osmani *et al.*, 1984) and yeast (Rohde *et al.*, 1986) enzymes, it would appear that with the exception of *Pseudomonas* enzyme, all Pycs contain this structure (see Fig. 1.3). In fact, it has been suggested that despite its different subunit structure, the enzyme from *Pseudomonas* also quite probably has a tetrahedral-like structure (Wallace and Easterbrook-Smith, 1985), judging from the electron microscope images it produces (Cohen *et al.*, 1979a).

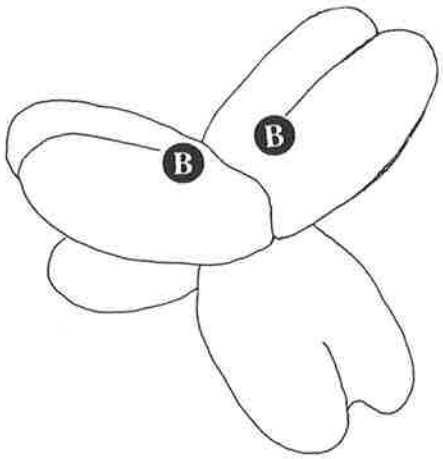


Figure 1.3 Quaternary structure of pyruvate carboxylase, adapted from Booker (1990)

1.5.4 Active oligomeric form

Insight into the active oligomeric form(s) of Pyc has been obtained by studying the structural effects of cold inactivation (Scrutton and Utter, 1965; Irias *et al.*, 1969) and dilution inactivation (Ashman *et al.*, 1972; Khew-Goodall *et al.*, 1991; Attwood *et al.*, 1993), and by using the technique of reactive enzyme sedimentation (Taylor *et al.*, 1972; 1978).

The cold inactivation observed with the chicken liver enzyme has been found to be due to dissociation into inactive monomers (Irias *et al.*, 1969), while following the course of dilution inactivation (below 4U/mL) by gel filtration revealed that only the tetramers of the sheep (Khew-Goodall *et al.*, 1991) and chicken enzymes (Attwood *et al.*, 1993) were active. Similarly, the sedimentation technique of Taylor *et al.* (1972; 1978) revealed that yeast enzyme was a stable active tetramer, and apart from an active higher molecular weight (22.7S) presumably "octameric" species (formed at high concentration), the chicken enzyme was also only active as a tetramer. However the rat enzyme appears to be active in all

oligomeric forms (Nakashima *et al.*, 1975; Taylor *et al.*, 1978), although it has been suggested that this may be due to the special assay conditions used by these workers (Khew-Goodall *et al.*, 1991).

While it is clear from the above studies that the active oligomeric form of Pyc is predominantly, if not exclusively the tetramer, Khew-Goodall *et al.* (1991) and Attwood *et al.* (1993) found evidence for the reassociation of the monomers into inactive tetramers, thus indicating that the activity of Pyc is also dependent on the correct conformational form of the component monomers.

1.5.5 Metal binding site

It has long been known that pyruvate carboxylase enzymes contain one tightly bound divalent metal ion per subunit. The chicken (Scrutton *et al.*, 1966), sheep (Bais, 1974) and a number of other vertebrate Pycs (Scrutton *et al.*, 1973a) all contain a bound Mn(II) ion, while the yeast (Scrutton *et al.*, 1970) and thermophilic *Bacillus* (Libor *et al.*, 1979) enzymes both contain Zn(II) ions.

Scrutton and co-workers (1973b) used NMR techniques to show that the Mn(II) ion is located in the region of the pyruvate binding site. However, several lines of evidence suggest that the metal ions have a structural rather than mechanistic role (reviewed in Scrutton *et al.*, 1973b). Firstly, NMR spectroscopy experiments revealed that the Mn(II) ion was at least 1 nm from the ATP binding site, and therefore could not be directly involved with 1st partial reaction (Reed and Scrutton, 1974). Other NMR studies (Fung *et al.*, 1973; Reed and Scrutton, 1974) have shown that Mn(II) is too far away (0.6-0.7 nm) from pyruvate for it to be an inner sphere ligand, and is thus also not a part of the 2nd partial reaction site. Taken together these data imply that the bound metal ion has a structural rather than an a mechanistic role. This view has received recent support in the finding that the inactivation of chicken liver Pyc by the metal ion chelator 1,10-phenanthroline is associated with a loss of the enzymes tetrameric structure (Carver *et al.*, 1988).

1.5.6 Acetyl-CoA binding site

As outlined in section 1.4.1.5, a number of kinetic studies have revealed that acetyl-CoA binds to an allosteric site. Frey and Utter (1977) showed that there is one acetyl-CoA binding site per subunit, and several studies have since been directed at characterising this site.

Chemical modification studies using the enzymes from sheep (Keech and Farrant, 1968; Ashman *et al.*, 1973), *Aspergillus* (Osmani and Scrutton, 1981), *Bacillus* (Libor *et al.*, 1978), rat, chicken, and yeast (Scrutton and White, 1973), have each reported the desensitisation of acetyl-CoA stimulation upon modification of a single lysine residue, which can be protected by the binding of acetyl-CoA. However, the actual role of this modified lysine residue remains to be proven. It may be directly involved in the acetyl-CoA binding site, as originally proposed by Keech and Farrant (1968), and most recently by Chapman-Smith *et al.* (1991). Alternatively, it may simply be involved in activation by acetyl-CoA, and inaccessible to chemical modification (possibly due to conformational changes) while the nearby acetyl-CoA binding site is occupied (Ashman *et al.*, 1973; Scrutton *et al.*, 1977). In any event the modified lysine residue has so far not been located in the primary amino acid sequence, and sequence comparisons between various Pycs and other acetyl-CoA binding proteins have failed to identify the acetyl-CoA binding site (Booker, 1990).

1.5.7 Effects of substrates

There are several lines of evidence suggesting that substrates of the 1st and 2nd reaction sites affect both the conformation of Pyc, and its oligomeric form.

Scrutton and Utter (1965) concluded that the protection against avidin inactivation afforded by ATP (see section. 1.4.1.3) was most likely due to changes in the conformation of the 1st partial reaction site in the area occupied by the biotin ring. Also, in studying cold inactivation of the chicken liver enzyme, Irias *et al.* (1969) showed that ATP was able to induce both reassociation and reactivation.

Similarly, several observations have indicated that substrate binding at the 2nd partial reaction site is also associated with conformational change. Firstly, as outlined in section 1.4.2.2, the binding of pyruvate induces the dissociation of the carboxy-biotin complex from

the 1st reaction site. Also, Attwood and Graneri (1992) noted that the rate of the biotin carboxylation at the 1st reaction site is markedly less when carried out in the absence of pyruvate, presumably due to the enzyme being in a less active form. Furthermore, by employing electron microscopy to monitor the formation of complexes between avidin and Pyc (sheep or chicken), Attwood *et al.* (1986a) found that in the absence of acetyl-CoA high concentrations of pyruvate (or the analogue 2-oxobutarate) were able to mimic the acetyl-CoA effect in maintaining the quaternary structure of Pyc in the tight tetrahedron-like conformation.

1.5.8 Effect of activators

For both Mg^{2+} and acetyl-CoA binding there is more direct evidence for induced conformational changes than was the case for the enzyme substrates. Bais and Keech (1972) reported two lines of evidence demonstrating considerable conformational change associated with the Mg^{2+} activation of sheep kidney Pyc. By measuring the effect of temperature on Mg^{2+} activation, they calculated that the change in entropy on Mg^{2+} binding was high (58.6 U/mole), and in the range consistent with the occurrence of a conformational change. Secondly, the fluorescence emission intensity of 1-anilino-naphthalene-8-sulfonate, which binds to hydrophobic regions of proteins, was markedly lowered by the binding of Mg^{2+}.

In the case of acetyl-CoA, numerous studies point to the fact that this activator produces the greatest structural change. For example, conformational changes upon acetyl-CoA binding have been reported by measuring changes in the ultraviolet absorption spectra (Frey and Utter, 1977), fluorescence emission (McGurk and Spivey, 1979), and the sedimentation coefficient of yeast Pyc tetramers (Taylor *et al.*, 1978). Acetyl-CoA also protects against dissociation caused by low temperatures (Irias *et al.*, 1969), dilution (Khew-Goodall *et al.*, 1991; Attwood *et al.*, 1993), mild denaturants (Irias *et al.*, 1969; Scrutton and Utter 1967), and moderate changes in pH (Irias *et al.*, 1969). Furthermore, electron microscopy studies with the enzymes from yeast (Rohde *et al.*, 1986), and various vertebrate species (Mayer *et al.*, 1980; Attwood and Wallace, 1986; Khew-Goodall *et al.*, 1991; Attwood *et al.*, 1993), have shown that the presence of acetyl-CoA leads to formation of

more compact tetramers. In the case of the vertebrate enzymes, this is accompanied by a reduction in the size of the longitudinal midline cleft along the monomers (Mayer *et al.*, 1980).

Regarding the interaction between Pyc subunits, sigmoidal velocity curves initially led to the belief that the subunits exhibit cooperativity in the binding of acetyl-CoA. However, Easterbrook-Smith *et al.* (1979) have since clearly shown that this apparent cooperativity is due to the effects of acetyl-CoA in protecting against dilution inactivation, since assays are frequently performed at less than 4U / mL, and in reducing the K_m for pyruvate.

1.6 THE BIOTIN CARRIER DOMAIN

1.6.1 Structure

As outlined in section 1.5.2, the biotin prosthetic group is covalently attached to the most C-terminal domain of the Pyc protein, which we will refer to as the biotin domain. In fact, with the exception of the vertebrate Acc enzymes (chicken, Takai *et al.*, 1987; rat, Bai *et al.*, 1989), biotin domains are always located at the C-terminus. The biotin moiety is attached via an amide bond to a specific lysine residue (Samols *et al.*, 1988; Cronan 1990) which is generally 34-35 residues from the C-terminus (Samols *et al.*, 1988) and is located in a conserved AMKM tetra peptide. The only exceptions to the tetra peptide attachment sequence are once again in Acc enzymes, some of which have either the alanine replaced with a valine, or the C-terminal methionine replaced by a leucine (see Fig. 4.10).

The biotin domain itself is known to be a proteolytically stable (Fall and Vagelos, 1973; Cronan, 1990) independently folded domain (Reed and Cronan, 1991) of 75-80 residues (Cronan, 1990; Li and Cronan, 1992). Sequence comparisons have revealed considerable identity between the domains from a wide range of enzymes and species (Lim *et al.*, 1988; Samols *et al.*, 1988; Kondo *et al.* 1991; Li and Cronan, 1992), and Lim *et al.*, (1988) found that biotin domains also have sequence similarity with the lipoyl domains of pyruvate dehydrogenase (Pdh).

Interestingly, this sequence similarity between biotinyl and lipoyl domains is further reflected in many other structural features. For example, the lipoyl cofactor, much like biotin, acts as a swinging arm carrying a reaction intermediate between active sites (Reed, 1974), and it too is covalently attached to a lysine residue similarly located within the ~80 amino acid domain (Perham 1991; Reed and Hackert, 1990). This structural similarity is further revealed by the fact that avidin, in addition to its well established high affinity for biotin, has been found to interact with the lipoyl domains in the Pdh multi-enzyme complex (Visser and Kellogg, 1978; Hale *et al.*, 1992). It has also been noted that biotin domains frequently contain upstream pro-X-pro or pro / ala rich sequences (Wood & Zwolinski, 1976; Lamhonwah *et al.*, 1987; Cronan, 1990) which resemble the flexible linker sequences (Texter *et al.*, 1988; Radford *et al.*, 1989) that flank the lipoyl domains.

The first insights into the three-dimensional structure of biotin domains were provided by the solved structures of lipoyl domains. The 3D structure of the Pdh lipoyl domains from *B. stearothermophilus* (Dardel *et al.*, 1993) and *E. coli* (Green *et al.*, 1995) have been determined by NMR, and the lipoyl domain from the pea leaf H-protein by x-ray crystallography (Pares *et al.*, 1994). Based on the retention of certain amino acids in key positions, Brocklehurst and Perham (1993) proposed a three-dimensional structure for the biotin domain of yeast Pyc, predicting that it contains the same fold as the Pdh lipoyl domain. Brocklehurst *et al.* (1995) have since added credibility to this prediction by determining the 3D solution structure of the biotin domain of the BCCP subunit of *E. coli* Acc.

These workers found that the BCCP biotin domain showed a striking resemblance to the lipoyl domain structures, being a compact globular domain adopting a Greek fold consisting primarily of two four-stranded anti-parallel β-sheets (see Fig. 1.4). Furthermore, the biotin and lipoyl groups were found to occupy analogous positions in their respective domains, each protruding from the tip of a corresponding hairpin loop.

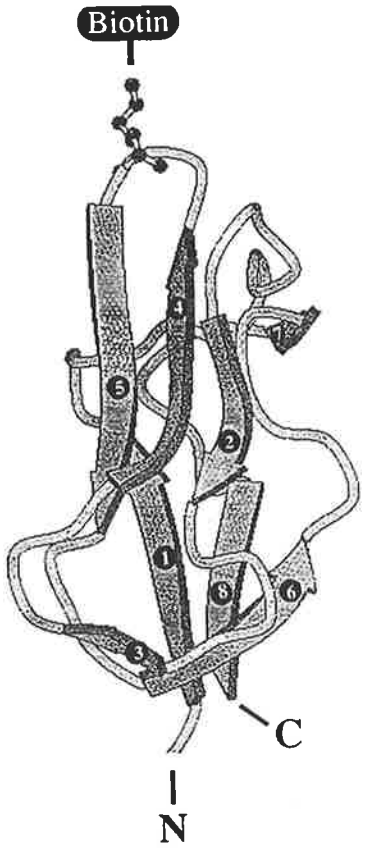


Figure 1.4 Schematic drawing of the three-dimensional structure of the *E. coli* BCCP biotin domain, adapted from Brocklehurst *et al.*, (1995)

The eight strands of β-sheet making up the bulk of the secondary structure of the biotin domain are numbered from the N-terminus, and the N and C-termini of the structured biotin domain are as shown. The carbon chain of the biotinylated lysine residue is shown in "ball and stick" format.

1.7 PROJECT BACKGROUND

The primary research objective of our laboratory over the years has been to characterise the structure / function relationships of pyruvate carboxylase. Research towards this end was advanced considerably when previous members of our group determined the complete sequence of a pyruvate carboxylase gene from *Saccharomyces cerevisiae* (Lim *et al.*, 1988). In addition to establishing the first complete amino acid sequence, this finding

was important in opening the door to the use of a new and powerful technique in the study of this enzyme. Site-directed mutagenesis could now be applied to precisely locate the various catalytically and structurally important residues involved in each of the partial reactions.

With this objective in mind Dr. M. E. Walker then set about constructing yeast *pyc* null mutants devoid in Pyc activity to use as hosts for the expression of various constructs in structure / function studies. However, Dr Walker found that null mutants strains containing the *PYC* gene disrupted with either the *LEU2* or the *HIS3* marker genes (strains MW21.3 and MW15.2.2 respectively) still contained 10 - 20 % of the Pyc activity of the parental wild type strain (Walker *et al.*, 1991). Also, Western analysis showed that the null mutants contained a biotinylated protein of identical size to the purified Pyc protein. Moreover, a yeast Pyc poly clonal antibody / protein A-gold complex detected the presence of a cytosolic immunoreactive protein in the null mutant strain MW21.3. Taken together these results suggested that *Saccharomyces cerevisiae* contained another gene encoding a second cytosolic form of Pyc. The existence of a second gene was further supported by the fact that genomic Southern blots from MW21.3, MW15.2.2, and parental strain, all displayed a hybridising band of a size which did not correspond to the restriction map of the isolated gene.

Clearly, in order for the structure / function studies to be carried out in yeast, this second gene for Pyc would have to be disrupted so that a host strain truly devoid of Pyc activity could be obtained. In addition, it was expected that characterisation of the second gene may well provide further insights into the structure / function relationships of this important enzyme.

An additional approach to the study of the structure and function of Pyc made possible by the findings of Lim *et al.* (1988), was that of expressing and characterising each of the structural domains of the protein in isolation. This approach may provide a means of determining the three-dimensional structure of each of the domains of Pyc, which would be valuable in the light of the fact that previous attempts to crystallise the whole enzyme have been unsuccessful (Lim, 1988).

1.8 AIMS

The specific aims of my project were to:

a) localise the *PYC* genes on the physical map of *Saccharomyces cerevisiae*

b) clone and characterise the *PYC2* gene from *Saccharomyces cerevisiae*

c) use site-directed mutagenesis to investigate the role of various amino acid residues or motifs which appear to be important on the basis of resultant sequence comparisons

d) express and purify a yeast Pyc biotin domain peptide for three-dimensional structure determination by NMR

e) disrupt the *PYC2* gene in a *pyc1* null so as to construct a strain with no Pyc activity suitable as a host for the expression of mutant Pyc molecules

f) use site-directed mutagenesis to investigate the hypothesis that Pyc enzymes contain an essential cysteine residue which functions as a catalytic base in the 2nd partial reaction.

CHAPTER 2

MATERIALS AND METHODS

2.1 MATERIALS

2.1.1 General chemicals

The following chemicals were obtained from Sigma Chemical Co., St Louis, MO., USA: agarose (type I), ampicillin, arabinose, aspartate, ATP (disodium, grade I), D-biotin, bis[2-Hydroxymethyl]imino-tris[hydroxymethyl]methane (Bis-Tris), bromochloroindolephosphate (DCIP), chloramphenicol, coenzyme A, Coomassie Brilliant Blue (G_{250}, R_{250}), dithioerythritol (DTE), dithiothreitol (DTT), ethidium bromide, ethylenediaminetetraacetic acid (EDTA), glycine, 8-hydroxyquinoline, N-lauroylsarcosine (sarkosyl), 2-mercaptoethanol, 3-[N-Morpholino]propane-sulfonic acid (MOPS), nitroblue tetrazolium (NBT), phenylmethylsulphonyl fluoride (PMSF), salmon testis (sperm) DNA, sodium dodecyl sulphate (SDS), sodium pyruvate, sorbitol, N,N,N',N'-tetramethyl-ethylenediamene (TEMED), tetracyclin, and tetramethylammonium chloride. Glycogen was obtained from Boehringer Manheim; acrylamide and Tris-(hydroxymethyl)methylamine (Tris) were purchased from Merck Pty. Ltd., Kilsyth, Vic., Australia; NN''-methylene-bis-acrylamide was purchased from BioRad Laboratories Inc., Herates, CA, USA; bromophenol blue, xylene cyanol, chloroform and acetic acid were obtained from BDH Chemicals, Australia, Pty, Ltd.; phenol (special grade) was obtained from Wako Pure Chemical Industries Ltd., Osaka, Japan; polyethylene glycol (PEG) 6000 was obtained from AJAX Chemicals, Sydney, NSW, Australia; 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (BCIG, X-GAL) was obtained from Progen Industries Ltd.; isopropyl-β-D-thiogalactopyranoside (IPTG) was obtained from Diagnostic Chemicals Ltd., Charlottown, Canada; and HiSafe OptiScint and OptiPhase scintillation fluids were purchased from Pharmacia, Australia, Pty. Ltd.

2.1.2 Radiochemicals

[α-^{32}P] dATP, [α-^{35}S] dATP, [γ-^{32}P] ATP, and [α-^{32}P] dCTP (3000 Ci / mmol) were all purchased from Bresated Ltd., Adelaide, South Australia. $NaH^{14}CO_3$ (50-60 mCi / mmol) and [^{3}H]-biotin (35 Ci / mmol) were obtained from Amersham Australia, North Ryde, NSW, Australia.

2.1.3 Blotting membranes and glass beads

Nytran and Nitrocellulose membranes were obtained from Schleicher and Schell, and glass beads (0.45 micron) were obtained from Sigma Chemical Co., St Louis, MO.

2.1.4 Enzymes and proteins

Avidin alkaline phosphatase conjugate, *E. coli* DNA polymerase I (Klenow fragment), T4 DNA ligase, and T4 polynucleotide kinase were obtained from Bresatec Ltd, Adelaide, South Australia. Avidin (egg white), bovine serum albumin (fraction V), DNase I, lysozyme, lyticase (high grade), proteinase K, and RNase A were purchased from Sigma Chemical Co., St Louis, MO. Calf intestinal phosphatase, Mung Bean Nuclease and exonuclease III were obtained from Boehringer Manheim, unmodified T7 DNA polymerase was obtained from New England Biolabs Inc., MA, USA., Sequenase was obtained from United States Biochemical, Taq DNA polymerase (Pyrostase) was obtained from Molecular Genetic Resources Inc., Florida , USA., and restriction enzymes were principally purchased from Pharmacia LKB and New England Biolabs Inc., MA, USA. Biotinylated molecular weight markers (BSA, 66 kDa; glutamate dehydrogenase, 55.6 kDa; carbonic anhydrase 29 kDa; lysozyme 14.3 kDa) were the kind gift of Dr. S. Gargosky.

2.1.5 Oligonucleotides

Oligonucleotides were purchased from Bresatec Ltd., Adelaide, South Australia. The restriction sites in the oligonucleotides are indicated by underlining, and the mutagenic changes are shown by the larger bold lettering. The sequences of the oligonucleotides are as follows:

O4288, 5'-GATAATCAAAGTCTTACC-3';

O3671, 5'-CTGTTGCTAAACC**C**A**T**GGCTGATGTCCAC-3';

O6612, 5'-CTACCCCCAT **C**CCCAAAAAAAGTAA-3';

O6613, 5'-GTTTTCATTA**G**GGATGGTGAAAGTG-3';

pSeq, 5'-CTAGATAAAATATACGCC-3';

PCR*Bam* HI, 5'-GAATGGA**T**C**C**ATGGTAAGTG-3';

O5011, 5'-CGAATTTTTCTCAT**GA**CACCATTCAAATC-3';

O5012, 5'-CATGCATGTCTGCCATGGATTTAGTAAC-3';

O2021, 5'-GCGACATTGTCCTCGAGGAAGGGGA-3';

YPC04, 5'-GCACCACCCCAAGATTCATCTAAGG-3';

YPC05, 5'-TCCCCAGAGAAAGAAACAGTGGCTT-3';

YPC06, 5'-CCCGCCAGAGCAGACGCAGTCATTG-3';

YPC07, 5'-TCAGCCTCGAAAGAAGAGTATAACA-3';

2.1.6 DNA clones and vectors

The pMW4A clone containing the full length *PYC1* gene used as the source of *PYC1* DNA probes, the Pyc1 biotin domain coding sequence, and the full length *PYC1* gene for the cysteine mutagenesis studies was supplied by Dr. M. Walker (University of Adelaide).

The clones used to prepare the various chromosome specific probes referred to in Chapter 3 were kindly supplied by Dr. R. Devenish (Monash University), and they are as follows: JW#16 (Welch *et al.*, 1989), *CUP2* gene from the *RAD2* proximal fragment of chromosome VII; YIp333 (Eibel & Philippsen, 1983), *LYS2* gene from chromosome II; pPM408 (Carle and Olson, 1985), centromere sequences from chromosome XIV (Carle and Olson, 1985).

The pCY216 plasmid used to express the *E. coli Bir* A gene product from an arabinose inducible promoter was kindly provided by Dr. J. E. Cronan Jr. (University of Illinois, Urbana-Champaign). The *TRP1* gene used for preparing the T3 homologous recombination construct was obtained from the plasmid YRp7 (Struhl *et al.*, 1979; Stinchcomb *et al.*, 1979), while the YEp6 plasmid (Struhl *et al.*, 1979) was the source of the *HIS3* gene used to prepare the H11 homologous recombination construct .

The cloning vector KS+ bluescript was obtained from Stratagene, while pHSG396 was supplied by Dr. S. Takeshita (Takeshita *et al.*, 1987). The T7 expression vector pET-16B was obtained from Novagen Inc., the yeast expression vector pVT100-U was provided by Dr. T. Vernet (Vernet *et al.*, 1987), and the yeast-*E. coli* shuttle vector YCpLAC33 was supplied by Dr. R. Geitz (Gietz and Sugino, 1988).

2.1.7 Yeast strains, genomic libraries and chromosomal localisation filters

DBY746 *MATα, his3, leu2, trp1, ura3, PYC1, PYC2*

DBY747 *MATa, his3, leu2, trp1, ura3, PYC1, PYC2*

MW21.3 *MATα, his3, trp1, ura3, pyc1::LEU2, PYC2*

YPH80 *MATα, his7, lys2, ade, trp1, ura3, PYC1, PYC2*

YPH149 *MATa, his7, lys2, ade, TRP1, URA3, PYC1, PYC2*. This strain carries a *URA3*$^+$ proximal *RAD2* chromosome VII fragment, and a *TRP1*$^+$ distal *RAD2* chromosome VII fragment.

The pEMBL Ye23 *Sau* 3A genomic library from *S. cerevisiae* strain S288C was obtained from Dr. C. P. Morris (formerly at the University of Adelaide), and the λEMBL3 library was obtained from Dr. F. Lim (Lim *et al.*, 1988).

The prime λ-clone grid filters used for the physical mapping of the yeast genes were obtained from the laboratory of Dr. M. V. Olson (Riles *et al.*, 1993).

2.1.8 Bacterial strains

E. coli DH5α *supE*44, Δ*lac*U169, (ϕ80 *lacZ*ΔM15), *hsdR*17, *recA*1, *endA*1, *gyrA*96, *thi*-1, *relA*1

E. coli ED8799 *hsdS*$^-$, (rk$^-$, mk$^-$), *metB7*, *supE*, (*glnV*)44, *supF*, (*tyrT*)*S8*, Δ(lacZ) M15

E. coli LE392 *supE*44, *supF*58, *hsdR*514, g*alK*2, *galT*22, *metB*1, t*rpR*55, *lacY*1

E. coli MV1190 Δ(*lac-proAB*), *thi*, *supE*, Δ(*srl-recA*)306::Tn*10*(*tet*^{r}), *F*' [*raD36*, *proAB*, *lacI*^{q}Z Δ*M15*]

E. coli CJ236 *dut1*, *ung1*, *thi1*, *relA1*/pCJ105 (Cm^r)

E. coli BL21λDES used as a host for expression studies, is an *E. coli* B strain (F- *ompT* r_B- m_B-) lysogenised with a prophage (λDES) which expresses T7 RNA polymerase upon IPTG induction (Studier and Moffat, 1986).

2.1.9 Culture media

All media were prepared using deionised water, and then autoclaved prior to the addition of antibiotics, amino acids or other heat sensitive supplements. Media plates were then prepared by the addition of 12g of agar per litre of broth. The *E. coli* media L-broth, 2xYT, NZCYM and SOC were prepared according to standard methods (Sambrook *et al.*, 1989), and the yeast broths, YPD and glucose minimal media were prepared as follows:

SD glucose minimal broth:	Yeast Nitrogen Base (without amino acids)	1.7 g / L
	$(NH_4)_2SO_4$	5 g / L
	Glucose	20 g / L

The required amino acid supplements were added to a final concentration of 100 mg / L. Histidine, uracil, leucine, adenine were added prior to autoclaving, while tryptophan was added after autoclaving as a filter sterilised solution.

YPD glucose rich broth:	Yeast Extract	100 g / L
	Glucose	200 g / L
	Bactopeptone	200 g / L
	$K_2H_2PO_4$	0.5 g / L
	KH_2PO_4	0.5 g / L

2.1.10 Chromatographic media

Q-Sepharose, Sephacryl S-200, Sepharose CL-6B Superose-6, and the 1 mL Resource-Q column were all purchase from Pharmacia LKB, Uppsala, Sweden. The reverse-phase C4 HPLC column was from Brownlee labs, Applied Biosystems.

2.1.11 Molecular biology kits

GIGAprime oligonucleotide labelling kits (Cat. No. GPK-1) and 5' Terminal Kinasing kits (Cat. No. TKK-1) were purchased from Bresatec Ltd. Adelaide, South Australia. ^{T7}Sequencing kits (Cat. No. 27-1682-01) and the *double-stranded* Nested Deletion kit (Cat. No. 27-1691-01) were purchased from Pharmacia LKB. The Muta-Gene

Phagemid *In Vitro* Mutagenesis kit (Cat. No. 170-3576) was obtained from BioRad Laboratories Inc., Herates, CA, USA.

2.2 METHODS

2.2.1 DNA methods

2.2.1.1 Isolation of double-stranded plasmid DNA

Double stranded "mini-preps" of plasmid DNA were prepared from 3 mL overnight cultures in L-broth by the alkaline lysis method based on the procedure of Birnboim and Doly (1979). Large scale preparation of plasmid DNA was similarly performed by a scaled up version of the same method using 50 mL overnight cultures.

2.2.1.2 Denaturation and preparation of plasmid DNA for DNA sequencing

Two µL of RNase A (10 mg / mL) and 8 µL of sterile deionised water were added to 10 µL of "mini-prep" plasmid DNA (from *E. coli* DH5α), mixed and incubated at 37 °C for 15 minutes. The DNA was denatured by the addition of 5 µL of 1 M NaOH / 1 mM EDTA and incubating at 37 °C for 15 minutes. The denatured RNased DNA was then purified by spinning the samples through a 0.5 mL Sepharose CL-6B column in a PCR tube at 1800 rpm using a bench centrifuge as described in Sambrook *et al.* (1989).

2.2.1.3 Purification of large scale plasmid DNA by FPLC

Large scale DNA samples requiring further purification prior to use in techniques such as the transformation of yeast were purified after RNase A digestion by gel filtration chromatography. Briefly, 0.5 mL DNA samples were applied to a 30 mL Superose-6 column and eluted at 0.5 mL / min in 100 mM sodium acetate pH 7 / 0.05 % SDS / ethanol. The absorbance at 260 nm was monitored and the DNA peak was collected directly into an Eppendorf tube.

2.2.1.4 Isolation of single-stranded phagemid DNA

Single stranded uDNA (from *E. coli* CJ236) to be used in mutagenesis was prepared from 2 mL overnight 2xYT cultures using the protocol supplied with the BioRad Muta-Gene phagemid *in vitro* mutagenesis kit.

2.2.1.5 Isolation of λDNA

Amplification of the *PYC2* λEMBL3 clones was achieved by setting up 2 x 2 mL NZCYM cultures per clone, inoculating each with 0.2 mL of plating bacteria (*E. coli* LE392 grown to stationary phase in 50 mL of NZCYM + 0.2% maltose, and resuspended in 10 mM $MgSO_4$) and 0.2 mL of λ eluate followed by incubation at 37 °C for 2 hours. At the same time 2 x 20 mL cultures of NZCYM were set up with 0.3 mL of plating bacteria. After the completion of the 2 hour incubations the 2 mL lysis cultures were added to the 20 mL bacterial cultures, and cultures were incubated (as previously described) for a further 3.5 hours, or until cell lysis was complete.

Three drops of chloroform, 7 mL of 10 mg / mL RNase A, and 7 mL of 5 mg / mL DNase I was added, and the cultures were left at 4 °C overnight. NaCl was added to 1 M and the cell debris was removed by centrifugation (10,000 rpm for 15 minutes at 4 °C in an SS34 rotor), the phage particles were then obtained by centrifugation of the supernatant at 18,000 rpm for 2 - 3 hours at 4 °C in an SS34 rotor. The protein coat was removed by a 65 °C incubation for 45 minutes in buffer K (0.02 M EDTA pH 8.0, 0.05% SDS, 0.5 mg / mL proteinase K), followed by phenol / chloroform extraction and ethanol precipitation.

2.2.1.6 Isolation of genomic DNA from yeast

Intact genomic DNA was prepared from yeast strains essentially as outlined in Cryer *et al.* (1975). Ten mL cultures in YPD media were centrifuged (5000 rpm for 5 minutes), washed with deionised water, resuspended in 0.5 mL of lyticase buffer (0.9 M sorbitol, 0.05 M sodium phosphate buffer pH 7.5, 14 mM 2-mercaptoethanol) containing 1 mg / mL lyticase, and transferred to Eppendorf tubes. The cells were briefly vortexed and incubated at 37 °C for 30 minutes. Fifty μL of 0.5 M EDTA (pH 8), 50 μL of 10% SDS and 100 μL of 5 mg / mL proteinase K were added, mixed and the cells were incubated at 65 °C for 30

minutes. The resultant solution was extracted with 0.5 mL of phenol / chloroform and precipitated with 0.5 mL of ethanol.

2.2.1.7 Screening the yeast genomic libraries

Screening of the pEMBL Ye 23 and λEMBL3 *Saccharomyces cerevisiae* genomic libraries was performed using standard techniques (Maniatis *et al.*, 1982; Sambrook *et al.*, 1989), transforming the pEMBL Ye 23 library into *E. coli* ED8799, and using *E. coli* LE392 as the plating bacteria for the phage from the λEMBL3 library.

2.2.1.8 Pulsed-field electrophoresis of yeast DNA

Separation of chromosomes from the yeast strains DBY746, MW21.3, YPH80 and YPH149 was performed using a modification of the method of Schwartz and Cantor (1984) as outlined in the "Instruction Manual and Application Guide" supplied with the BioRad CHEF-DR II pulsed-field electrophoresis system.

2.2.1.8.1 Preparation of yeast DNA in agarose plugs

Yeast strains were grown to stationary phase in YPD media, centrifuged (3,000 rpm for 10 minutes), re-suspended in an equal volume of 50 mM EDTA (pH 8.0), and digested with lyticase at 0.5 mg / mL for 37 °C for 30 minutes. The cell suspension was mixed with molten (50°C) 1% low melting point agarose in 0.125 M EDTA (pH 7.5) to a final agarose concentration of 0.75%, pipetted into the mold chambers and allowed to set. The agarose plugs were then removed and incubated for a further 24 hours at 37 °C in LET buffer (0.45 M EDTA pH 8.0, 10 mM Tris pH 7.5, 7.5% 2-mercaptoethanol), washed in 5 mM EDTA pH 8.0, then incubated for a further 24 hours at 50 °C in NDS buffer (0.45 M EDTA pH 8.0, 10 mM Tris pH 7.5, 1% laurylsarcosine, 1 mg / mL proteinase K). The agarose plugs were then thoroughly washed and then stored in 5 mM EDTA pH 8.0.

2.2.1.8.2 Electrophoresis The agarose plug samples were loaded into the wells of a 1% agarose gel and electrophoresis was performed in 0.5 TBE buffer (45 mM Tris base, 45 mM boric acid, 1 mM EDTA) at 130V for 30 hours ramped at 80 - 150 seconds.

2.2.1.9 General molecular biology techniques

Agarose gel electrophoresis, restriction enzyme digestion and plasmid ligations were performed using standard procedures (Maniatis *et al.*, 1982; Sambrook *et al.*, 1989).

Southern transfer was performed using the traditional blotting method (Southern, 1975) with nitrocellulose or Nytran membranes as outlined by Whal *et al.* (1981), and the DNA was permanently UV cross-linked to the membranes using a Stratagene UV Stratalinker (120 mJ for 2 minutes). DNA hybridisations and subsequent washing of nitrocellulose or nytran membranes were performed using the respective methods recommended for these two membranes by Whal *et al.* (1981).

2.2.1.10 DNA slot blot analysis

Slot blot analysis was performed essentially as described by Whal *et al.* (1981). Sample and control DNAs were prepared as follows: 1 - 10 µg of DNA in 100 µL of 1 mM EDTA was added to 200 µL of 2 M NaCl / 0.1 M NaOH and incubated at 95 °C for 5 minutes, 25 µL of 3 M sodium acetate, and the samples were cooled on ice. The wells of the Milliblot^{TM}-S slot blot apparatus (Millipore) were prepared for DNA loading by the addition (under vacuum) of 0.5 mL of SSC / well, the DNA was then loaded (under vacuum) and post addition washing was performed by the addition of 200 µL of 2M NaCl. The DNA was permanently UV cross-linked to the membranes using a Stratagene UV Stratalinker (120 mJ for 2 minutes).

2.2.1.11 Techniques performed using molecular biology kits

2.2.1.11.1 5' terminal kinasing of DNA Kinasing of mutagenic oligonucleotides with ATP, and kinasing oligonucleotides in the presence of [γ-^{32}P] ATP for preparing nucleic acid probes were both performed using T4 Polynucleotide kinase with the 5' Terminal Kinasing Kit (Cat. no. TKK-1) from Bresatec Ltd., Adelaide, South Australia.

2.2.1.11.2 Preparation of overlapping sequencing clones

Preparation of overlapping sets of sequencing clones from the *PYC2* gene was performed by removing internal restriction fragments, and by the exonuclease III deletion method of

Henikoff (1987), using the Nested Deletion Kit (Cat. no. 27-1691-01) from Pharmacia LKB as outlined in Fig. 4.3a & b. Blunt ending of deletion clones prior to re-ligation was achieved using either the S1 nuclease step recommended in the kit, or by using mung bean nuclease as outlined in Sambrook *et al.* (1989). Screening for the desired deletion sizes was done using the "Lid Lysate" technique of Hoekstra (1988), and by DNA sequencing.

2.2.1.11.3 Site-directed mutagenesis Single stranded uDNA site-directed mutagenesis was performed using the BioRad Muta-gene Phagemid *In Vitro* Mutagenesis Kit (170-3576), which employs the technique of Kunkel *et al.* (1987).

2.2.1.11.4 Preparation of hybridisation probes Radioactive labelling of DNA fragments (to 10^7-10^8 cpm / μg) for use as probes in nucleic acid hybridisations was performed using the Giga-prime labelling kit from Bresatec Ltd., which employs the method of Feinberg and Vogelstein (1983). Purification of the labelled DNA fragments was achieved by phenol / chloroform extraction and ethanol precipitation.

2.2.1.11.5 DNA sequencing DNA sequencing was performed on denatured double-stranded "mini-prep" DNA (sec. 2.2.1.2) by di-deoxy chain termination method of Chen and Seeburg (1985), initially by using the T7 Sequencing Kit from Pharmacia LKB, and during the latter stages of the work by employing the protocol from the Pharmacia Kit with separate reagents, ie labelling mix containing 1.5 μM dNTPs and termination mixes for each nucleotide (80 μM dNTP, 8 μM ddNTP, 50 mM NaCl). The reactions were performed in the presence of either 5 μCi [α-^{32}P]dATP or 10 μCi [α-^{35}S]dATP and terminated by the addition of 4 μL of stop solution (95% deionised formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol). Sequencing reactions were denatured at 100 °C for 5 minutes, loaded onto a 0.3 mm 7 M urea / 6% polyacrylamide gel (25:1, acrylamide:bis-acrylamide), and run at 40 watts constant power for the required migration distance. The gel was fixed by soaking for 30 minutes in 20% ethanol / 10% acetic acid, transferred to Whatman 3MM paper and dried under vacuum at 65 °C for 1 hour. Dried gels were exposed to X-ray film at room temperature overnight (^{32}P) or for 2 - 7 days (^{35}S).

Sequencing compressions were resolved by substitution of dITP and ddITP for dGTP and ddGTP respectively in the labelling and termination mixes.

2.2.1.12 *In situ* hybridisation of bacterial colonies

In situ hybridisation of bacterial colonies was used to screen for recombinant clones containing the desired inserts during the construction of the homologous recombination plasmids T3 and H11 (see Fig. 6.1), the isolation of pET-16B-*PYC2* biotin domain recombinants (sec. 5.2.4), and in other cloning steps where the frequency of correct recombinants was expected to be low and / or screening for the desired recombinant by restriction analysis would be difficult. Colony lifts with nitrocellulose were carried out using the method of Grunstein and Hogness (1975), and the subsequent hybridisation and washing of filters were performed as previously described.

2.2.1.13 Hybridisation of the yeast chromosome mapping filters

The prime λ-clone grid filters used for the physical mapping of the *PYC* genes were hybridised with either the ^{32}P oligonucleotide labelled 527 bp *Pvu*II-*Bgl*II *PYC2* promoter fragment, or the ^{32}P labelled 655 bp *Hind*III-*Nde*I *PYC1* promoter fragment overnight at 68 °C in hyb / pre-hyb mix (0.125 M NaCl, 0.1 M Na_2HPO4, 5 mM Na_2EDTA, 1% sarkosyl, 0.1 mg / mL salmon sperm DNA). The filters were then washed for a total of 20 minutes in buffer A (1 mM Tris pH 8, 1% sarkosyl) followed by 20 minutes in 1 mM Tris pH 8.

2.2.1.14 Polymerase chain reaction

Polymerase chain reactions (PCR) were performed using 100 ng of oligonucleotide and 10 ng of template DNA in 20 µL of 1X Pyrostase buffer (supplied by Molecular Genetic Resources Inc., Florida USA) containing 1.25 mM $MgCl_2$ and 200 mM dNTP's. Twenty cycles were performed with a Perkin Elmer Cetus Thermal Cycler (Norwalk, CT, USA). The cycle conditions were as follows: denaturation, 95 °C for 2 min; annealing, 60°C for 1.5 min; and synthesis, 72 °C for 3 min.

2.2.2 Protein methods

2.2.2.1 Protein Quantitation

Protein concentrations in samples to be used in SDS-PAGE were determined by Bradford assays (Bradford, 1976), while concentrations of samples for N-terminal sequencing or mass spectrometry were determined either by the microbiuret assay (Munkres and Richards, 1965) or HPLC peak integration (Buck *et al.*, 1989), depending on the origin of the sample.

2.2.2.2 SDS-PAGE

2.2.2.2.1 Tris-glycine gels Electrophoretic analysis of Pyc levels in yeast lysates was performed under reducing conditions with a Tris-glycine buffer system and 10 x 8 cm 10% (4% stacking gels) polyacrylamide gels (w/v, 40:1, acrylamide:bisacrylamide), as described by Laemmli (1970). The gels were electrophoresed at 60 V using the Hoefer Mighty Small electrophoresis system, and stained with Coomassie Brilliant Blue-R_{250} (Laemmli, 1970).

2.2.2.2.2 Tris-tricine gels Electrophoretic analysis of the expression of the Pyc biotin domain peptides was performed under reducing conditions using the Tris-tricine method for obtaining high resolution of low molecular weight proteins (Schagger and von Jagow, 1987). The 12% Tris-tricine gels were also electrophoresed using the Hoefer Mighty Small electrophoresis system as described above.

2.2.2.3 Western blotting

Western transfer was performed at 300 mA for 1 hour in transfer buffer (10 mM Tris pH 7.2, 20 mM glycine, 20% ethanol) using the Hoefer Mighty Small Western transfer system. The nitrocellulose filters were blocked overnight in blocking solution (1% BSA, 0.05% Tween 20, 0.1 M Tris pH 7.5, 0.1 M NaCl, 5 mM $MgCl_2$) and the biotinylated proteins were detected by avidin alkaline phosphatase as described by Lim *et al.* (1987).

2.2.2.4 Laser densitometry

Quantitation of Coomassie gels and Western blots was performed by laser densitometry with a Molecular Dynamics model 300A densitometer using ImageQuant software. The nitrocellulose Western blot filters were soaked in paraffin oil to make them translucent prior to laser densitometry .

2.2.2.5 Induction of expression from the pET-16B T7 expression vector

BL21λDES clones transformed with the various pET-16B expression constructs were grown in shake flasks at 37 °C in 2xYT media supplemented with 10 μM biotin and 0.5% arabinose in the presence of 20 μg / mL chloramphenicol and 100 μg / mL ampicillin for plasmid maintenance. Overnight cultures were diluted 1:10 in fresh media and grown to an optical density at 600 nm of 0.4 - 0.5, before the addition of IPTG to a final concentration of 0.1 mM. The cells were grown for a further 3 hours before harvesting.

2.2.2.6 *In vitro* biotinylation of the Pyc2 C-terminal peptide

The IPTG induced cells expressing the Pyc2 C-terminal biotin domain peptide obtained from a 250 mL culture were washed, and resuspended in 20 mL ligase buffer (40 mM sodium phosphate pH 7.0, 1.5 mM $MgCl_2$, 5% glycerol, 0.2 mM DTE (Barker and Campbell, 1981). The cells were then lysed with a French press (12-15,000 psi) and the lysate was extracted with an equal volume of freon (1,1,1-trichloro-1,1,2-trifluoro-ethane) to remove membrane and other lipids.

In vitro biotinylation and labelling with ^{3}H-biotin was performed essentially as outlined in Chapman-Smith *et al.* (1994). A 2 mL aliquot of the induced lysate was labelled with ^{3}H-biotin by incubation for 1 hour in the presence of 3 mM ATP and 100 nM ^{3}H-biotin (35 Ci / mmol). This labelled aliquot was then combined with the rest of the lysate and the biotinylation of the expressed biotin domain peptide was forced to completion by an overnight incubation at 37 °C in the presence of 3 mM ATP and 0.5 mM unlabelled biotin. Unincorporated ^{3}H-biotin was removed by overnight dialysis at 4 °C against 50 mM MOPS pH 7.2 / 0.1 mM EDTA.

2.2.2.7 Purification of the Pyc2 C-terminal peptide

After the overnight *in vitro* biotinylation step the resultant samples were halved, diluted 1/3 with buffer A (20 mM Bis-Tris pH 6.1 / 0.1 mM EDTA), and this 30 mL sample was then loaded at 2.5 mL / min onto a 12 x 2.6 cm Q-Sepharose column (equilibrated in buffer A). The column was washed with buffer A until all the unbound proteins had been removed such that the OD_{280} trace had returned to background levels. The Pyc2 biotin domain peptide was then eluted at 2.5 mL / min with a 0 - 1.5 M NaCl gradient (0 -100% buffer B, ie buffer A + 1.5 M NaCl) over 320 min, followed by an 80 min wash with buffer B. Ten mL fractions were collected, and those containing ^3H counts were pooled and dialysed against 2 mM ammonium acetate. The peptide was concentrated by a second anion exchange step using a 1 mL Resource-Q column (Pharmacia) employing a scaled down version of the above Q-Sepharose protocol. It was then passed through (1 mL / min) a 300 mL Sephacryl S-200 gel filtration column (previously equilibrated) in buffer G (50 mM KPO_4 pH 7.0 / 0.15 M NaCl / 0.2 mM DTE / 0.1 mM EDTA) collecting 4 mL fractions. Fractions containing the Pyc2 biotin domain peptide (^3H counts) were dialysed and subjected to C4 reverse phase HPLC using a 1% / min acetonitrile gradient in 0.082% trifluoroacetic acid. The resultant peptide peaks from each run were pooled and further purified by HPLC with a 0.1% / min acetonitrile gradient.

2.2.2.8 N-terminal sequencing and mass spectrometry

Amino-terminal sequencing of the purified Pyc2 C-terminal biotin domain peptide was performed by Ms Denise Turner. The peptide was subjected to automated Edman degradation in an Applied Biosystems 470A sequencer.

Molecular weight determination by mass spectrometry was performed by electrospray ionisation at the University of Illinois with a VG Quattro mass spectrometer.

2.2.2.9 Pyruvate carboxylase enzyme assays

Pyc activity was determined in triplicate for each lysate sample by the $^{14}CO_2$ isotopic incorporation method of Utter and Keech (1963) using acetyl-CoA prepared by a modification of the succinyl-CoA synthesis procedure of Simon and Shemin (1953). The

reactions were initiated by the addition of 100 µL of pre-warmed (30 °C) assay mix (0.5 M Tris pH 8.0, 12.5 mM $MgCl_2$, 6.25 mM ATP pH 7.0, 0.625 mM acetyl-CoA, 25 mM sodium pyruvate, 25 mM $NaH^{14}CO_3$) to yeast lysate samples (20 - 50 µL) diluted to a total volume of 150 µL just prior the addition of the assay mix. After exactly 2 minutes incubation at 30 °C the reaction was terminated by the addition of 25 µL of 2 M HCl, the unincorporated acid labile counts were allowed to escape as $^{14}CO_2$, and 50 µL of the reaction mix was spotted onto Whatman 3MM paper for scintillation counting using OptiScint scintillation fluid.

The Pyc specific / biotin dependent counts were determined from the difference between the total counts and the counts obtained in the presence of 1 mg / mL avidin. The specific activity of the $NaH^{14}CO_3$ was routinely checked by scintillation counting resulting from spotting 40 µL of a 1 mM solution of the $NaH^{14}CO_3$ (40 nmol) onto dry Whatman paper that had been pre-soaked in 1% $BaCl_2$. One unit of Pyc activity = 1 µmol / min of acid stable counts (oxaloacetate).

2.2.3 Microbiological methods

2.2.3.1 *E. coli* transformations

The method chosen for the transformation of *E. coli* depended on the transformation efficiency required for that particular experiment. Routine transformations were performed by the standard $CaCl_2$ procedure (Maniatis *et al.*, 1982), while transformations requiring higher efficiencies were performed by either the "TFB-based" chemical transformation method of Hanahan *et al.* (1991), or by electroporation (Dower *et al.*, 1988) using a BioRad Gene Pulser apparatus.

2.2.3.2 Yeast transformations

Routine transformations were performed by the standard Li^+ method (Ito *et al.*, 1983), while the more efficient electroporation procedure of Becker and Guarente *et al.* (1991) was used to transform yeast strain MW21.3 in some of the experiments directed at isolating the *pyc* double null mutants (see Chapter 6).

2.2.3.3 Growth curves for the Pyc1 cysteine mutants

The growth curves for the Pyc1 cysteine mutants expressed from the multi-copy and single-copy vectors were determined in standard glucose minimal media supplemented with 8 mM aspartate (where required), 10 nM D-biotin, and 0.1 mg / mL L-tryptophan. Fifty mL cultures of each clone were simultaneously set up with a starting OD_{600} of 0.1 from fresh stationary phase cultures, and incubated in 250 mL flasks at 30 °C with orbital shaking at 180 rpm. The optical density of the cultures were monitored at regular time intervals.

CHAPTER 3

CHROMOSOMAL LOCALISATION OF THE YEAST *PYC* GENES

3.1 INTRODUCTION

The development of the technique of pulse field gel electrophoresis (PFGE) made it possible for the first time to separate very large molecular weight DNAs on the basis of size. This technique has proven to be extremely useful in separating the chromosomes of *S. cerevisiae*, thus facilitating the localisation of specific genes to a single yeast chromosome (Carle and Olson, 1984; Schwartz and Cantor, 1984). Having separated the chromosomes, the localisation of the gene of interest is then determined by hybridising a suitable DNA probe to a Southern blot containing the isolated yeast chromosomes in known positions on the filter.

In the light of the growing evidence at the outset of this project for a second *PYC* gene in yeast, the techniques of PFGE and Southern hybridisation were used as a means of a) providing further evidence for the existence of *PYC2*, and b) initiating the characterisation of "*PYC2*".

This chapter describes the:

(i) mapping of the yeast *PYC* genes to individual chromosomes,

(ii) further localisation of *PYC1* and *PYC2* on the physical map of *Saccharomyces cerevisiae*.

3.2 RESULTS AND DISCUSSION

3.2.1 Mapping of the *PYC* genes to individual chromosomes

Preliminary Southern hybridisation experiments were performed using yeast chromosomes separated by PFGE in collaboration with Dr. R. Devenish at Monash University. These experiments indicated that *PYC1* hybridised to two DNA bands which appeared to correspond to the chromosome XV / VII doublet, and either chromosome II or the closely migrating chromosome XIV (data not shown). To investigate these results further, localisation experiments were conducted with the following yeast strains: MW21.3, a *pyc1* null mutant strain (*pyc1*:: *LEU2*) with the 2.4 kb *Bgl*II fragment of *PYC1* (containing the promoter and ATP domain coding sequences; Lim *et al.*, 1988) replaced with the *LEU2*

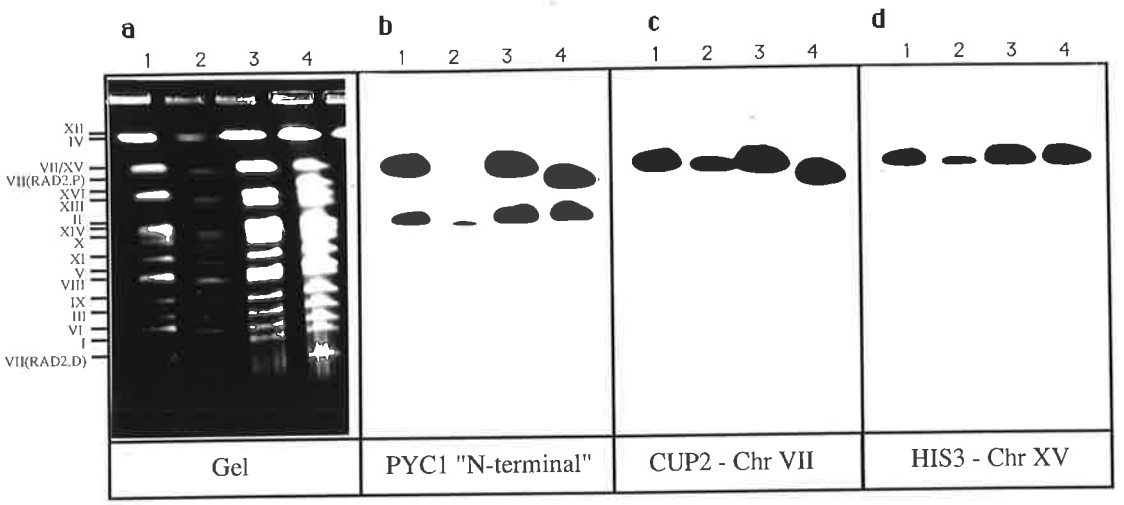
gene (Walker *et al.*, 1991); DBY746, the parental strain of MW21.3; YPH149, a *PYC*$^+$ strain which has chromosome VII split into two fragments at the *RAD2* locus (Vollrath *et al.*, 1988; see Fig. 3.1a & Fig. 3.7b); and YPH80, the parental strain of YPH149.

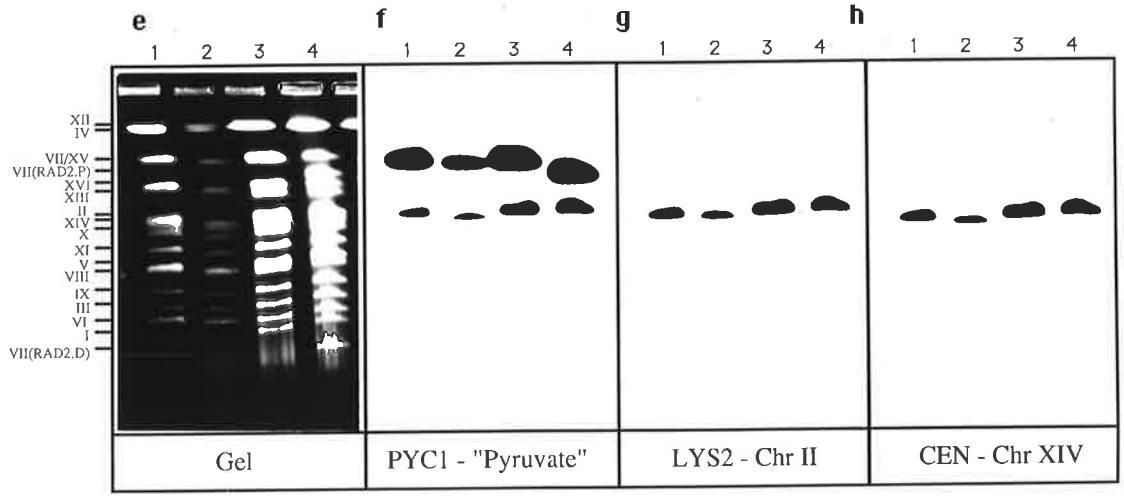
The intact chromosomes from these 4 yeast strains were loaded into adjacent lanes of a single agarose gel, separated out by pulse field gel electrophoresis, and the Southern blot filter from this gel was probed with the 2.4 kb "N-terminal" *Bgl*II fragment of *PYC1*. Two chromosomal bands of differing intensity were visible in the lanes from strains DBY746, YPH80 and YPH149 (Fig. 3.1b). The strongest hybridisation signal corresponding to the *PYC1* gene aligned with the position of the chromosome XV / VII doublet in DBY746 and YPH80, while its position in the lane containing YPH149 chromosomes clearly demonstrated that the *PYC1* gene is located on chromosome VII proximal to the *RAD2* gene (Fig. 3.1a & b). Further confirmation of this assignment was obtained by reprobing the same filter with the *CUP2* gene as a marker for chromosome VII (Welch *et al.*, 1989) and the *HIS3* gene as a marker for chromosome XV (Struhl *et al.*, 1979) (see Fig. 3.1c & d). These experiments (alignment of the figures 3.1a - d) revealed that the *PYC1* gene, like the *CUP2* gene, is located on the *RAD2* proximal (RAD2.P) fragment of chromosome VII, rather than on chromosome XV (the band hybridising to the *HIS3* gene) which in strains other than YPH149 co-migrated with chromosome VII. The chromosome XV / VII doublet from MW21.3 did not hybridise to this "N-terminal" *PYC1* probe (Fig. 3.1b) as the DNA sequences complementary to this probe had been removed in the construction of this null mutant. The weaker hybridisation at chromosome II (Fig. 3.1b) indicates that yeast contains a second gene for pyruvate carboxylase (*PYC2*) which has some sequence similarities to *PYC1*.

To ensure that the hybridisation signals produced by the "N-terminal" probe were indeed due to pyruvate carboxylase genes, the filter was re-probed with a second DNA probe from the *PYC1* gene, ie the 1.04 kb *Bam*HI-*Hinc*II fragment. This fragment encodes for the pyruvate domain which was known from sequence comparisons to be unique to biotin enzymes which bind either pyruvate or oxaloacetate (Pyc, Tc, and Odc; Lim *et al.*, 1988). As yeast do not contain Tc or Odc, this probe was expected to be specific for the *PYC* genes. The two hybridisation signals resulting from the hybridisation with this "pyruvate domain"

Figure 3.1 Chromosomal localisation of *PYC1* and *PYC2*

Intact chromosomal DNA prepared from (**1**) DBY746, (**2**) MW21.3, (**3**) YPH80, and (**4**) YPH149 was separated by pulse field gel electrophoresis as described in Chapter 2. The resultant gel was stained with ethidium bromide and photographed (panels **a** & **e**) to enable alignment with the autoradiographs from Southern hybridisation. The DNA was transferred to Nytran and subsequently probed, stripped, and reprobed such that the filter had been probed separately with each of the following probes; (**b**) 2.4 kb *Bgl*II N-terminal fragment of *PYC1*, (**c**) 2.1 kb *Dra*I fragment containing the *CUP2* gene from chromosome VII (Welch *et al.*, 1989), (**d**) 1.76 kb *Bam*HI fragment containing the *HIS3* gene from chromosome XV (Struhl *et al.*, 1979), (**f**) 1.04 kb *Bam*HI-*Hinc*II *PYC1* "pyruvate domain" fragment, (**g**) 6.9 kb *Eco*RI-*Pst*I fragment containing the *LYS2* gene from chromosome II (Eibel and Philippsen, 1983), (**h**) 1.4 kb *Eco*RI-*Hind*III *CEN* fragment containing centromere sequences from chromosome XIV (Carle and Olson, 1985). The order of chromosomes are as indicated, with the exact positions referring to the separated chromosomes from strain YPH80 in lane (**3**).


a
b
c
d
1 2 3 4
XII
IV
VII/XV
VII(RAD2.P)
XVI
XIII
II
XIV
X
XI
V
VIII
IX
III
VI
I
VII(RAD2.D)
Gel
PYC1 "N-terminal"
CUP2 - Chr VII
HIS3 - Chr XV


e
f
g
h
1 2 3 4
XII
IV
VII/XV
VII(RAD2.P)
XVI
XIII
II
XIV
X
XI
V
VIII
IX
III
VI
I
VII(RAD2.D)
Gel
PYC1 - "Pyruvate"
LYS2 - Chr II
CEN - Chr XIV

probe aligned with those observed with the "N-terminal" probe (see Fig. 3.1f), providing further evidence that they correspond to the *PYC* genes. In addition, a band was observed for *PYC1* (chromosome VII) with the null mutant MW21.3 as the pyruvate domain is intact in this mutant.

The physical mapping of *PYC2* to chromosome II was confirmed by re-probing the filter with the *LYS2* gene as a probe for chromosome II (Eibel & Philippsen, 1983), and centromere sequences from chromosome XIV (Carle and Olson, 1985). These experiments (alignment of figures 3.1e - h) revealed the *PYC2* gene was indeed located on chromosome II, and not on the closely migrating chromosome XIV.

In conclusion, this series of experiments (reported in Walker *et al.*, 1991) clearly demonstrates that there are two *PYC* genes in yeast, *PYC1* which is located on chromosome VII, and *PYC2* which is located on chromosome II.

3.2.1 Further localisation of the *PYC* genes

To refine the localisation of the *PYC* genes, specific probes for *PYC1* and *PYC2* were hybridised to a set of prime λ-clone grid filters (kindly supplied by the M. V. Olson lab.) containing an ordered set of λ–clones of known location on the physical map of *S. cerevisiae* (Riles *et al.*, 1993). The *PYC2* probe used in the following experiments was obtained from the 5' clone described in Chapter 4.

3.2.1.1 The *PYC2* gene

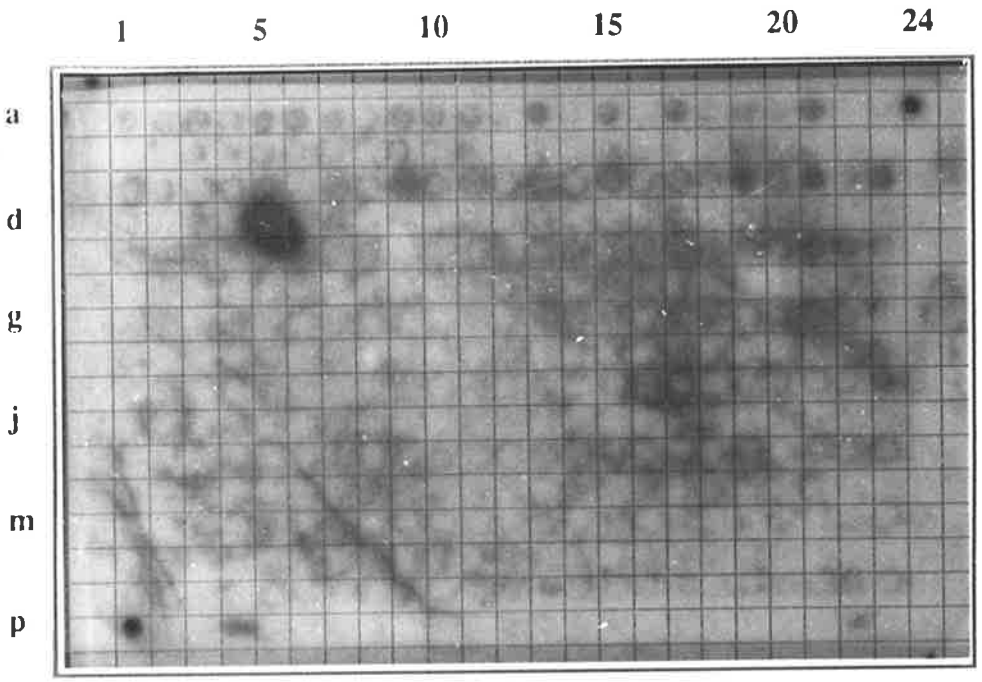
The grid filters were hybridised with a 527 bp *Pvu*II-*Bgl*II promoter fragment of *PYC2* (gene specific), which was found to bind to filter 2 position D5 (Fig. 3.2a) corresponding to clone 5564 (see Appendix A) from the right arm of chromosome II (Fig 3.3). The flanking clones which partially overlap with the proximal and distal ends of clone 5564 are clones 6471 (known to contain the *CDC28* gene), and 6734 respectively (Fig. 3.3). The *PYC2* promoter fragment did not hybridise to clone 6471, but it did hybridise to the DNA at position M7 on a supplementary filter (also supplied by the Olson lab.) (Fig. 3.2b) corresponding to clone 6734. Hence these results suggest that the *PYC2* gene is

Figure 3.2 Fine mapping of *PYC2*

The three standard and one supplementary "Olson" prime λ-clone grid filters were hybridised to a 527 bp *Pvu*II-*Bgl*II gene specific promoter fragment from *PYC2* . Part **(a)** shows that the *PYC2* promoter fragment hybridised to the "Olson" clone (5564) located at position D5 on filter 2, and part **(b)** shows the hybridisation to the "Olson" clone (6734) located at position M7 on the supplementary filter. The smaller dark spots in the top corners, and botton left hand corner of the filters were the result of the spots of radioactive ink used to orientate the filters.

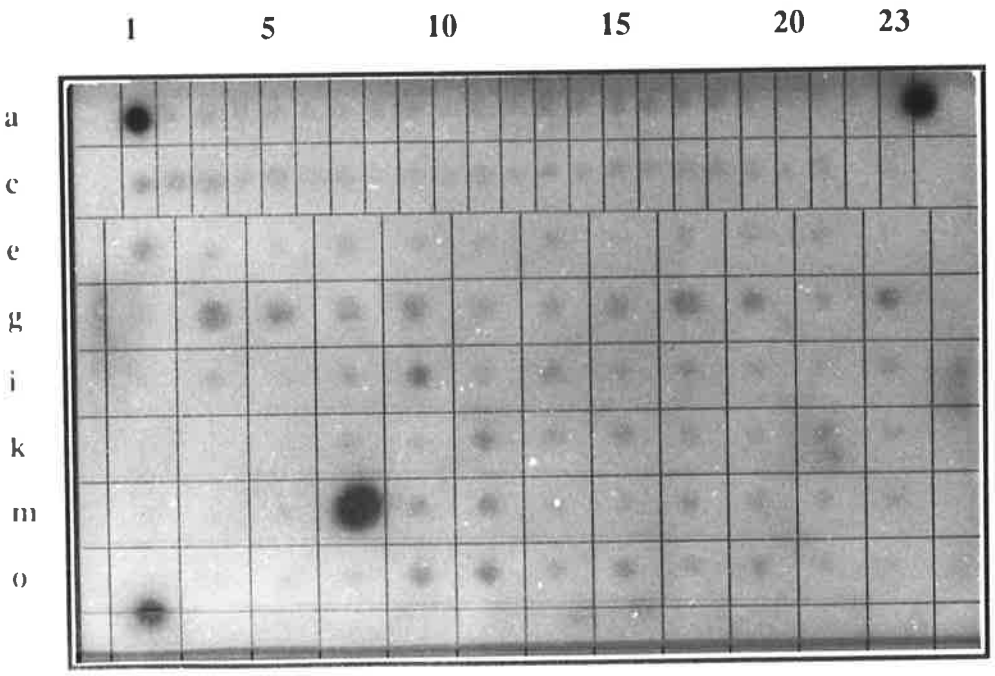
a)

Filter 2



b)

Supplementry filter



contained within the small region of DNA in common to clones 5564 and 6734 on the right arm of chromosome II distal to the *CDC28* gene (Fig. 3.3).

In the same year that we reported the chromosomal localisation of the *PYC* genes by PFGE, Stucka *et al.*(1991) later reported the same findings, also using PFGE. In addition, these workers reported that *PYC2* was located just distal of the *DUR1,2* loci, *ie.* on a cosmid clone (c411) which overlaps the distal end of another clone that they had earlier found to contain the *DUR1,2* loci and a $tRNA^{Glu3}$ gene. Feldman *et al.* (1994) have since published the complete DNA sequence of chromosome II assembled from the sequence data contributed by a consortium of 28 European laboratories, including the *PYC2* sequence data provided by R. Stucka. According to their map, the *PYC2* gene is approximately 15 kb distal of the *DUR1,2* loci (see Fig 3.4).

Figure 3.7a shows the comparison of the physical location of genes on chromosome II that have been determined by hybridisation to the prime λ-clone grid filters prepared by Dr. M. V. Olson's laboratory (unpublished data provided by Dr. L. Riles; included in the Stanford data base), and by the complete sequencing of chromosome II from strain S288C (Feldman *et al.*, 1994). This shows that the *PYC2* localisation using the "Olson filter technique" described in this chapter coincided exactly with the location revealed more recently by DNA sequencing. Indeed, with the exception of *CDC28*, there was a close match between the locations determined for all the genes which had been localised by both techniques, providing good evidence for the accuracy of the "Olson filter" hybridisation technique. In future though, judging from the speed at which the yeast genome project is progressing (chromosome III, Oliver *et al.*, 1992; chromosome XI, Dujon *et al.*, 1994; chromosomes II, Feldman *et al.*, 1994;), this technique may no longer be required as the entire yeast genome will soon be sequenced.

3.2.1.1 The *PYC1* gene

As the *PYC1* gene had not previously been localised to a specific region of chromosome VI, the same grid filters were also used to further localise this gene. When the grid filters were hybridised with a 655 bp *Hind*III-*Nde*I promoter fragment from *PYC1*, the fragment hybridised to filter 2 position P4 and filter 3 position K21 (Fig. 3.5a & b),

Figure 3.3 "Olson" contig map of chromosome II

Unpublished contig map of chromosome II showing the position of each of the prime λ-clones along the length of chromosome II (prepared by the research group of Dr. M. V. Olson; Riles *et al.*, 1993). The short vertical lines extending above and below the central horizontal line showing either the length of each λ-clone, or the length of the whole chromosome, indicate the positions of the *Eco*RI-*Hind*III (RH) fragment boundaries used to construct the contig map of chromosome II. Unordered fragments within an RH fragment group are shown by the short vertical lines that extend above the horizontal line. The arrow indicates the position of the overlapping clones which hybridised to the *PYC2* probe (see Fig. 3.2).

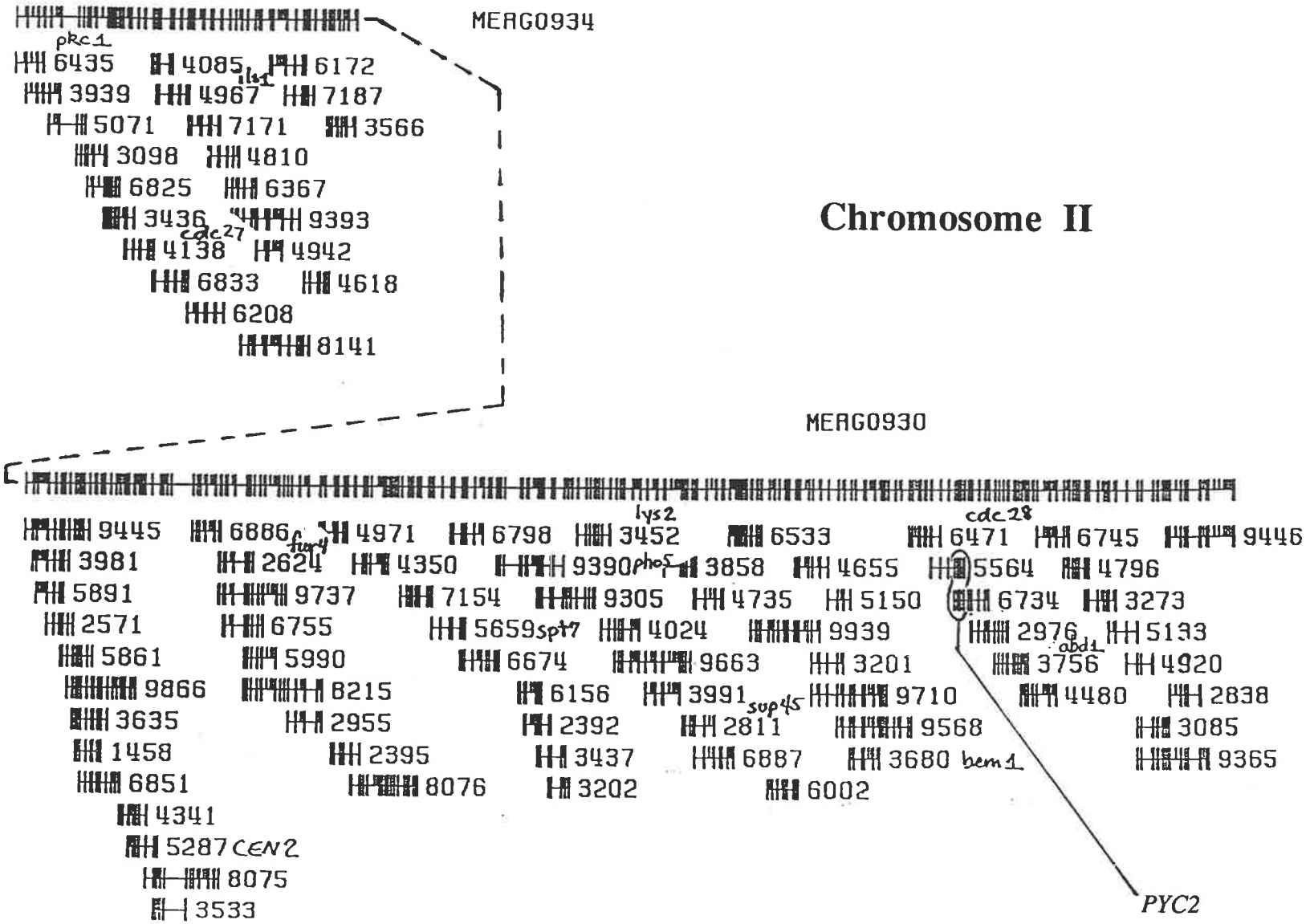

MERG0934
pkc1
6435
4085
lst1
6172
3939
4967
7187
5071
7171
3566
3098
4810
6825
6367
3436
cdc27
9393
4138
4942
6833
4618
6208
8141
Chromosome II
MERG0930
9445
6886
fun4
4971
6798
lys2
3452
6533
cdc28
6471
6745
9446
3981
2624
4350
9390
pho5
3858
4655
5564
4796
5891
9737
7154
9305
4735
5150
6734
3273
2571
6755
5659
spt7
4024
9939
2976
abd1
5133
5861
5990
6674
9663
3201
3756
4920
9866
8215
6156
3991
sup45
9710
4480
2838
3635
2955
2392
2811
9568
3085
1458
2395
3437
6887
3680
bem1
9365
6851
8076
3202
6002
4341
5287
CEN2
8075
3533
PYC2

Figure 3.4 The complete map of chromosome II, adapted from Feldman *et al.* (1994)

Map (to scale) indicating the positions of the known genes on chromosome II as determined by sequencing of the complete chromosome (Feldman *et al.*, 1994). The scale shown by the top line is in kilobase pairs, the genetic elements on the two strands are shown as coloured bars, the 410 identified open reading frames (ORFs) are shown as blue and purple boxes, and the ORFs corresponding to known genes are shown by the black bars. Ty elements are shown as green bars, while the red bars indicate the position of a tRNA genes symbolised by a t and the one-letter code for the amino acid accepted. The arrow highlights the position of the *PYC2* gene.

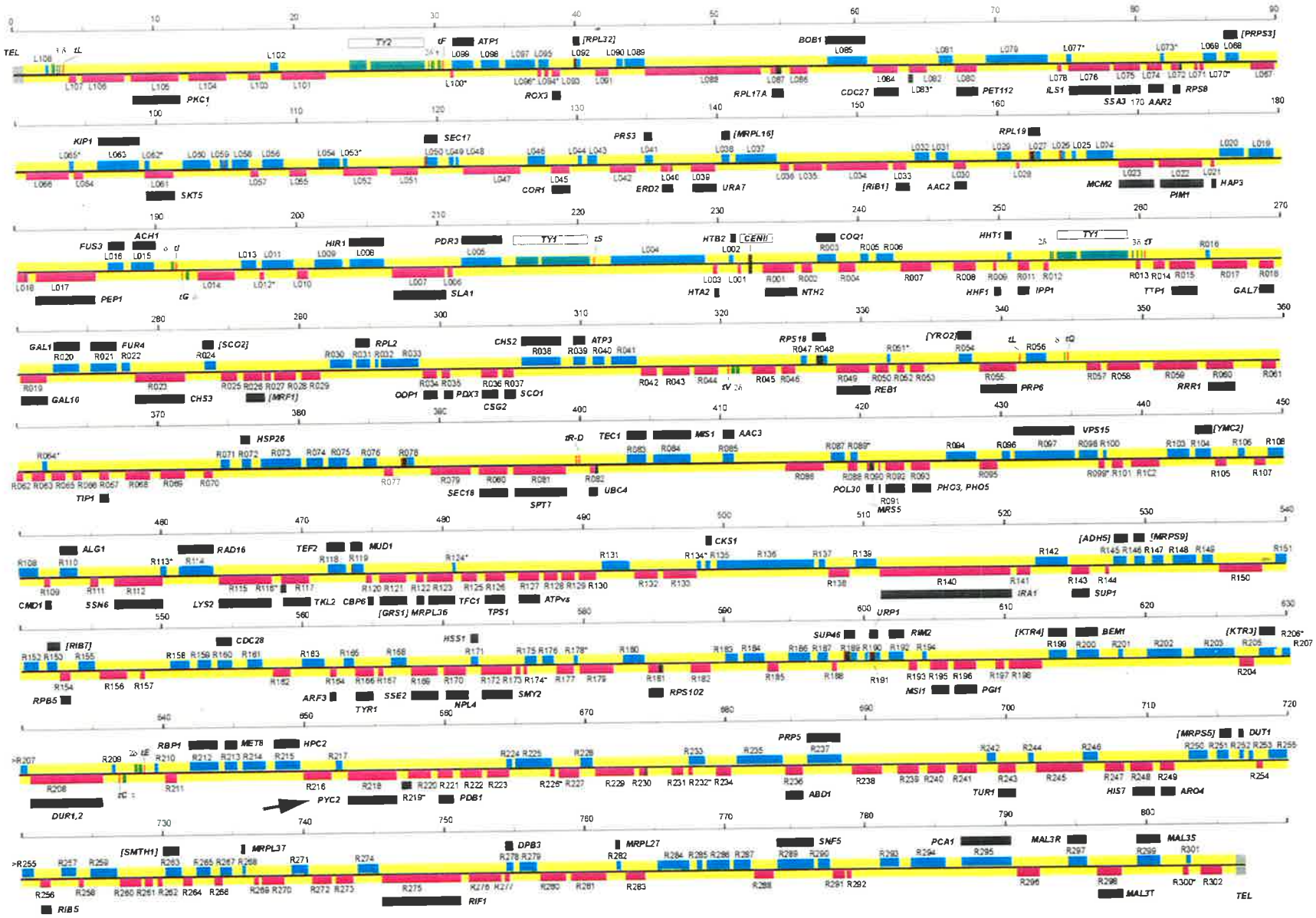
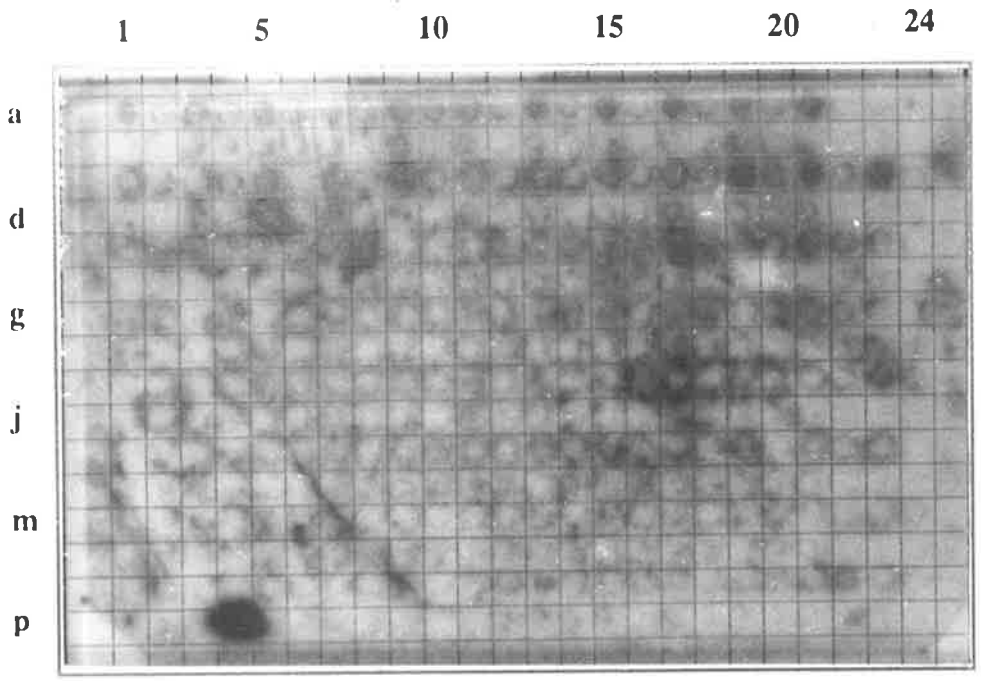


Figure 3.5 Fine mapping of *PYC1*

The three standard "Olson" prime λ-clone grid filters were hybridised to a 655 bp *Hind*III-*Nde*I gene specific promoter fragment from *PYC1*. Part (**a**) shows the hybridisation of the *PYC1* promoter fragment to the "Olson" clone (6619) located at position P4 on filter 2, and part (**b**) shows the hybridisation to the "Olson" clone (7086) located at position K21 filter 3.

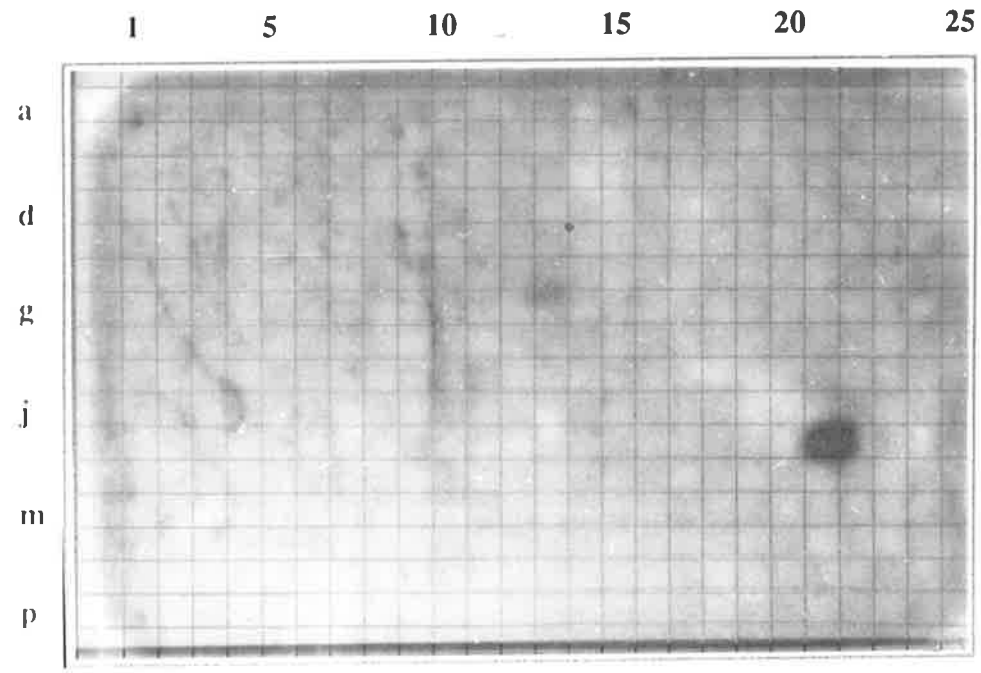
a)

Filter 2



b)

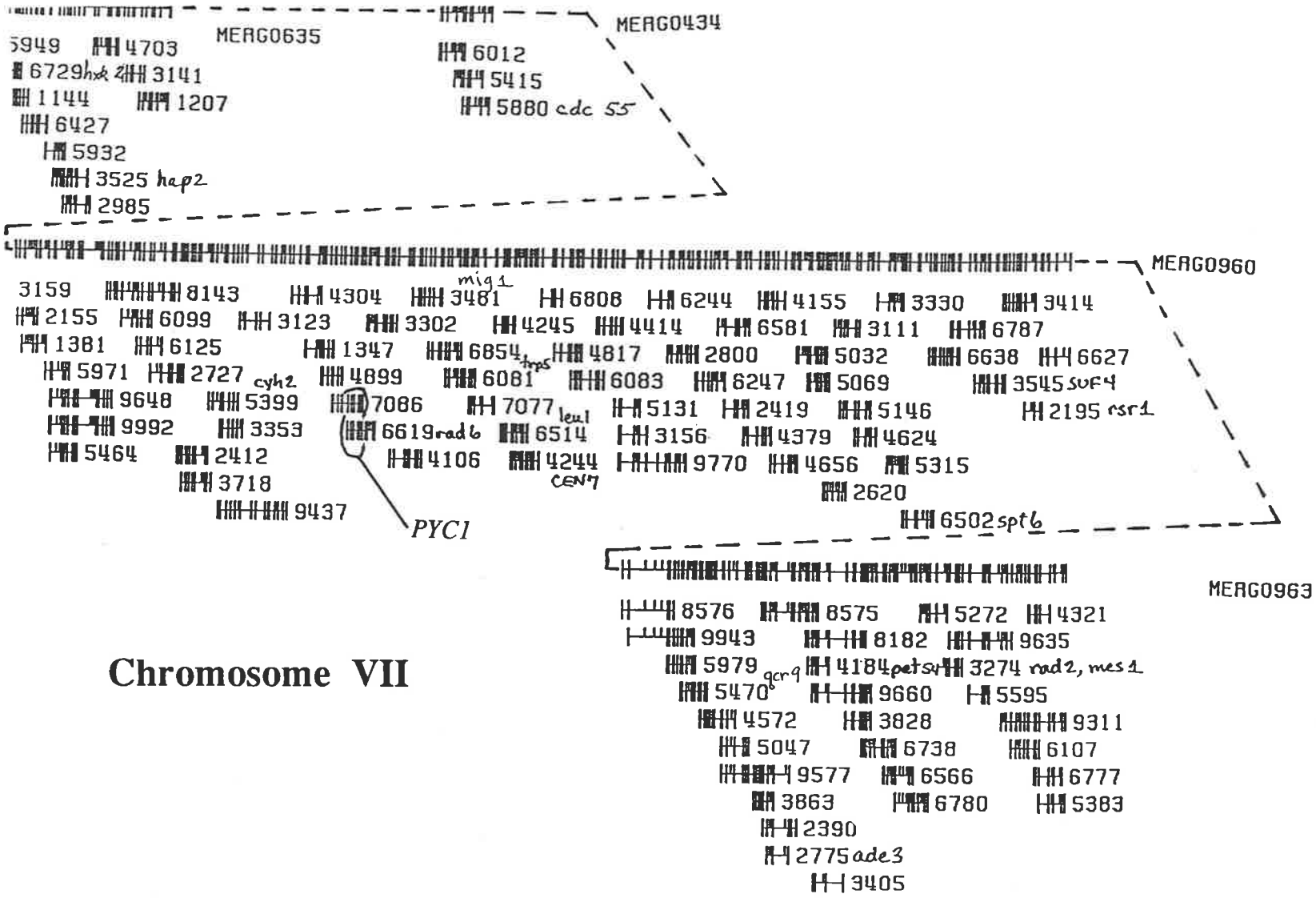
Filter 3



corresponding to the overlapping λ clones 6619 and 7086 respectively (see Appendix B, Fig. 3.6). Therefore the *PYC1* gene is contained within the region of DNA in common to both of these clones, and as the *RAD6* gene has previously been localised to clone 6619, the *PYC1* gene is either at, or very close to the *RAD6* locus (Fig. 3.7b).

Figure 3.6 "Olson" contig map of chromosome VII

Unpublished contig map of chromosome VII showing the position of each of the prime λ-clones along the length of chromosome VII (prepared by the research group of Dr. M. V. Olson; Riles *et al.*, 1993). The short vertical lines extending above and below the central horizontal lines showing either the length of each λ-clone, or the length of the whole chromosome, indicate the positions of the *Eco*RI-*Hind*III (RH) fragment boundaries used to construct the contig map of chromosome II. Unordered fragments within an RH fragment group are shown by the short vertical lines that extend above the horizontal line. The arrow indicates the position of the overlapping clones which hybridised to the *PYC1* probe (see Fig. 3.5).

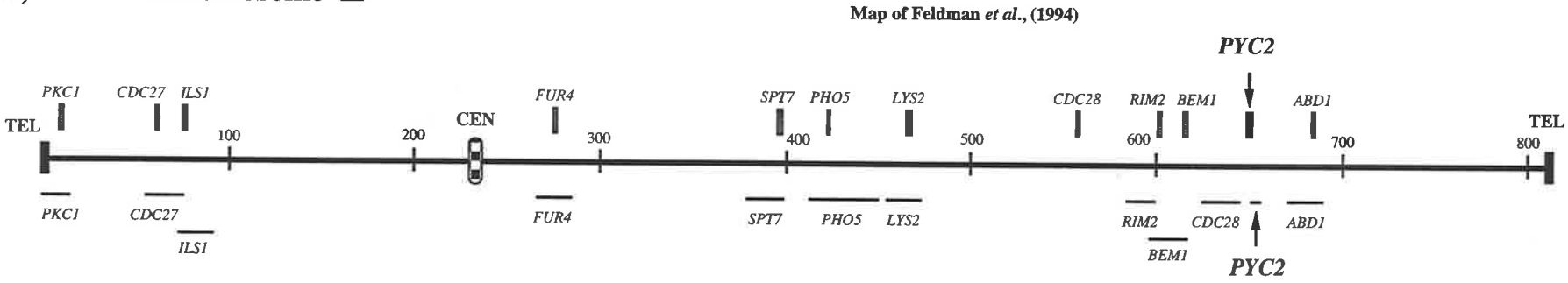


Chromosome VII

Figure 3.7 Physical maps of chromosomes II and VII

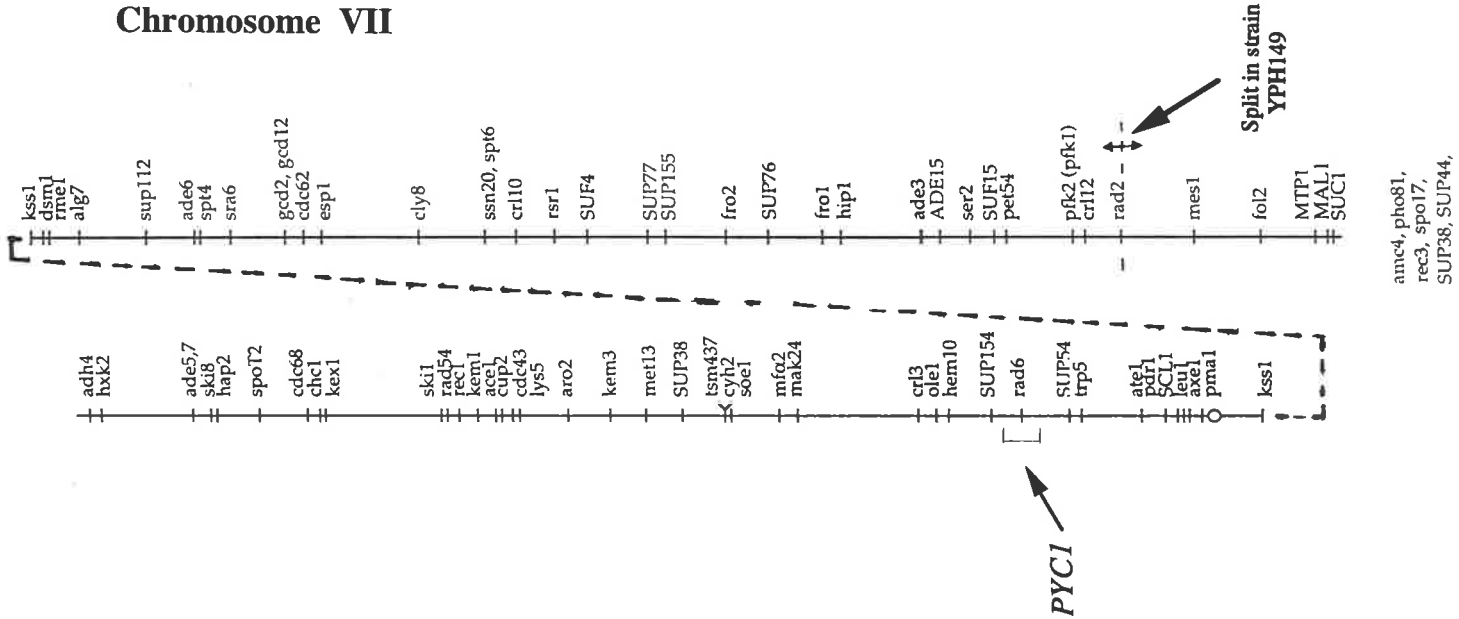
(a) Comparison of the gene locations on chromosome II determined by hybridisation, with the locations known by DNA sequencing. The vertical shaded blocks above the central horizontal line indicate the positions of the indicated genes in the DNA of chromosome II (Fig. 3.4; Feldman *et al.*, 1994), while the short horizontal lines in the lower half of the diagram indicate the positions of the same genes as determined by hybridisation to the "Olson" grid filters. The scale is in kilobase pairs, and the arrows indicate the position of *PYC2* as determined by DNA sequencing (top arrow), and by the hybridisation results presented in figure 3.2 (bottom arrow).
(b) The physical map of chromosome VII. From Mortimer *et al.* (1989). The bottom arrow indicates the position of *PYC1* as determined by the results presented in figures 3.5 and 3.6, while the top arrow indicates the position of the *RAD2* locus, the site at which chromosome VII is split into two fragments (RAD2.P & RAD2.D) in the yeast strain YPH149.

a) **Chromosome II**



Unpublished map from Dr. M. V. Olson

b) **Chromosome VII**



CHAPTER 4

ISOLATION AND SEQUENCING OF THE YEAST *PYC2* GENE

4.1 INTRODUCTION

The vast majority of the studies that have been conducted on the structure and function of pyruvate carboxylase have been carried out with the enzymes from the livers of chicken and sheep. This has primarily been for practical reasons, namely the abundance of the enzyme in vertebrate liver mitochondria, and the ready availability of these two domestic species. However, the yeast *S. cerevisiae* was the organism first chosen for gene studies because of its lack of introns, ease of genetic manipulation, and the simplicity of controlling the carbon sources and other growth conditions when monitoring the effect of any introduced mutations. *E. coli* could not be used as this organism does not contain pyruvate carboxylase.

The initial isolation of a yeast *PYC* clone (3' clone) was performed in our laboratory by Dr. P. Morris and co-workers (Morris *et al.*, 1987). These workers screened a yeast genomic library with a mixture of degenerate oligonucleotides (14 mers) specific for the 5 amino acids encompassing the biotin attachment site, which they had determined by direct protein sequencing. A 5' clone including the promoter and most of the coding region was subsequently isolated and sequenced in our laboratory by Dr. F. Lim (Lim *et al.*, 1988)

Upon discovering that yeast contains two *PYC* genes (see Chapter 1, 1.7 & Chapter 3) it became clear that studying the structure / function relationships of yeast Pyc using yeast as the host for expression of the mutant Pyc constructs would require the isolation, characterisation, and disruption of *PYC2*. Hence, one of the primary goals of my research project has been to isolate and characterise this gene.

At the outset of this study only partial amino acid sequences had been determined for pyruvate carboxylases from organisms other than yeast. A direct protein sequencing approach had been used to sequence tryptic peptides (19 - 24 amino acids) containing the amino acids encompassing the biotin attachment site from chicken, turkey and sheep (Rylatt *et al.*, 1977). Also the sequence of the C-terminal 44 amino acids had been determined for the human enzyme from a partial cDNA clone (Freytag and Collier, 1984). Finally, the sequence of the C-terminal 750 amino acids (64 %) of the rat enzyme had been obtained

from a number of cDNA and genomic clones by the combined efforts of Dr's. I. A. Cassady (1987) and G. W. Booker (1990) from our laboratory (unpublished results).

However, during the course of this study a number of complete Pyc sequences were reported. Zhang *et al.* (1993) published the cDNA sequence of mouse Pyc, and several groups have determined the cDNA sequence of the human enzyme. In March of 1994 a cDNA sequence for human Pyc was entered into the data base, in July of the same year a different group published a sequence from their human cDNA clone(s) (MacKay *et al.*, 1994), and Dr M. E. Walker from our laboratory has also determined the human cDNA sequence (Walker *et al.*, 1995). Finally, in September of 1994 the first sequence of an insect Pyc was reported when the cDNA sequence of mosquito Pyc was entered into the data base.

The most influential report with regard to my research was a paper published in October of 1991 which contained the sequence of the yeast *PYC2* gene (Stucka *et al.*, 1991). At that stage significant progress had been made in sequencing the *PYC2* clone that I had isolated, and for this and other reasons outlined in this chapter I decided to complete the sequencing of this gene rather than request the clone from these workers.

This chapter describes the:

(i) isolation, restriction mapping, and sequence analysis of a 5' *PYC2* clone (pBDV2)

(ii) isolation and sequencing of the 3' end of the gene

(iii) sequence comparisons with the published *PYC2* gene.

(iv) sequence comparisons with the Pyc1 and other functionally similar enzymes

4.2 RESULTS

4.2.1 Isolation of *PYC2* clones

Having found that the *PYC2* gene was located on chromosome II (Walker *et al.*, 1991), the first approach used in attempting to isolate the *PYC2* gene involved the preparation of a chromosome II specific library. The strategy chosen to prepare the library involved partial *Sau*3A digestion and subsequent cloning of chromosome II DNA. The DNA had been obtained from a low melting point agarose gel containing the yeast chromosomes

separated by pulse field gel electrophoresis. As it is technically simpler to isolate total genomic DNA, the pilot partial *Sau*3A digestion experiments were carried out on total wild type (DBY746) yeast genomic DNA. However, technical difficulties in obtaining sufficient intact DNA from the individual chromosomes resulted in this chromosome II library strategy being abandoned for a more straightforward approach.

As Dr. M. E. Walker in our laboratory had recently constructed a *pyc1* null mutant strain of yeast (MW21.3), another obvious approach to the isolation of the *PYC2* gene involved the preparation of a genomic library (partial *Sau*3A) from MW21.3. The library could then be screened with the 2.4 kb 5'*Bgl*II fragment of *PYC1* which been deleted from the genome of MW21.3, thus ensuring that no *PYC1* clones would be isolated by the screening procedure. However once again technical problems in generating a large enough library hampered my progress. Consequently, whilst continuing with this approach, an existing genomic library from a wild type laboratory strain of *Saccharomyces cerevisiae*.was screened.

The wild type genomic library screened using this third method had been prepared by Dr. P. Morris of our laboratory, and contained partial *Sau*3A fragments cloned into the plasmid EMBL Ye23. The library was screened with the above mentioned 2.4 kb *Bgl*II *PYC1* fragment with the colonies plated out at a density of 1-3 x10^3 colonies / plate. An unexpectedly large number of hybridising colonies were identified in this first round of screening, leading to my suspicion that this library had somehow been contaminated with a plasmid bearing a *PYC1* clone. Regardless of this suspicion 10 duplicate first round "positives" were picked and 5 of these carried through to the end of the second round of screening. As the colonies were sufficiently well spaced to allow the individual "positive" colonies to be picked with confidence, the usual third round screening step was omitted. Instead, the 5 "positive" clones and a number of the surrounding "negative" clones were picked, and plasmid DNA "mini-preps" were prepared.

Slot blot analysis was performed using equal amounts (5 µg) of DNA from the five "positive" clones, a range of different "negative" clones, and a positive control plasmid containing the whole of the *PYC1* gene sequence (pMW4A). By comparing the hybridisation signals produced after probing the slot blot filter with the 2.4 kb *Bgl*II *PYC1*, and washing

at a range of different temperatures (50, 70, and 80 °C), four of the five "positive" clones were found to have an equal affinity for the probe which was noticeably less than that of the control *PYC1* DNA (data not shown). The DNA from the fifth clone did not appear to hybridise to the *PYC1* probe. This suggested that despite my earlier suspicion, these four "positive" clones (EMBY1 - 4) quite possibly all contained the *PYC2* gene.

4.2.2 Restriction analysis of the *PYC2* clones

Large scale DNA preps of clones EMBY1 - 4 were prepared and further compared by Southern analysis after restriction with the enzymes *Eco*RI, *Eco*RI-*Hin*dIII, and *Pst*I (data not shown). Clones EMBY1, 2 and 4 appeared to be identical, indicating that previous amplification of this library had somehow selectively amplified this particular clone, while clone EMBY3 no longer seemed to hybridise to the *PYC1* probe.

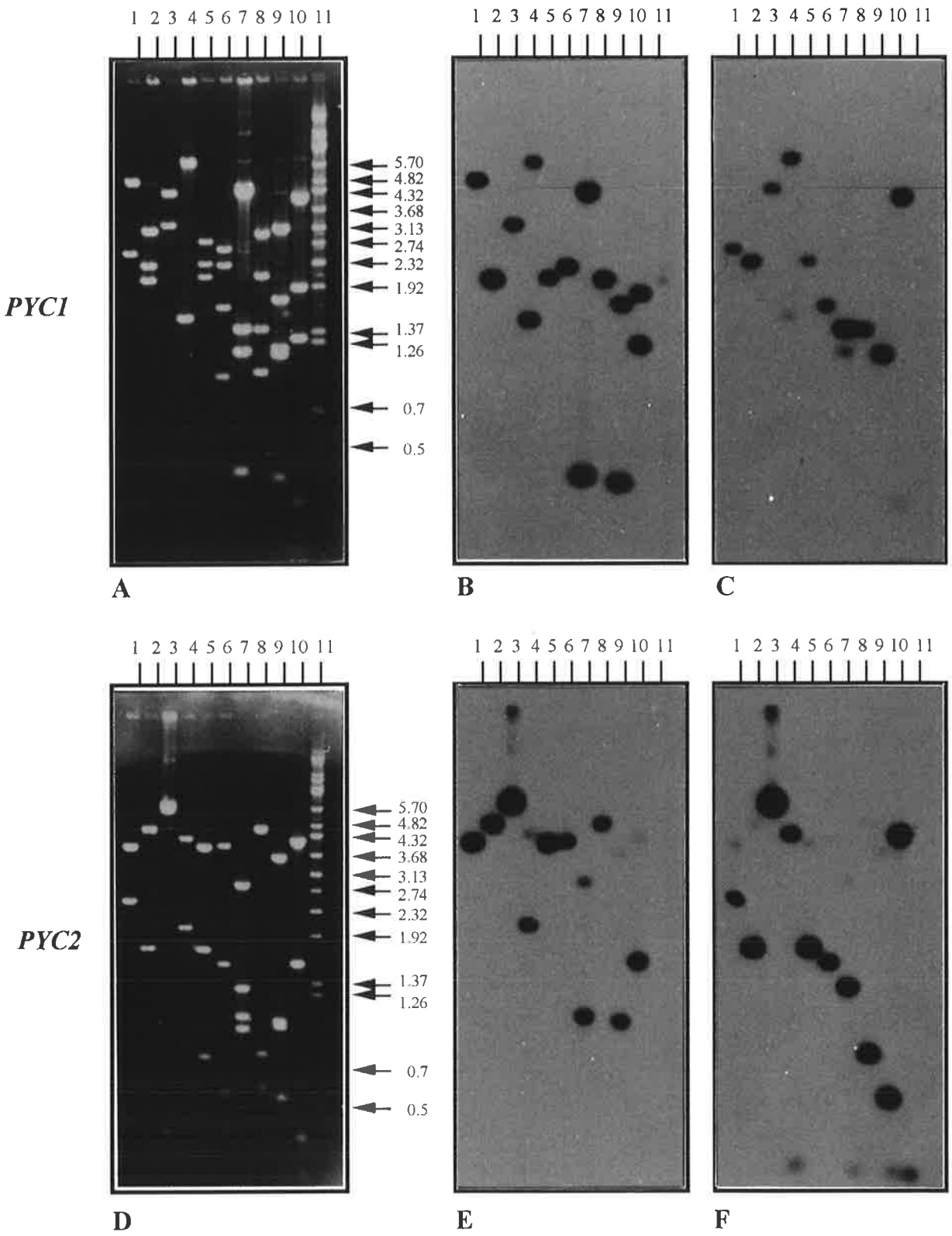
To facilitate further mapping and subsequent sequencing, the approximately 3.6 kb insert from clone EMBY2 was subcloned as a *Sma*I-*Hind*III fragment into KS+ bluescript (Stratagene). The resultant clone (pBDV2) and a control plasmid containing the whole of the *PYC1* gene (pMW4A) were each digested with a more extensive range of restriction enzymes, and the Southern blots obtained from the two gels were probed with either the 2.4 kb *Bgl*II 5' *PYC1* probe, or a "pyruvate domain" probe containing sequences believed to be entirely unique to Pyc as yeast do not contain the Tc or Odc enzymes (Fig. 4.1). Once again the increased autoradiographic exposure time required for the pBDV2 fragments suggested that this insert contained a different *PYC* gene, in agreement with the different pattern of restriction fragments observed between pBDV2 and pMW4A (see Fig. 4.1). There were however some sites that appeared to be common to both genes. Taken together, the hybridisation and restriction site data suggested that the clone that had been isolated was most likely the 5' end of the *PYC2* gene.

4.2.3 Sequencing of the 5' *PYC2* clone

The 5' prime *PYC2* clone that had been isolated from the EMBL Ye23 genomic library was sequenced in both directions using the strategy outlined in figure 4.2. This

Figure 4.1 Restriction mapping and Southern analysis of the 5' *PYC2* clone.

The 5' "putative" *PYC2* clone pBDV2, and the plasmid pMW4A containing the full length *PYC1* gene were each digested with the same range of restriction enzymes, electrophoresed under identical conditions on a 0.8% agarose gel, and stained with ethidium bromide (panels A and D respectively). Southern transfer was performed on both gels and the resultant nitrocellulose filters were hybridised with the 1.04 kbp *Bam*HI-*Hinc*II *PYC1* "pyruvate domain" probe (panels B and E), stripped, and then re-probed with the 2.4 kbp *Bgl*II *PYC1* "N-terminal probe" (panels C and F). **Lanes**: **1**, *Bam*HI digest; **2**, *Bgl*II digest; **3**, *Eco*RI digest; **4**, *Kpn*I digest; **5**, *Bam*HI-*Bgl*II digest; **6**, *Bam*HI-*Eco*RI digest; **7**, *Bam*HI-*Kpn*I digest; **8**, *Bgl*II-*Eco*RI digest; **9**, *Bgl*II-*Kpn*I digest; **10**, *Eco*RI-*Kpn*I digest; **11**, SPP1 DNA size markers.


RESTRICTION DIGESTS
SOUTHERN ANALYSIS
Pyruvate domain probe
5' end probe
PYC1
PYC2
1 2 3 4 5 6 7 8 9 10 11
5.70
4.82
4.32
3.68
3.13
2.74
2.32
1.92
1.37
1.26
0.7
0.5
A
B
C
D
E
F

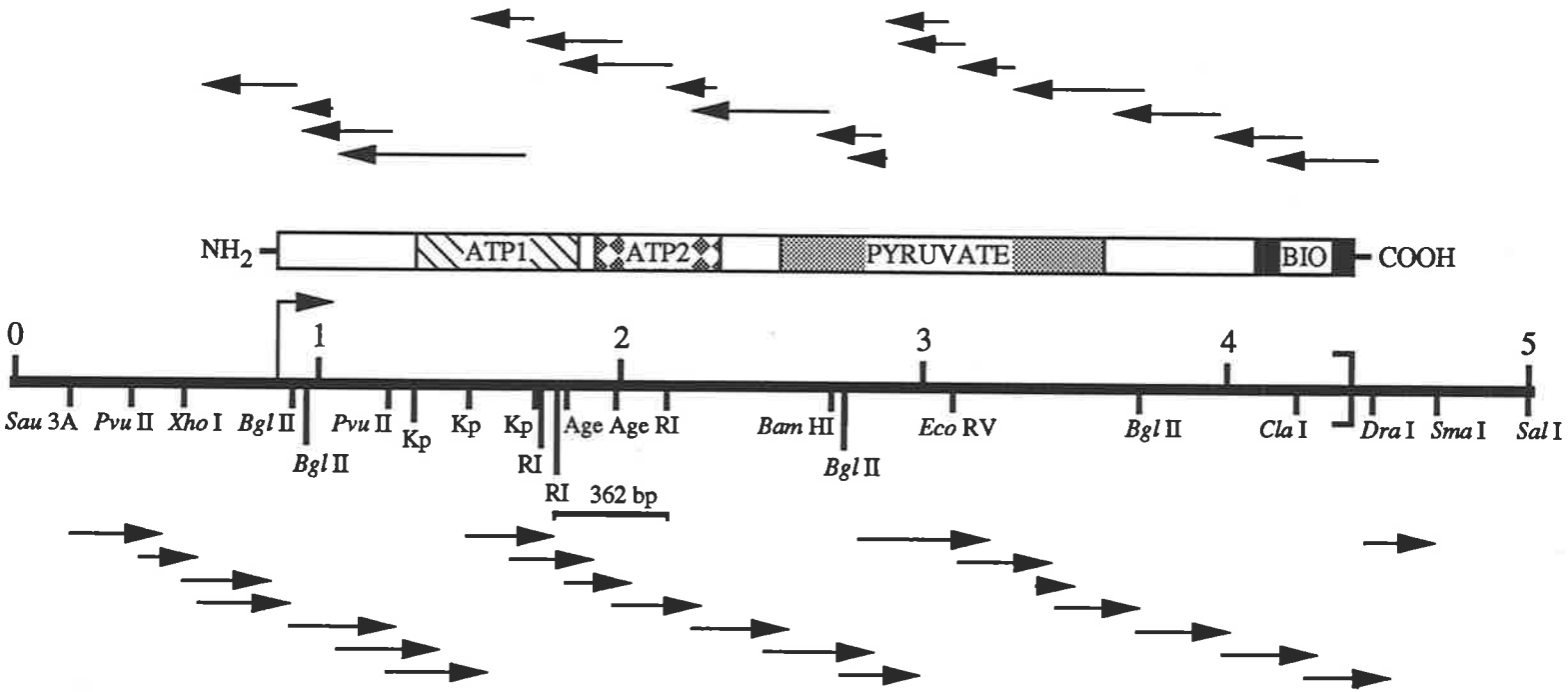


Figure 4.2 Sequencing strategy for the *PYC2* gene.

Arrows indicate the direction, position and length of the sequences used in compiling the final sequence presented in figure 4.4. The numbers on the central scale are in kbp, and the positions of the translation start and termination sites are indicated by the bent arrow and square brackets respectively. The short line below the central scale spanning the two RI sites indicates the position of the 362 bp *Eco*RI fragment used as a probe to isolate the 3' *PYC2* clone from a λEMBL3 genomic library. The shaded sections in the diagram above the scale indicates the positions of the structural domains in the protein corresponding to those determined for Pyc1 (Lim *et al.*, 1988). Abbreviations: Kp, *Kpn*I; RI, *Eco*RI; Age, *Age*I.

involved the combination of two approaches, that of sequencing from the natural restriction sites by either sub-cloning or removing various internal restriction fragments, and a nested deletion approach (Henikoff, 1987).

Sets of nested deletion clones were obtained from both ends of pBDV2 using increasing periods of exonuclease III digestion to remove progressively more DNA from the end having the 5' overhang, with the other end being protected by a 3' overhang (Fig. 4.3a & b). The resultant deletion clones at each time point were then recircularised after blunt ending using methods employing either S1 or mung bean nuclease. Recircularized clones from each time point were then transformed into DH5α, and screened for the appropriate sized deletions using the "lid lysate" technique (Hoekstra, 1988).

Sequencing of the insert in pBDV2 confirmed that it did indeed contain *PYC2*, ie a different *PYC* gene very similar in sequence to *PYC1*. The insert was found to be 3.54 kb in length extending from the *Sau*3A sites at -675 (Fig. 4.2) and +2866 (sites inclusive) relative to the first ATG. Subsequent sequencing of an overlapping 3' clone (described below) revealed that the 3' terminal *Sau*3A site at +2866 was actually contained within a *Bgl*II site (Fig. 4.2). The nucleotide and inferred amino acid sequences derived from the above mentioned pBDV2 clone are presented in figure 4.4.

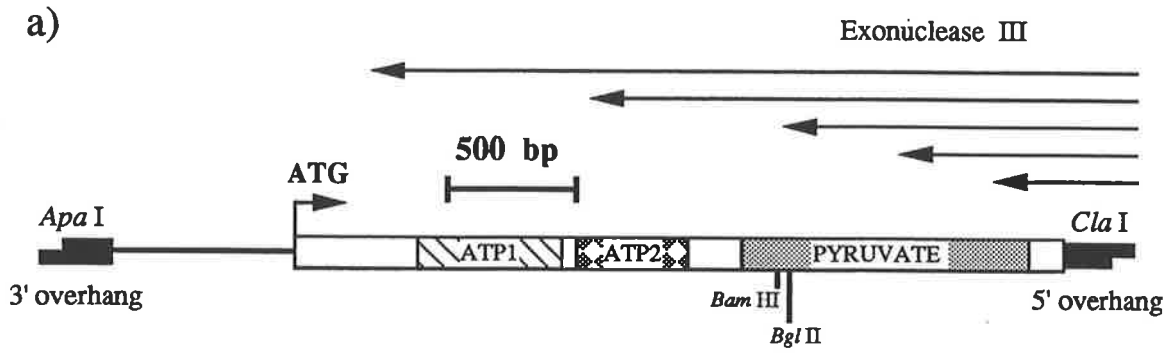
4.2.4 Isolation of a *PYC2* gene by another group

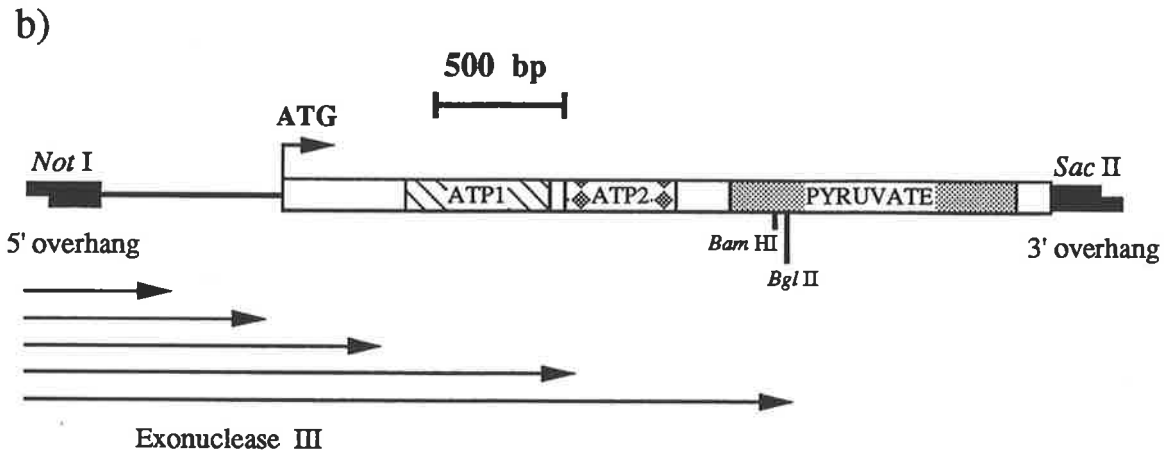
Unbeknown to us, at the same time I was working on the *PYC2* gene, a collaborative effort between three European labs was being directed at isolating mutants lacking in Pyc activity in order to study the interaction between glycolysis and other pathways leading to oxaloacetate formation in *S. cerevisiae*. In the course of these investigations these workers found evidence for a second *PYC* gene in yeast. They subsequently isolated and sequenced this gene (*PYC2*), and constructed and characterised null mutants with either or both of the *PYC* genes disrupted (Stucka *et al.*, 1991).

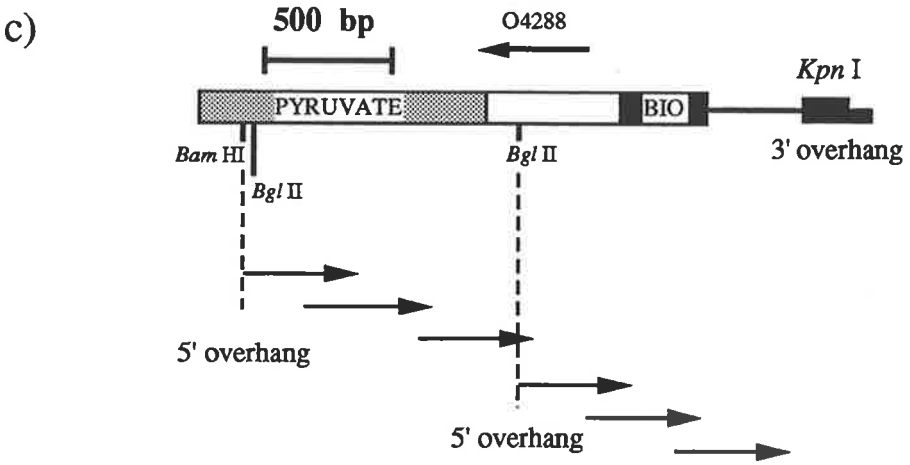
Upon discovering this paper it was necessary to decide whether to continue my research in characterising and disrupting the *PYC2* gene, or to request their *PYC2* clone and null mutant strains. Although on the surface it may appear to be an unnecessary duplication,

Figure 4.3 Schematic representation of the Nested Deletion approach used to obtain the overlapping DNA sequencing clones.

(a) The 3' end of clone pBDV2. (b) The 5' end of clone pBDV2. (c) The 3' end of clone p58. Sets of overlapping sequencing clones were obtained by using Exonuclease III to progressively remove increasing amounts of DNA from the end of the linearised plasmid containing the 5' overhang, with the other end being protected by a restriction site with a 3' overhang. Restriction sites in the flanking poly linker sequences used to create the necessary overhangs at the ends of the linearised fragments are shown in solid black boxing, while restriction sites used within the coding region are indicated by the dotted lines. Non-coding and non-Pyc regions are shown by the solid line (not to scale), the Pyc coding region is shown by the boxed sections with the domains as indicated (to scale). The horizontal unlabelled arrows show the typical pattern of deletion sizes obtained upon Exonuclease III digestion, while the O4288 arrow shows the position of the sequencing primer used to sequence across a gap in the sequence obtained from the non-coding strand of clone p58.


a)
Exonuclease III
500 bp
ATG
Apa I
ATP1
ATP2
PYRUVATE
Cla I
3' overhang
Bam HI
Bgl II
5' overhang


b)
500 bp
ATG
Not I
ATP1
ATP2
PYRUVATE
Sac II
5' overhang
Bam HI
Bgl II
3' overhang
Exonuclease III


c)
500 bp
O4288
Kpn I
PYRUVATE
BIO
Bam HI
Bgl II
3' overhang
Bgl II
5' overhang
5' overhang

as a step towards further structure / function studies I decided to continue my research towards characterising and disrupting *PYC2* for the following reasons.

Firstly, at the time I had already completed sequencing the coding strand of the pBDV2 clone and was in process of sequencing the other strand, and preparing the constructs to produce *PYC2* null mutants. Comparison of the two DNA sequences suggested that there were either a number of errors in the reported sequence, or differences due to strain polymorphism (or a combination of the two). Hence it appeared that further sequencing would be necessary to substantiate these differences anyway. Also, given the sequence differences in the 5' clone, it seemed quite likely that there would be other differences in the 3' end of the gene.

Secondly, Stucka *et al* (1991) had used a different parental strain (W303) in constructing the null mutants than Dr. M. E. Walker had used (DBY746) in the construction of the *pyc1* null mutant (MW21.3) in our laboratory. Stucka and co-workers found that disruption of either of the *PYC* genes had no effect on the growth phenotype with regard to aspartate dependence. In contrast MW21.3 exhibits growth dependence on aspartate (Dr Walker; pers. comm.), further implying that there may be allelic variation in the *PYC2* gene between laboratory strains.

Finally, as we had published our evidence for the existence of a second *PYC* gene and the chromosomal localisation data first (Walker *et al.*, 1991), it was felt that there may well be considerable reluctance by these workers in supplying us with their clones and strains, especially as our research interests and thus future work would most certainly overlap.

4.2.5 Isolation of the 3' end of *PYC2*

Repeated attempts to isolate the 3' end of the *PYC2* gene from the same genomic library using various *PYC* probes were unsuccessful. Instead, isolation of a clone containing the 3' end of the gene (p58) was achieved by screening a λEMBL3 genomic library with the 367 bp *PYC2* *Eco*RI fragment (Fig. 4.2). The 3' end of this clone was then subcloned as a 2.3 kb *Bam*HI-*Sal*I fragment into pHSG396 (Takashita *et al.*, 1987) for sequencing from both directions (Fig. 4.2).

4.2.6 Sequencing the 3' end of *PYC2*

Sequencing of the coding strand of the 3' clone was again achieved by the combined approaches of sequencing from the natural restriction sites and nested deletions from both ends of the clone (Fig 4.3c). In the case of the non-coding strand, sufficient overlaps were obtained by sequencing from the natural restriction sites (Fig. 4.2) and the use of the synthetic primer O4288 to complete one of the overlaps (Fig. 4.3c).

Considerable care was taken to ensure that any differences between my sequence data and the reported sequence with both this clone and the above mentioned 5' clone were not due to errors on my part.

4.2.7 Analysis of the nucleotide and inferred protein sequence

In total, 4.3 kb of *PYC2* DNA has been sequenced in both directions including the entire 3.54 kb open reading frame (ORF), 675 bp of 5' non-coding, and 97 bp of 3' non-coding sequence extending to the *Dra*I site at +3681 (Fig. 4.4). The putative TATA box elements and transcription factor binding site consensus sequences in the 5' non-coding region (Stucka *et al*, 1991) are shown in figure 4.5.

The ORF encodes for a protein of 1180 amino acids in length (Fig. 4.4) having a predicted molecular weight of 130,250.7 Da and isoelectric point of 6.52. Although the identity of the first amino acid of Pyc2 has not been determined experimentally, there are several pieces of data indicating that the first ATG of the ORF shown in figure 4.5 is almost certainly the translation initiation codon. Firstly, in the case of Pyc1, sequencing of an N-terminal peptide extending from residues 15 - 42 (as the N-terminus of Pyc1 was blocked) indicated that the ORF extends at least as far as the "+1" ATG, and beyond this point there is an in-frame termination codon at base -6 (Lim *et al.*, 1988). Secondly, upstream of the similarly located first ATG in the ORF of Pyc2 there are in-frame termination codons at positions -6, -18, -27, and -39 (see Fig. 4.5), and there is an adenine at position -3 (Fig 4.4) which is typical of eukaryotic initiator AUG's (Kozak, 1981).

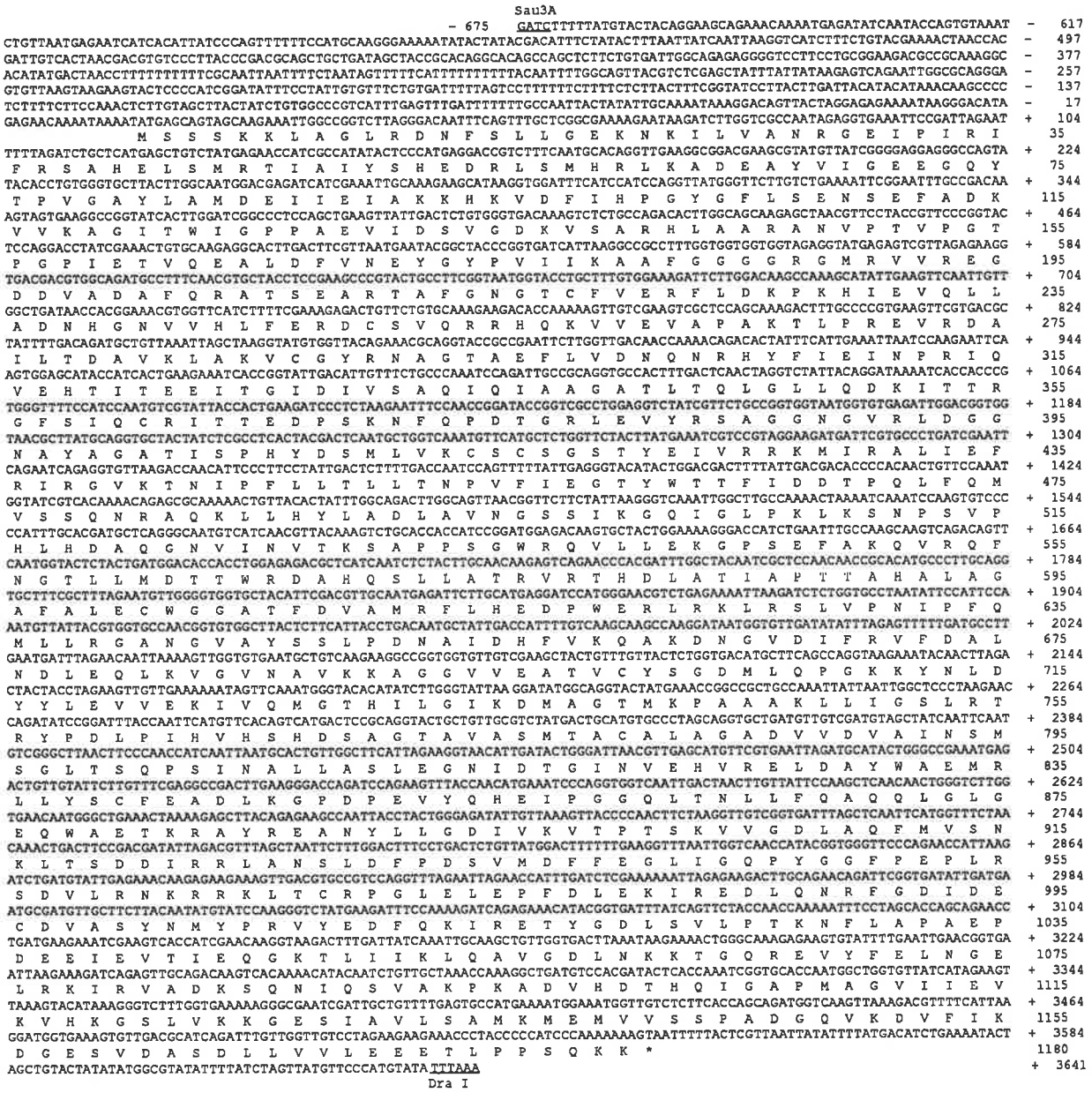


Figure 4.4 **Nucleotide and inferred protein sequence for the *PYC2* gene.**

The top sequence refers to the nucleotide sequence of the coding strand, the lower sequence refers to the inferred protein sequence, and the numbers in the right hand column correspond to the positions for each of these respective sequences relative to the first ATG. The underlining shows the positions of the flanking *Sau*3A and *Dra*I restriction sites.

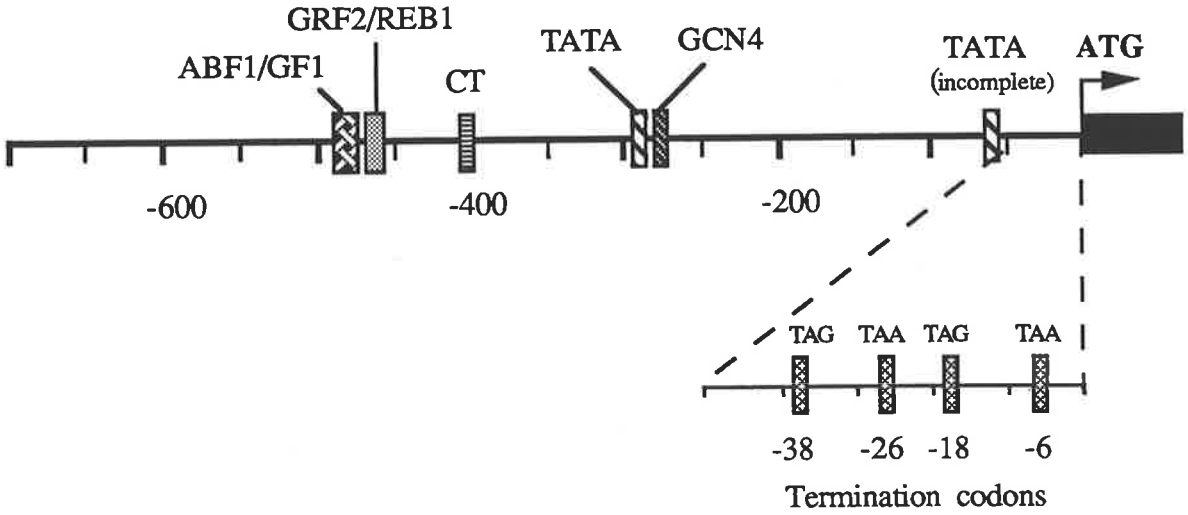


Figure 4.5 Schematic diagram of the promoter region of the *PYC2* gene

The start of the *PYC2* open reading frame is indicated by the position of the bent arrow and the solid boxing. The yeast promoter elements identified by Stucka *et al.* (1991) are shown by the shaded upright boxes positioned along the top line, while the shaded upright boxes along the lower line show the positions of the upstream TAG and TAA termination codons. Numbers indicate the sequence position in base pairs relative to the first ATG of the *PYC2* open reading frame.

4.2.8 Sequence comparisons between the various *PYC* isoforms

4.2.8.1 Comparison between the two *PYC2* sequences

Comparing the *PYC2* non-coding regions revealed that in the 772 bp of sequenced non-coding DNA there are a total of 16 differences, giving a sequence identity of 97.9% (average of 1 sequence difference every 48 residues). These include 4 deletions, 3 insertions and 9 transversions in the published sequence, relative to the sequence of the gene presented here (see Table 4.1). However there were no sequence differences observed within either the incomplete TATAT sequence at -65, or the TATTTA sequence at -290 (Fig 4.5). Furthermore the putative binding sites for the regulatory factors ABF1 / GF1 (-492), GRF2 / REB1 (-496) and the CT block at -493 (Fig. 4.5) (Stucka *et al.*, 1991) were unchanged.

Position	Deletion	Insertion	Transversion
- 675			G - A
- 665			T - G
- 664			G - C
- 557			T - C
- 499			C - T
- 439	G		
- 353	T*		
- 316		T*	
- 195			C - T
- 182		T	
- 178			A - T
- 104	T		
- 95		A	
- 23			G - A
- 7			A - G
+3626	C		

TABLE 4.1 Sequence differences between the two *PYC2* genes in the non-coding region

The differences between the two *PYC2* sequences refer to the mutations that would have to be introduced into *PYC2* DNA sequence reported in figure 4.4 in order for it to have an identical sequence to that reported by Stucka *et al.*(1991). The first base in each pair listed in the transversion column refers to the base present in the sequence reported in figure 4.4. The asterisk (*) indicate that the stated nucleotide position is only approximate as it occurs at a site of multiple Ts.

Comparison of the two ORF's revealed 36 sequence differences in a total of 3.54 kb, equivalent to one sequence difference every 98 bp, or a sequence identity of approximately

99%. Of the 36 sequence differences, 35 were transversions (Table 4.2), and the most 3' of the differences was a single base insertion (relative to my data) at +3533. Of the 35 transversions, 25 (71%) occurred in the third or "wobble" position. Consequently only 3 of these resulted in an amino acid change in the predicted protein sequence. The majority of the remaining 10 transversions, being in either the first or second codon positions, resulted in changes in the predicted protein sequence giving a total of 12 amino acid differences due to transversions (Table 4.2 & Fig. 4.6). By way of proof, the data from which the nucleotide sequence of *PYC2* gene reported here (Fig 4.4 & 4.6) was deduced, at each of the sites which encode the above mentioned amino acid differences due to transversions, are presented in figure 4.7.

As mentioned, the most 3' of the sequence differences in the ORF was an insertion in the published sequence at +3533. This results in a frame shift changing Q1178 into a proline (P), and reading through the stop codon thus lengthening the ORF by 15 bp (see Fig. 5.8b (i)). Hence the inferred protein sequence (1185 aa) has a 5 amino acid (VIFTR) extension at the C-terminus relative to that predicted on the basis of the nucleotide sequence reported here (Fig. 4.6 & 5.8b (i)). By way of proof the DNA sequence of the *PYC2* non-coding strand at position +3533 from the gene I have isolated is presented in figure 5.8b (ii).

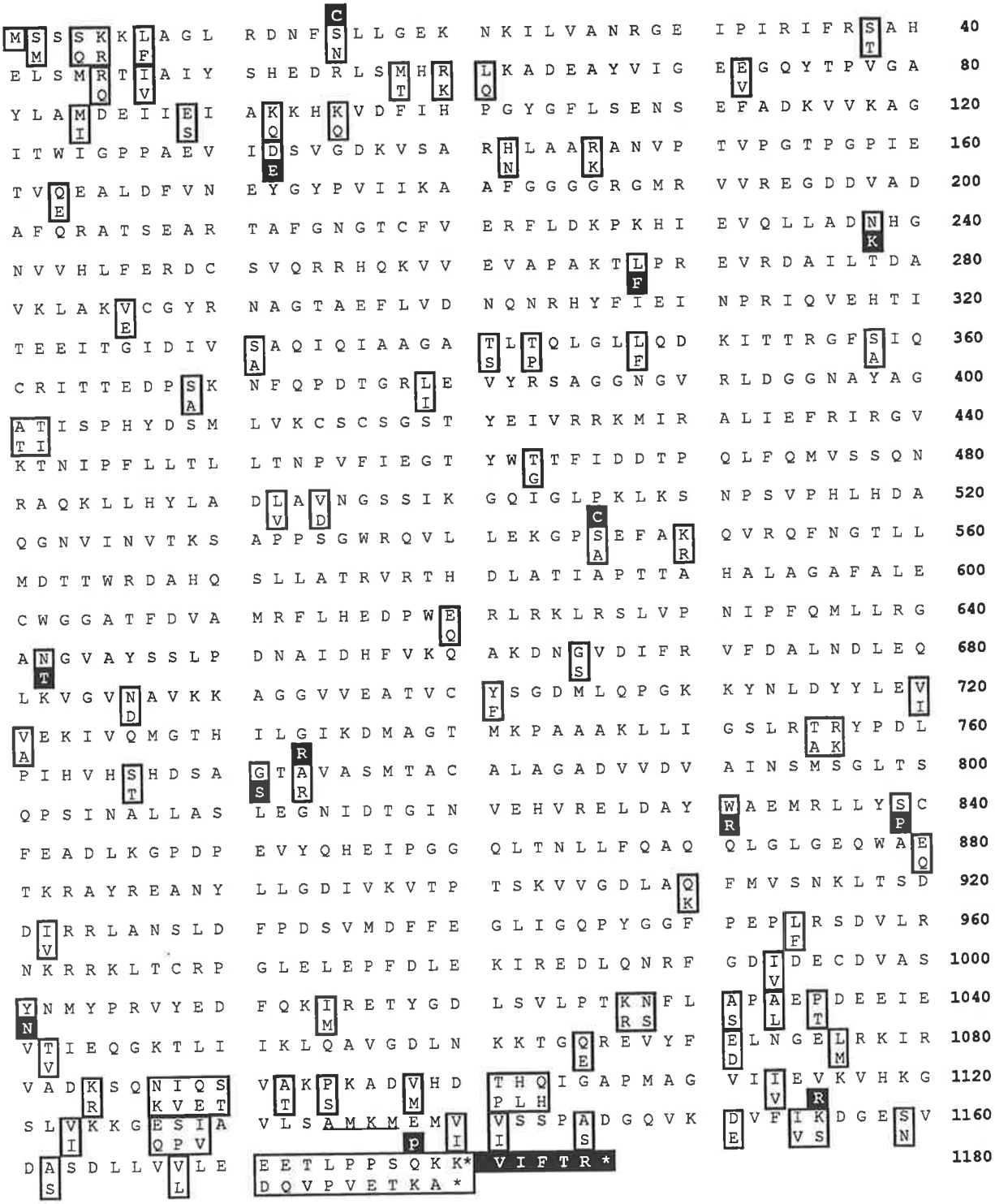
4.2.8.2 Comparisons between *PYC1* and *PYC2* and their protein products

The ORF of the *PYC2* gene described in this chapter is 6 bp longer than the ORF of *PYC1*, and has a percent identity of 85.4%. At the protein level Pyc2 has 92.6% identity to Pyc1, having one additional amino acid at the N and C termini, and 92 internal amino acid sequence differences (Fig. 4.6).

Stucka *et al.*(1991) calculated the codon bias index (Bennetzen and Hall, 1982) of *PYC2* to be 0.57, slightly higher than the value for *PYC1* (0.55) suggesting that it may be more highly expressed (Sharp *et al.*, 1986). However Northern analysis and RNase protection assays have shown that Pyc1 is in fact the more highly expressed of the two genes

Figure 4.6 Sequence comparisons between the yeast Pyc enzymes.

The main sequence shown corresponds to the inferred protein sequence from the *PYC2* gene I isolated, and the numbers on the right refer to the residue positions for this sequence. The letters connected to the main sequence by open boxes indicate the corresponding residue present in Pyc1 (Lim *et al.*, 1988), while those in the shaded boxes indicate the residues present in the published sequence for Pyc2 (Stucka *et al.*, 1991). The letters in the shaded box directly following the terminus of the main sequence indicate the residues present in the 5 amino acid extension present in the published sequence of Pyc2.


40
80
120
160
200
240
280
320
360
400
440
480
520
560
600
640
680
720
760
800
840
880
920
960
1000
1040
1080
1120
1160
1180

in strains DBY745 and S288C (Brewster *et al*, 1994). In addition, their expression paturns differ markedly depending on growth phase and the carbon source (Brewster *et al*, 1994).

Nucleo-tide Position	Nucleo-tide Change#	Codon Position	Amino acid Change	Amino acid No.
+ 43	A - T	1	S - C	15
+ 138	C - T	3		
+ 156	T - C	3		
+ 177	C - T	3		
+ 372	C - T	3		
+ 396	C - A	3	D - E	132
+ 408	C - T	3		
+ 411	A - G	3		
+ 465	T - C	3		
+ 714	C - A	3	N - K	238
+ 804	G - C	3	L - F	268
+1068	G - T	3		
+1320	T - G	3		
+1380	A - T	3		
+1431	G - A	3		
+1637	C - G	2	S - C	546
+1785	T - G	3		
+1815	T - G	3		
+1914	A - G	3		
+1925	A - C	2	N - T	642
+1926	C - T	3	" "	
+2013	T - C	3		
+2145	C - T	3		
+2311	G - A	1	G - S	771
+2317	G - C	1	A - R	773
+2318	C - G	2	" "	" "
+2472	T - C	3		
+2490	C - T	3		
+2491	T - A	1	W - R	831
+2515	T - C	1	S - P	839
+2541	A - C	3		
+3001	T - A	1	Y - N	1001
+3372	A - G	3		
+3453	C - T	3		
+3464	A - G	2	K - R	1155

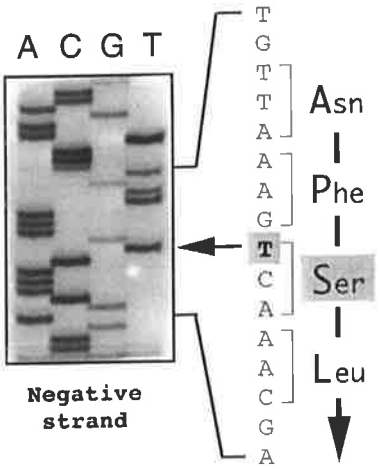
TABLE 4.2 **Nucleotide and amino acid sequence differences in the coding region of *PYC2***

The first base listed in the change column refers to the base present in the (+) strand of thegene reported here, while the second base refers to sequence of Stucka *et al.*, (1991). The symbols listed in the amino acid change column refer to the amino acids encoded by the corresponding nucleotide sequences.

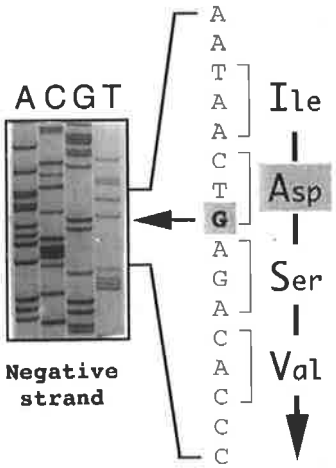
Figure 4.7 Verification of the *PYC2* DNA sequence at the sites in the coding region where the predicted amino acid sequence differs from the published data.

The photographs of the DNA sequences displayed in this figure show the data used to determine the sequence of the *PYC2* gene reported in figures 4.4 and 4.6 at each of the 12 "transversion " sites which encode for the amino acid differences between the data reported in this chapter (figures 4.4 and 4.6) and the published sequence of Stucka *et al.* (1991). The shading highlights the nucleotide and amino acid residues which differ between my data and the published sequence. The bold vertical arrows indicate the direction of translation.

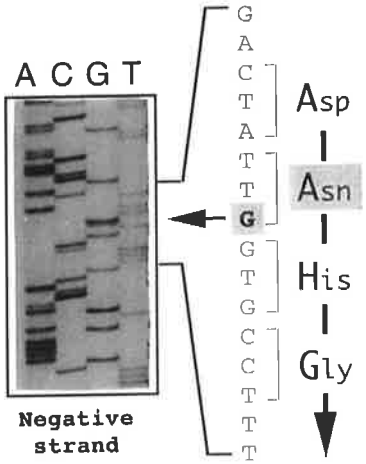
(i) S15



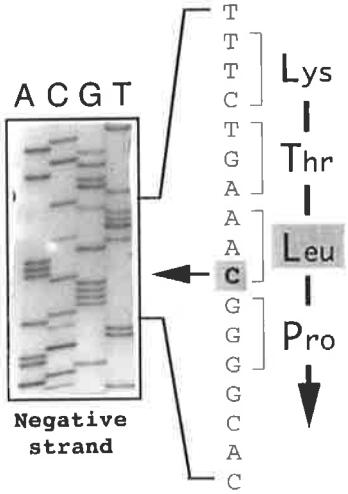
(ii) D132



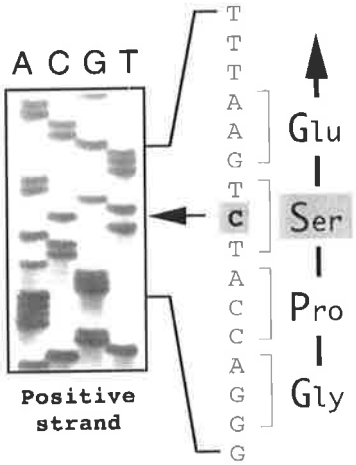
(iii) N238



(iv) L268



(v) S546



(vi) N642

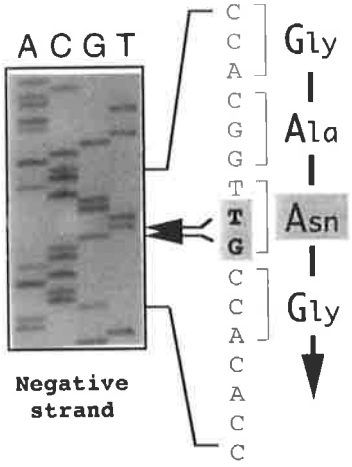
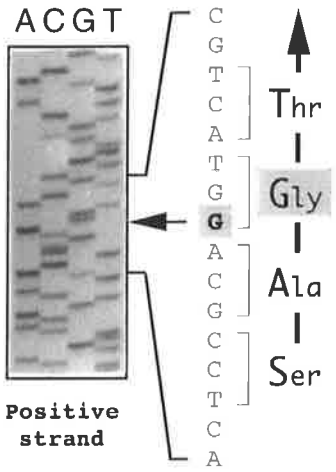
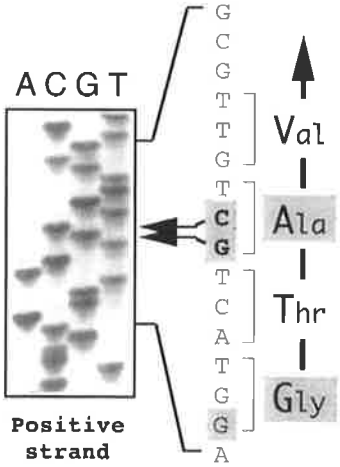


Figure 4.7 (continued)

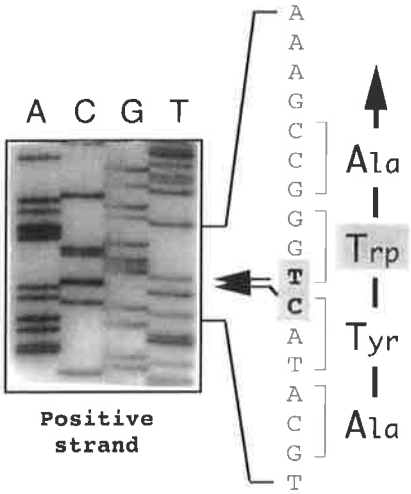
(vii) G771



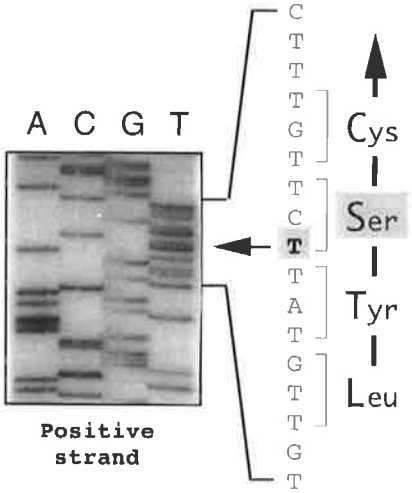
(viii) A773



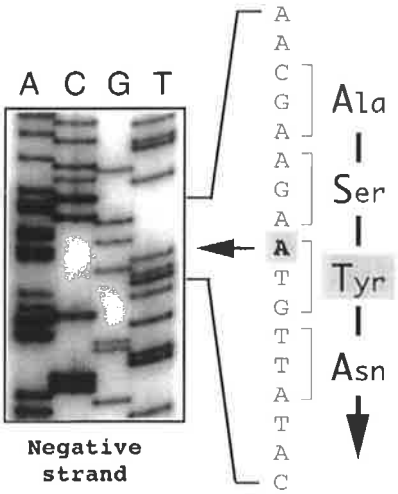
(ix) W831



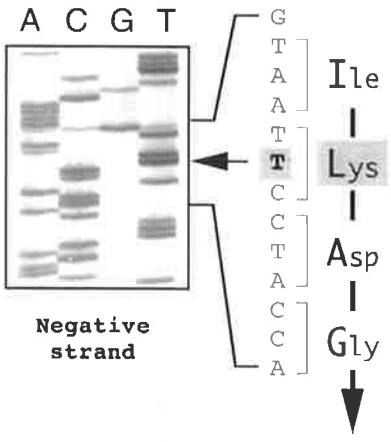
(x) S839



(xi) Y1001



(xii) K1155R



4.2.9 Sequence comparisons with other biotin enzymes

As previously mentioned (see Chapter 1, ie 1.5.2), once the sequence of yeast *PYC1* was determined by Lim *et al.* in 1988 it became obvious that the various domains of the Pyc protein had considerable sequence similarity with the domains of other enzymes which carry out the same or very similar chemical reactions. In the intervening years much progress has been made towards understanding the structure and function of many of these enzymes, and many more sequences having been reported. Hence, in an effort to identify groups of amino acid residues which may be responsible for the various substrate binding and other functions of Pyc, the ATP, pyruvate, and biotin domains of the Pyc2 sequence reported here have been compared with all of the available sequences of biotinyl and non-biotinyl enzymes which display sequence similarity with Pyc.

The N-terminal ATP domain has sequence similarity with all biotin carboxylases (Class I biotin enzymes) that have been sequenced thus far (urea amidolyase ,Ual; Pcc, & Acc), and the N-terminal ATP binding region of carbamoyl-phosphate synthetases (Cps) which also catalyse the formation of carboxy-phosphate from ATP and bicarbonate (Wimmer *et al.*, 1979) (Fig 4.8). Actually the ATP domain also has sequence similarity (slightly lower) with the ATP binding region in the C-terminal half of Cps enzymes, as the two halves of Cps are essentially repeats believed to be the result of an earlier gene duplication event (Nyunoya and Lusty, 1983). This C-terminal ATP binding region has been found to be the site of synthesis of carbamoyl-phosphate (Post *et al*, 1990), a structural analogue of carboxy-phosphate (Fig. 1.1). However due to the different chemistry involved in this reaction, these C-terminal sequences have not been included in figure 4.8.

The central pyruvate binding domain of Pyc2 is very similar in sequence to the other biotin enzymes which carry out the same (or the reverse) transcarboxylation reaction as Pyc (2nd partial reaction), ie transcarboxylase (Tc) and oxaloacetate decarboxylase (Odc) (see Fig. 4.9).

Finally, the biotin domain of Pyc2 is related in sequence to all other biotin domains, and the lipoyl domains of 2-oxo acid dehydrogenases (see Fig 4.10).

4.3 DISCUSSION

4.3.1 Differences between the two Pyc2 sequences

Mossé and coworkers (1993) recently published the compilation of a sequence data base for *S. cerevisiae* containing 1001 reported nucleotide sequences for 818 non-allelic genes. Although most of the genes had only been sequenced once, there were 149 genes for which more than one sequence had been determined. Of these, 30 genes showed differences at the nucleotide level within the ORF, but were totally identical at the predicted amino acid level (silent changes). They concluded that these "silent changes" were most likely due to natural polymorphisms. In addition to these "silent changes", by virtue of their frequency, it could be expected that at least some of the "non-silent" changes observed with the other genes could also be due to polymorphisms rather than sequencing errors. Furthermore, there are some examples in the literature of genes from different laboratory strains of *S. cerevisiae* having polymorphic variation at the amino acid level, eg: *URA3*, Rose *et al.*, (1984); and *POP2*, Sakai *et al.*, (1992). In the case of the two *PYC2* gene sequences, 61% (22 / 65; Table 4.2) of the nucleotide differences in the coding region were silent mutations, consistent with their being due to natural polymorphisms.

Considering the predicted amino acid substitutions in the coding region, only 2 of the 13 are strictly conservative (D132E, K1155R), although N642T and G771S (if residue 771 is exposed to the solvent) involve changes to one of the statistically preferred alternatives for these residues (Bordo & Argos, 1991). Of the remaining 9 amino acid differences, seven (S15C, N238K, L268F, S546C, A773R, S839P, Q1178P) occur infrequently with amino acids in equivalent environments for the protein families that have been looked at. The remaining two would be described as unlikely (W831R) or "avoided" (Y1001N) on the basis of the residue change allowing the same global protein topology (Bordo & Argos, 1991). Consequently, as most of these differences between the two predicted amino acid sequences are non-conservative, two Pyc2 enzymes containing these or similar amino acid changes might well be expected to display some differences in activity, leading to differences in growth phenotype between the respective *pyc1 / PYC2* null mutant strains. Hence, if real (ie not simply due to errors in the published sequence), these sequence differences may

contribute to the reported differences in Pyc2 specific activity (relative to Pyc1), and the differences in aspartate dependence of the *pyc1 / PYC2* single null mutants (Walker *et al.*, 1991; Stucka *et al.*, 1991).

It must be pointed out however that in the absence of the raw data for the *PYC2* sequence reported by Stucka *et al.* (1991), the possibility that some or even all of the differences between the two *PYC2* sequences may be due to errors in the published sequence can not be ruled out . The sequence comparisons between Pyc2 and related enzymes presented in figures 4.8 and 4.9 indicate that the residue changes which would be considered to be the most likely sites at which sequencing errors may have occurred (in the absence of direct evidence to the contrary) are L268F, N642T, G771S.

Figure 4.8 part b indicates that the L268F difference in the published Pyc2 sequence occurs within the ATP domain at a position in which all Pycs, Pccs, Cpss, Ual, and *Anabaena* Acc all have a leucine residue. Furthermore the only other enzymes (the remaining Acc's) which do not have leucines in this position all have either a conservative change to an isoleucine, or alanine, which is also an aliphatic hydrophobic residue. Figure 4.9 reveals that the N642T difference in the published Pyc2 sequence occurs within the pyruvate domain at a position in which all the enzymes which carry out the same transcarboxylation reaction (or the reverse in the case of ODC) as Pyc have an asparagine. Finally, figure 4.9 also reveals that the G771S difference occurs at a site at which all the enzymes have a glycine residue. In addition, this glycine is only 4 residues C-terminal of the HXHXH sequence may well turn out to be the divalent cation binding site in Pyc (see section 4.3.4.3), and thus could be an important residue in maintaining tertiary structure which makes up this site. However , it is worth mentioning again that although theoretically the N642T and G771S "differences" are not strictly conservative, empirical evidence suggests that these changes are actually among the statistically preferred alternative residues for structurally similar proteins, and thus a Pyc2 isozyme bearing these two changes as a result of polymorphism may in fact not suffer any detrimental alterations in structure.

The most notable difference between the two Pyc2 protein sequences is the presence of an additional 5 amino acids (VIFTR extension) at the C-terminus in the protein sequence reported by Stucka *et al.* (1991) (Fig. 4.6). Comparison of the two nucleotide sequences

revealed that this results from the insertion of an additional C into a stretch of 3 C's, thus altering the reading frame and extending the ORF by 15 nucleotides (Fig. 5.8b (i)). Given that mistakes can easily be made in the reading of multiple residues in sequencing gels, it would seen quite possible that the recorded 4 Cs, instead of the 3 clearly evident in the sequence reported here (actually the 3 Gs of the opposite strand are shown; see Fig. 5.8b (ii)), may simply be due to a sequencing error. However, in contrast to most of the protein sequence (see section 4.3.2), figure 4.6 shows that the C-terminal ~200 amino acids of Pyc1 and Pyc2 also differ quite markedly. Therefore it seems quite possible that the 5 amino acid VIFTR extension may indeed be the result of a genuine polymorphic variation between different forms of the *PYC2* gene.

4.3.2 Sequence differences between Pyc1 and Pyc2

As outlined in section 4.2.8.2, the Pyc2 sequence reported here differs from Pyc1 by having one additional residue at the N and C-termini, and 92 internal single amino acid changes. Of these a total of 50 occur within the boundaries of either the ATP or pyruvate domains which have been defined by sequence homology, or in the C-terminal region corresponding to the biotin domain of *E. coli BCCP* as revealed by its NMR structure (Brocklehurst *et al.* 1995). More precisely, 12 occur in the 317 residues from 148 - 464 in Pyc2 which spans the ATP domain (see Fig 4.8), giving an identity of 96.2 %. A further 12 occur in the stretch of 352 residues from 560 - 911 in Pyc2 which comprises the pyruvate domain (see Fig 4.9), indicating that the Pyc1 and Pyc2 pyruvate domains have a sequence identity of 96.6 %.

The other 26 changes which fall within the boundaries of the defined domains all occur in the 79 residues comprising the biotin domain. Hence the sequence identity between the two biotin domains is only 67 %, markedly less than that of the other two domains. This result agrees well with the noticeably lower sequence identity found among all the biotin domains sequenced to date (Fig. 4.10), as compared with the remarkably high identity evident in ATP and pyruvate domains (Fig. 4.8 &4.9). Indeed, Stucka *et al.*, (1991) noticed that a symmetric CCCTACCCCCATCCC sequence occurs near the "breakpoint" of the

Figure 4.8 Amino acid sequence homology in the ATP-binding domain of yeast Pyc2

Shading indicates sequence identity with Pyc2, clear boxes marked with asterisks show the amino acid differences present in the published Pyc2 sequence (Stucka *et al.*, 1991). The black boxing highlights the potentially important residues or motifs as follows: Part A the black boxing shows the position of the glycine rich motif which was proposed to form a "P-loop" ATP binding site in the Cps and biotin carboxylase enzymes (Lusty *et al.*, 1983; Kondo *et al.*, 1991: Li and Cronan, 1992), Part B the black boxing and solid circle highlight the position of the conserved cysteine residue (249 in Pyc1) proposed to be the base catalyst in the biotin carboxylation reaction (Li and Cronan, 1992), Parts B & C the single black boxed residues in the *E. coli* Acc sequence shown in parts B and C indicate the position of the residues located around the "putative" active site of the biotin carboxylase subunit of *E. coli* Acc (Waldrop *et al.*, 1994b), and Part C the black boxing and solid circle highlight the position of a cysteine residue in pyruvate carboxylases (360 in Pyc1) which appears to be an attractive candidate for the essential cysteine residue identified by chemical modification studies (see Chapter 7). The amino acid sequences presented were from the following sources: yeast Pyc2, data presented in this thesis; yeast Pyc1, (Lim *et al.*, 1988); human Pyc, (Walker *et al.*, 1995); mouse Pyc, (Zhang *et al.*, 1993); mosquito Pyc, Genbank; the α-subunit of rat Pcc, (Browner *et al.*, 1989); human Pcc, (Lamhonwah *et al.* 1989); rat Acc, (Lopez-Casillas *et al.*, 1988); human Acc, (HA *et al.*, 1994); *E. coli* Acc, (Kondo *et al.*, 1991); yeast Acc, (Al-Feel *et al.*, 1992); chicken Acc, (Takai *et el.*, 1988); *Pseudomonas aerogenosa* Acc, (Best and Knauf, 1993); *Anabaena* sp. Acc, strain PCC 7120 (Gornicki *et al.*, 1993); alfalfa, (Shorrosh *et al.*, 1994); wheat, (Gornicki *et al.*, 1994), *Arabadopsis*, (Roesler *et al.*, 1994), yeast Ual, (Genbauffe, and Cooper, 1991); human Cps, (Haraguchi *et al.*, 1991); rat Cps, (Nyunoya *et al.*, 1985); *Squalus acanthias* Cps, Genbank; *Drosophila melanogaster* Cps, (Freund and Jarry, 1987); *Plasmodium falciparum* Cps, Genbank; *Dictyostelium discoidium* Cps, (Elgar and Schofield, 1992); yeast Cps, (Lusty *et al.*, 1983); *E. coli* Cps, (Nyunoya and Lusty, 1983); *Bacillus subtilis* Cps, (Quinn *et al.*, 1991); *Neisseria gonorrhoeae* Cps, Genbank. Abbrebiations: Pyc, pyruvate carboxylase; Pcc, propionyl-CoA carboxylase; Acc, acetyl-CoA carboxylase; Ual, urea amidolyase; Cps, carbamoyl phosphate synthetase.

A

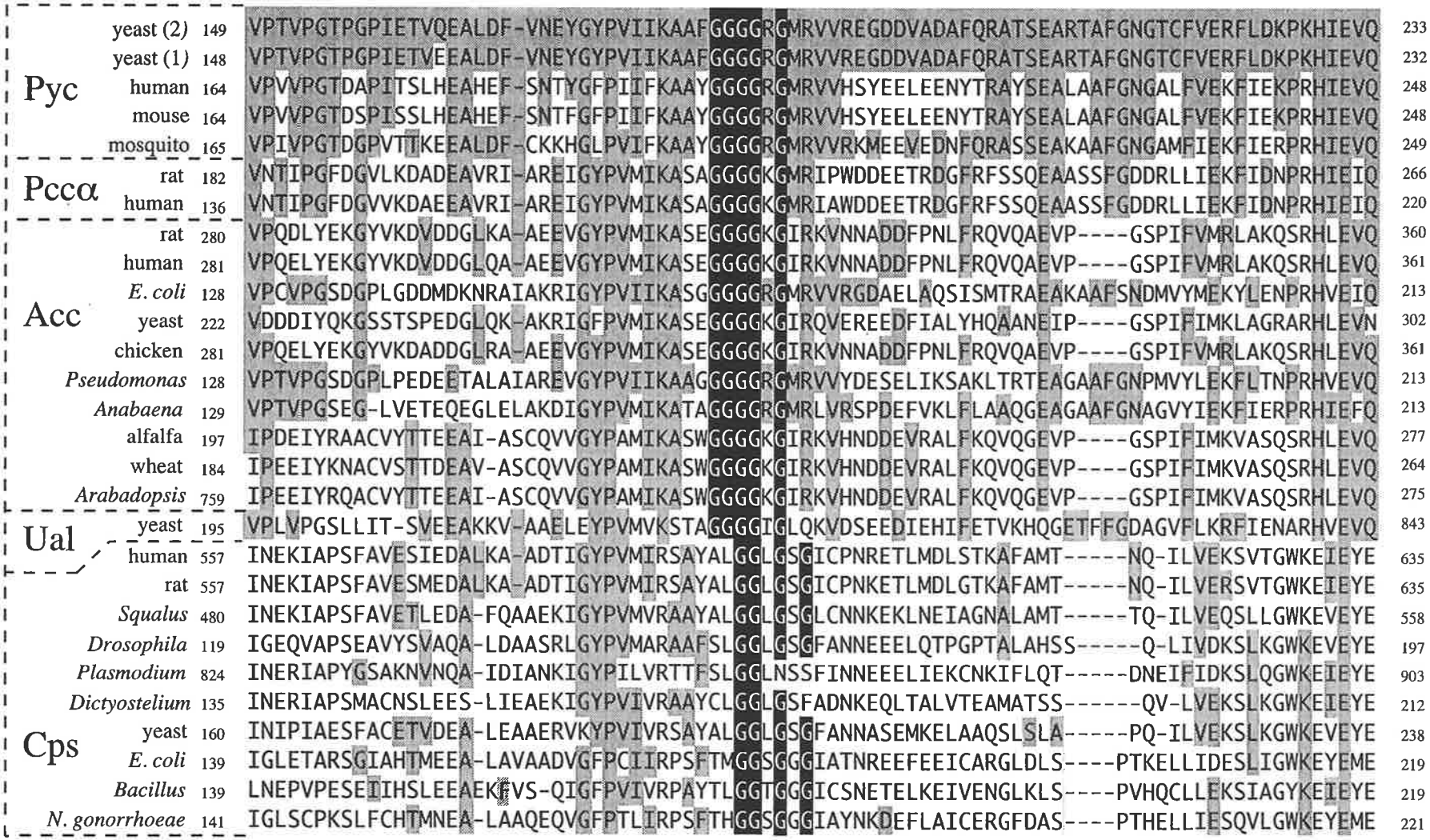

Pyc
Pccα
Acc
Ual
Cps
yeast (2) 149 VPTVPGTPGPIETVQEALDF-VNEYGYPVIIKAAFGGGGRGMRVVREGDDVADAFQRATSEARTAFGNGTCFVERFLDKPKHIEVQ 233
yeast (1) 148 VPTVPGTPGPIETVEEALDF-VNEYGYPVIIKAAFGGGGRGMRVVREGDDVADAFQRATSEARTAFGNGTCFVERFLDKPKHIEVQ 232
human 164 VPVVPGTDAPITSLHEAHEF-SNTYGFPIIFKAAYGGGGRGMRVVHSYEELEENYTRAYSEALAAFGNGALFVEKFIEKPRHIEVQ 248
mouse 164 VPVVPGTDSPISSLHEAHEF-SNTFGFPIIFKAAYGGGGRGMRVVHSYEELEENYTRAYSEALAAFGNGALFVEKFIEKPRHIEVQ 248
mosquito 165 VPIVPGTDGPVTTKEEALDF-CKKHGLPVIFKAAYGGGGRGMRVVRKMEEVEDNFQRASSEAKAAFGNGAMFIEKFIERPRHIEVQ 249
rat 182 VNTIPGFDGVLKDADEAVRI-AREIGYPVMIKASAGGGGKGMRIPWDDEETRDGFRFSSQEAASSFGDDRLLIEKFIDNPRHIEIQ 266
human 136 VNTIPGFDGVVKDAEEAVRI-AREIGYPVMIKASAGGGGKGMRIAWDDEETRDGFRFSSQEAASSFGDDRLLIEKFIDNPRHIEIQ 220
rat 280 VPQDLYEKGYVKDVDDGLKA-AEEVGYPVMIKASEGGGGKGIRKVNNADDFPNLFRQVQAEVP----GSPIFVMRLAKQSRHLEVQ 360
human 281 VPQELYEKGYVKDVDDGLQA-AEEVGYPVMIKASEGGGGKGIRKVNNADDFPNLFRQVQAEVP----GSPIFVMRLAKQSRHLEVQ 361
E. coli 128 VPCVPGSDGPLGDDMDKNRAIAKRIGYPVIIKASGGGGGRGMRVVRGDAELAQSISMTRAEAKAAFSNDMVYMEKYLENPRHVEIQ 213
yeast 222 VDDDIYQKGSSTSPEDGLQK-AKRIGFPVMIKASEGGGGKGIRQVEREEDFIALYHQAANEIP----GSPIFIMKLAGRARHLEVN 302
chicken 281 VPQELYEKGYVKDADDGLRA-AEEVGYPVMIKASEGGGGKGIRKVNNADDFPNLFRQVQAEVP----GSPIFVMRLAKQSRHLEVQ 361
Pseudomonas 128 VPTVPGSDGPLPEDEETALAIAREVGYPVIIKAAGGGGGRGMRVVYDESELIKSAKLTRTEAGAAFGNPMVYLEKFLTNPRHVEVQ 213
Anabaena 129 VPTVPGSEG-LVETEQEGLELAKDIGYPVMIKATAGGGGRGMRLVRSPDEFVKLFLAAQGEAGAAFGNAGVYIEKFIERPRHIEFQ 213
alfalfa 197 IPDEIYRAACVYTTEEAI-ASCQVVGYPAMIKASWGGGGKGIRKVHNDDEVRALFKQVQGEVP----GSPIFIMKVASQSRHLEVQ 277
wheat 184 IPEEIYKNACVSTTDEAV-ASCQVVGYPAMIKASWGGGGKGIRKVHNDDEVRALFKQVQGEVP----GSPIFIMKVASQSRHLEVQ 264
Arabadopsis 759 IPEEIYRQACVYTTEEAI-ASCQVVGYPAMIKASWGGGGKGIRKVHNDDEVRALFKQVQGEVP----GSPIFIMKVASQSRHLEVQ 275
yeast 195 VPLVPGSLLIT-SVEEAKKV-AAELEYPVMVKSTAGGGGIGLQKVDSEEDIEHIFETVKHQGETFFGDAGVFLKRFIENARHVEVQ 843
human 557 INEKIAPSFAVESIEDALKA-ADTIGYPVMIRSAYALGGLGSGICPNRETLMDLSTKAFAMT-----NQ-ILVEKSVTGWKEIEYE 635
rat 557 INEKIAPSFAVESMEDALKA-ADTIGYPVMIRSAYALGGLGSGICPNKETLMDLGTKAFAMT-----NQ-ILVERSVTGWKEIEYE 635
Squalus 480 INEKIAPSFAVETLEDA-FQAAEKIGYPVMVRAAYALGGLGSGLCNNKEKLNEIAGNALAMT-----TQ-ILVEQSLLGWKEVEYE 558
Drosophila 119 IGEQVAPSEAVYSVAQA-LDAASRLGYPVMARAAFSLGGLGSGFANNEEELQTPGPTALAHSS-----Q-LIVDKSLKGWKEVEYE 197
Plasmodium 824 INERIAPYGSAKNVNQA-IDIANKIGYPILVRTTFSLGGLNSSFINNEEELIEKCNKIFLQT-----DNEIFIDKSLQGWKEIEYE 903
Dictyostelium 135 INERIAPSMACNSLEES-LIEAEKIGYPVIVRAAYCLGGLGSFADNKEQLTALVTEAMATSS------QV-LVEKSLKGWKEIEYE 212
yeast 160 INIPIAESFACETVDEA-LEAAERVKYPVIVRSAYALGGLGSGFANNASEMKELAAQSLSLA-----PQ-ILVEKSLKGWKEVEYE 238
E. coli 139 IGLETARSGIAHTMEEA-LAVAADVGFPCIIRPSFTMGGSGGGIATNREEFEEICARGLDLS----PTKELLIDESLIGWKEYEME 219
Bacillus 139 LNEPVPESEIIHSLEEAEKFVS-QIGFPVIVRPAYTLGGTGGGICSNETELKEIVENGLKLS----PVHQCLLEKSIAGYKEIEYE 219
N. gonorrhoeae 141 IGLSCPKSLFCHTMNEA-LAAQEQVGFPTLIRPSFTHGGSGGGIAYNKDEFLAICERGFDAS----PTHELLIESQVLGWKEYEME 221

Figure 4.8 (part B)

B

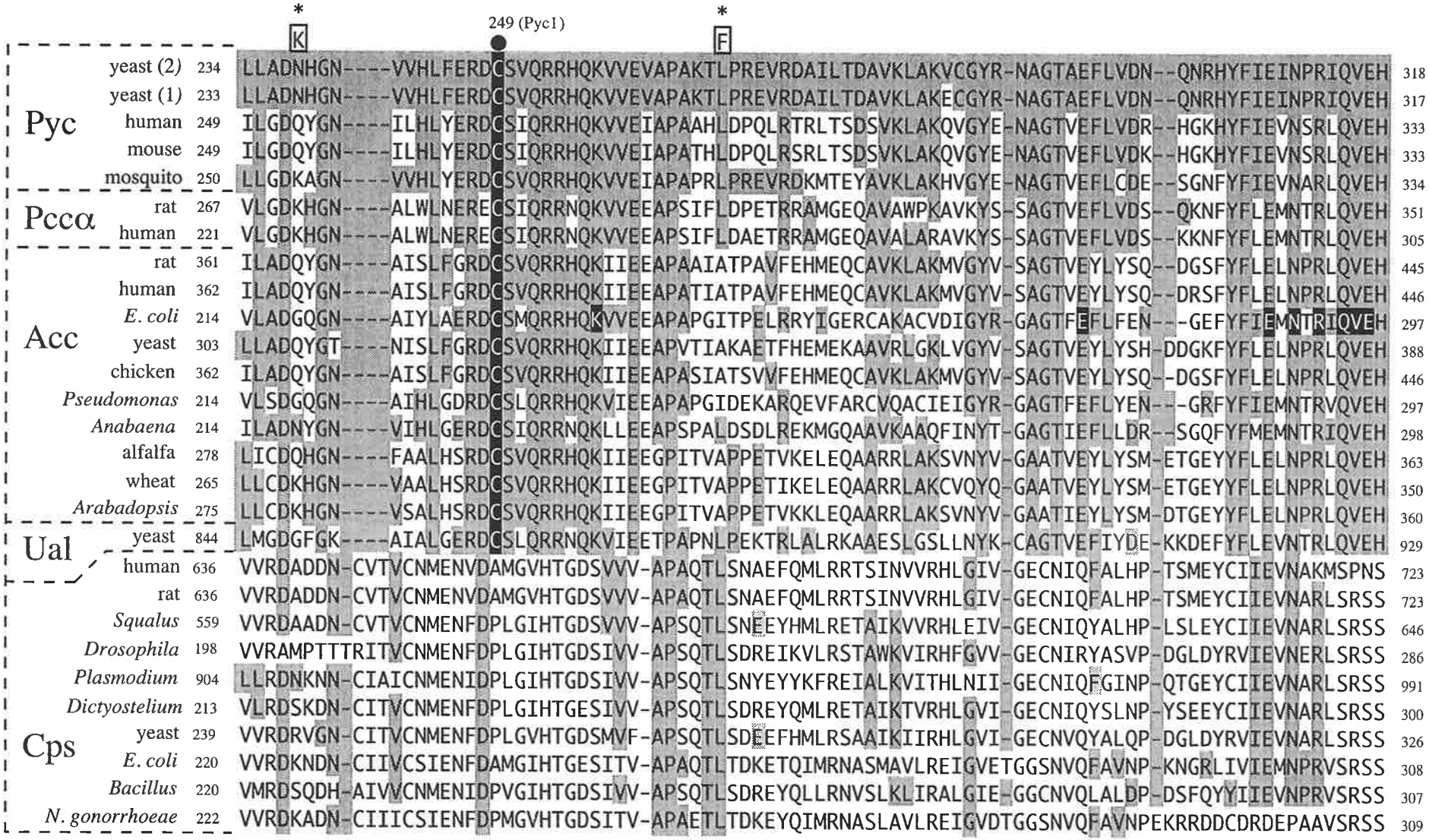

*
K
249 (Pyc1)
*
F
Pyc
Pccα
Acc
Ual
Cps
yeast (2) 234 LLADNHGN----VVHLFERDCSVQRRHQKVVEVAPAKTLPREVRDAILTDAVKLAKVCGYR-NAGTAEFLVDN--QNRHYFIEINPRIQVEH 318
yeast (1) 233 LLADNHGN----VVHLFERDCSVQRRHQKVVEVAPAKTLPREVRDAILTDAVKLAKECGYR-NAGTAEFLVDN--QNRHYFIEINPRIQVEH 317
human 249 ILGDQYGN----ILHLYERDCSIQRRHQKVVEIAPAAHLDPQLRTRLTSDSVKLAKQVGYE-NAGTVEFLVDR--HGKHYFIEVNSRLQVEH 333
mouse 249 ILGDQYGN----ILHLYERDCSIQRRHQKVVEIAPATHLDPQLRSRLTSDSVKLAKQVGYE-NAGTVEFLVDK--HGKHYFIEVNSRLQVEH 333
mosquito 250 LLGDKAGN----VVHLYERDCSVQRRHQKVVEIAPAPRLPREVRDKMTEYAVKLAKHVGYE-NAGTVEFLCDE--SGNFYFIEVNARLQVEH 334
rat 267 VLGDKHGN----ALWLNERECSIQRRNQKVVEEAPSIFLDPETRRAMGEQAVAWPKAVKYS-SAGTVEFLVDS--QKNFYFLEMNTRLQVEH 351
human 221 VLGDKHGN----ALWLNERECSIQRRNQKVVEEAPSIFLDAETRRAMGEQAVALARAVKYS-SAGTVEFLVDS--KKNFYFLEMNTRLQVEH 305
rat 361 ILADQYGN----AISLFGRDCSVQRRHQKIIEEAPAAIATPAVFEHMEQCAVKLAKMVGYV-SAGTVEYLYSQ--DGSFYFLELNPRLQVEH 445
human 362 ILADQYGN----AISLFGRDCSVQRRHQKIIEEAPATIATPAVFEHMEQCAVKLAKMVGYV-SAGTVEYLYSQ--DRSFYFLELNPRLQVEH 446
E. coli 214 VLADGQGN----AIYLAERDCSMQRRHQKVVEEAPAPGITPELRRYIGERCAKACVDIGYR-GAGTFEFLFEN--GEFYFIEMNTRIQVEH 297
yeast 303 LLADQYGT----NISLFGRDCSVQRRHQKIIEEAPVTIAKAETFHEMEKAAVRLGKLVGYV-SAGTVEYLYSH-DDGKFYFLELNPRLQVEH 388
chicken 362 ILADQYGN----AISLFGRDCSVQRRHQKIIEEAPASIATSVVFEHMEQCAVKLAKMVGYV-SAGTVEYLYSQ--DGSFYFLELNPRLQVEH 446
Pseudomonas 214 VLSDGQGN----AIHLGDRDCSLQRRHQKVIEEAPAPGIDEKARQEVFARCVQACIEIGYR-GAGTFEFLYEN--GRFYFIEMNTRVQVEH 297
Anabaena 214 ILADNYGN----VIHLGERDCSIQRRNQKLLEEAPSPALDSDLREKMGQAAVKAAQFINYT-GAGTIEFLLDR--SGQFYFMEMNTRIQVEH 298
alfalfa 278 LICDQHGN----FAALHSRDCSVQRRHQKIIEEGPITVAPPETVKELEQAARRLAKSVNYV-GAATVEYLYSM-ETGEYYFLELNPRLQVEH 363
wheat 265 LLCDKHGN----VAALHSRDCSVQRRHQKIIEEGPITVAPPETIKELEQAARRLAKCVQYQ-GAATVEYLYSM-ETGEYYFLELNPRLQVEH 350
Arabadopsis 275 LLCDKHGN----VSALHSRDCSVQRRHQKIIEEGPITVAPPETVKKLEQAARRLAKSVNYV-GAATIEYLYSM-DTGEYYFLELNPRLQVEH 360
yeast 844 LMGDGFGK----AIALGERDCSLQRRNQKVIEETPAPNLPEKTRLALRKAAESLGSLLNYK-CAGTVEFIYDE-KKDEFYFLEVNTRLQVEH 929
human 636 VVRDADDN-CVTVCNMENVDAMGVHTGDSVVV-APAQTLSNAEFQMLRRTSINVVRHLGIV-GECNIQFALHP-TSMEYCIIEVNAKMSPNS 723
rat 636 VVRDADDN-CVTVCNMENVDAMGVHTGDSVVV-APAQTLSNAEFQMLRRTSINVVRHLGIV-GECNIQFALHP-TSMEYCIIEVNARLSRSS 723
Squalus 559 VVRDAADN-CVTVCNMENFDPLGIHTGDSVVV-APSQTLSNEEYHMLRETAIKVVRHLEIV-GECNIQYALHP-LSLEYCIIEVNARLSRSS 646
Drosophila 198 VVRAMPTTTRITVCNMENFDPLGIHTGDSIVV-APSQTLSDREIKVLRSTAWKVIRHFGVV-GECNIRYASVP-DGLDYRVIEVNERLSRSS 286
Plasmodium 904 LLRDNKNN-CIAICNMENIDPLGIHTGDSIVV-APSQTLSNYEYYKFREIALKVITHLNII-GECNIQFGINP-QTGEYCIIEVNARLSRSS 991
Dictyostelium 213 VLRDSKDN-CITVCNMENFDPLGIHTGESIVV-APSQTLSDREYQMLRETAIKTVRHLGVI-GECNIQYSLNP-YSEEYCIIEVNARLSRSS 300
yeast 239 VVRDRVGN-CITVCNMENFDPLGVHTGDSMVF-APSQTLSDEEFHMLRSAAIKIIRHLGVI-GECNVQYALQP-DGLDYRVIEVNARLSRSS 326
E. coli 220 VVRDKNDN-CIIVCSIENFDAMGIHTGESITV-APAQTLTDKETQIMRNASMAVLREIGVETGGSNVQFAVNP-KNGRLIVIEMNPRVSRSS 308
Bacillus 220 VMRDSQDH-AIVVCNMENIDPVGIHTGDSIVV-APSQTLSDREYQLLRNVSLKLIRALGIE-GGCNVQLALDP-DSFQYYIIEVNPRVSRSS 307
N. gonorrhoeae 222 VVRDKADN-CIIICSIENFDPMGVHTGDSITV-APAETLTDKEYQIMRNASLAVLREIGVDTGGSNVQFAVNPEKRRDDCDRDEPAAVSRSS 309

Figure 4.8 (part C)

C

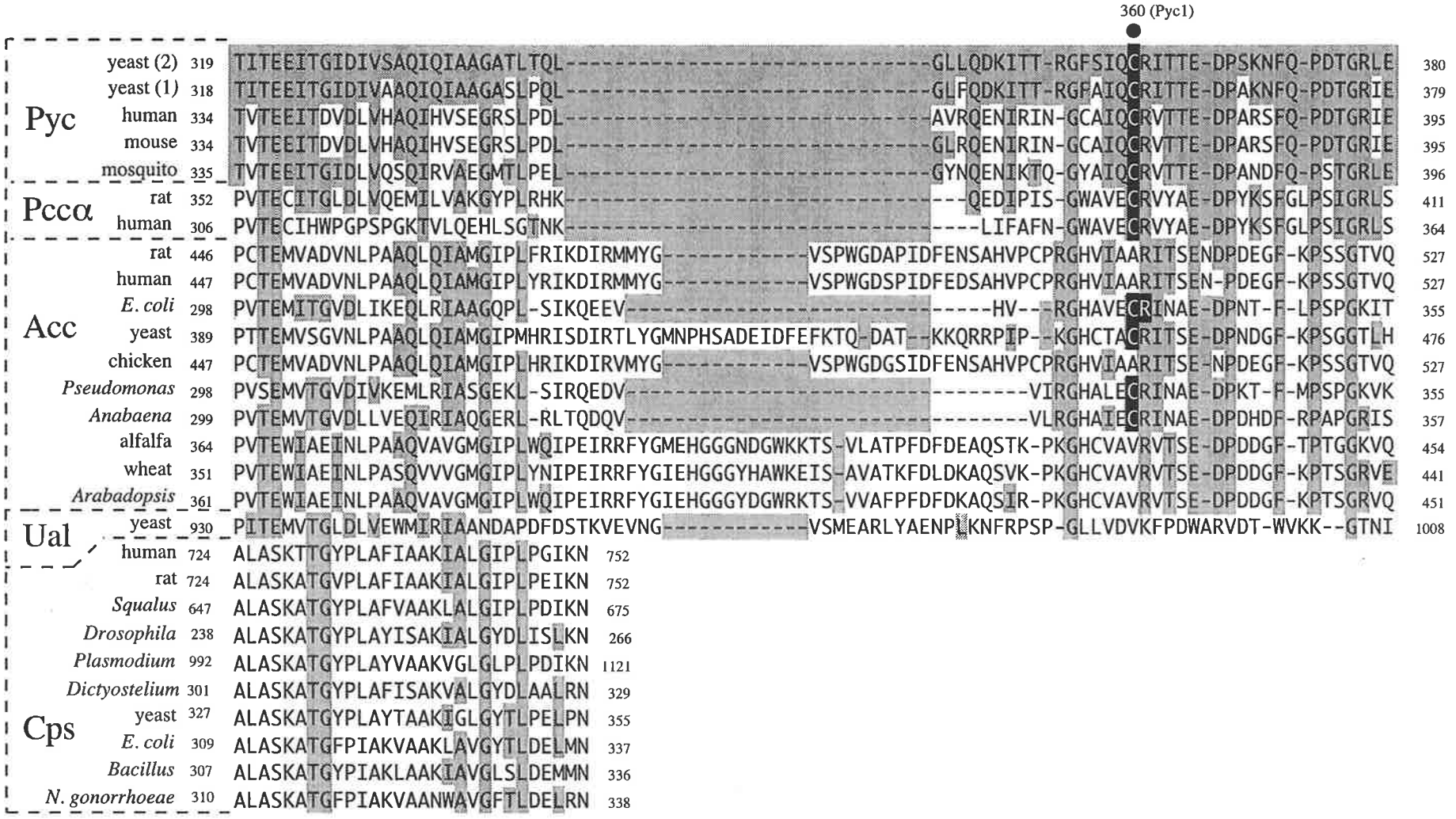

360 (Pyc1)
Pyc
Pccα
Acc
Ual
Cps
yeast (2) 319 TITEEITGIDIVSAQIQIAAGATLTQL----------GLLQDKITT-RGFSIQCRITTE-DPSKNFQ-PDTGRLE 380
yeast (1) 318 TITEEITGIDIVAAQIQIAAGASLPQL----------GLFQDKITT-RGFAIQCRITTE-DPAKNFQ-PDTGRIE 379
human 334 TVTEEITDVDLVHAQIHVSEGRSLPDL----------AVRQENIRIN-GCAIQCRVTTE-DPARSFQ-PDTGRIE 395
mouse 334 TVTEEITDVDLVHAQIHVSEGRSLPDL----------GLRQENIRIN-GCAIQCRVTTE-DPARSFQ-PDTGRIE 395
mosquito 335 TVTEEITGIDLVQSQIRVAEGMTLPEL----------GYNQENIKTQ-GYAIQCRVTTE-DPANDFQ-PSTGRLE 396
rat 352 PVTECITGLDLVQEMILVAKGYPLRHK----------QEDIPIS-GWAVECRVYAE-DPYKSFGLPSIGRLS 411
human 306 PVTECIHWPGPSPGKTVLQEHLSGTNK----------LIFAFN-GWAVECRVYAE-DPYKSFGLPSIGRLS 364
rat 446 PCTEMVADVNLPAAQLQIAMGIPLFRIKDIRMMYG----------VSPWGDAPIDFENSAHVPCPRGHVIAARITSENDPDEGF-KPSSGTVQ 527
human 447 PCTEMVADVNLPAAQLQIAMGIPLYRIKDIRMMYG----------VSPWGDSPIDFEDSAHVPCPRGHVIAARITSEN-PDEGF-KPSSGTVQ 527
E. coli 298 PVTEMITGVDLIKEQLRIAAGQPL-SIKQEEV----------HV---RGHAVECRINAE-DPNT-F-LPSPGKIT 355
yeast 389 PTTEMVSGVNLPAAQLQIAMGIPMHRISDIRTLYGMNPHSADEIDFEFKTQ-DAT--KKQRRPIP--KGHCTACRITSE-DPNDGF-KPSGGTLH 476
chicken 447 PCTEMVADVNLPAAQLQIAMGIPLHRIKDIRVMYG----------VSPWGDGSIDFENSAHVPCPRGHVIAARITSE-NPDEGF-KPSSGTVQ 527
Pseudomonas 298 PVSEMVTGVDIVKEMLRIASGEKL-SIRQEDV----------VIRGHALECRINAE-DPKT-F-MPSPGKVK 355
Anabaena 299 PVTEMVTGVDLLVEQIRIAQGERL-RLTQDQV----------VLRGHAIECRINAE-DPDHDF-RPAPGRIS 357
alfalfa 364 PVTEWIAEINLPAAQVAVGMGIPLWQIPEIRRFYGMEHGGGNDGWKKTS-VLATPFDFDEAQSTK-PKGHCVAVRVTSE-DPDDGF-TPTGGKVQ 454
wheat 351 PVTEWIAEINLPASQVVVGMGIPLYNIPEIRRFYGIEHGGGYHAWKEIS-AVATKFDLDKAQSVK-PKGHCVAVRVTSE-DPDDGF-KPTSGRVE 441
Arabadopsis 361 PVTEWIAEINLPAAQVAVGMGIPLWQIPEIRRFYGIEHGGGYDGWRKTS-VVAFPFDFDKAQSIR-PKGHCVAVRVTSE-DPDDGF-KPTSGRVQ 451
yeast 930 PITEMVTGLDLVEWMIRIAANDAPDFDSTKVEVNG----------VSMEARLYAENPLKNFRPSP-GLLVDVKFPDWARVDT-WVKK--GTNI 1008
human 724 ALASKTTGYPLAFIAAKIALGIPLPGIKN 752
rat 724 ALASKATGVPLAFIAAKIALGIPLPEIKN 752
Squalus 647 ALASKATGYPLAFVAAKLALGIPLPDIKN 675
Drosophila 238 ALASKATGYPLAYISAKIALGYDLISLKN 266
Plasmodium 992 ALASKATGYPLAYVAAKVGLGLPLPDIKN 1121
Dictyostelium 301 ALASKATGYPLAFISAKVALGYDLAALRN 329
yeast 327 ALASKATGYPLAYTAAKIGLGYTLPELPN 355
E. coli 309 ALASKATGFPIAKVAAKLAVGYTLDELMN 337
Bacillus 307 ALASKATGYPIAKLAAKIAVGLSLDEMMN 336
N. gonorrhoeae 310 ALASKATGFPIAKVAANWAVGFTLDELRN 338

Figure 4.8 (part D)

D


Pyc
Pccα
Acc
Ual
yeast (2) 381 VYR------SAGG-NGVRLDGGNAYAGATISPHYDSMLVK----CSCSGSTYEIVRRKMIRALIEFRIRGV-KTNIPFLLTLLTNPVFIEGTYWT 463
yeast (1) 380 VYR------SAGG-NGVRLDGGNAYAGTIISPHYDSMLVK----CSCSGSTYEIVRRKMIRALIEFRIRGV-KTNIPFLLTLLTNPVFIEGTYWG 462
human 396 VFR------SGEG-MGIRLDNASAFQGAVISPHYDSLLVK----VIAHGKDHPTAATKMSRALAEFRVRGV-KTNIAFLQNVLNNQQFLAGTVDT 479
mouse 396 VFR------SGEG-MGIRLDNASAFQGAVISPHYDSLLVK----VIAHGKDHPTAATKMSRALAEFRVRGV-KTNIPFLQNVLNNQQFLAGTVDT 479
mosquito 397 VFR------SGEG-MGIRLDSASAYAGAIISPYYDSLLVK----VISHASDLQSSAAKMNRALREFRIRGV-KTNIPFLLNVLENQKFLNGVLDT 480
rat 412 QYQ------EPIHLPGVRVDSGIQP-GSDISIYHDPMISKL----VTYGSDRAEALKRMEDALDSYVIRGV-THNIPLLREVIINTRFVKGDIST 494
human 365 QYQ------EPIHLPGVRVDSGIQP-GSDISIYYDPMISKL----ITYGSDRTEALKRMADALDNYVIRGV-THNIALLREVIINSRFVKGDIST 447
rat 528 ELN------FRSNK-NVWGYFSVAAAGGLHE-FADSQ----FGHCFSWGENREEAISNMVVALKELSIRGDFRTTVEYLIKLLETESFQQNRIDT 610
human 528 ELN------FRSNK-NVWGYFSVAAAGGLHE-FADSQ----FGHCFSWGESREEAISNMVVALKELSIRGDFRTTVEYLIKLLETESFQMNRIDT 610
E. coli 356 RFH------APGGF-GVRWESHIYA-GYTVPPYYDSMIGKL----ICYGENRDVAIARMKNALQELIIDGI-KTNVDLQIRIMNDENFQHGGTNI 437
yeast 477 ELNTLHELNFRSS-SNVWGYFSVGNNGNIHS-FSDSQ----FGHIFAFGENRQASRKHMVVALKELSIRGDFRTTVEYLIKLLETEDFEDNTITT 566
chicken 528 ELN------FRSNK-NVWGYFSVAAAGGLHE-FADSQ----FGHCFSWGENREEAISNMVVALKELSIRGDFRTTVEYLIKLLETESFQQNRTGW 610
Pseudomonas 356 HFHAPGGNGVRVDSHLYSGYSVPPNYDSLVGKVI----------TYGADRDEALARMRNALDELIVDGI-KTNTELHKDLVRDAAFCKGGVNI 437
Anabaena 358 GYL------PPGGP-GVRIDSHVYT-DYQIPPYYDSLIGKL----IVWGPDRATAINRMKRALRECAITG-LPTTIGFHQRIMENPQFLQGNVST 439
alfalfa 455 ELS------FKS-KPNVWAYFSVKSGGGIHE-FSDSQ----FGHVFAFGESRALAIANMVLGLKEIQIRGEIRTNVDYTIDLLNASDYRDNKIHT 537
wheat 442 ELN------FKS-KPNVWAYFSVKSGGAIHE-FSDSQ----FGHVFAFGESRSLAIANMVLGLKEIQIRGEIRTNVDYTVDLLNAAEYRENKIHT 524
Arabadopsis 452 ELS------FKS-KPNVWAYFSVKSGGGIHE-FSDSQ----FGHVFAFGESRALAIANMVLGLKEIQIRGEIRTNVDYTIDLLHASDYRDNKIHT 534
yeast 1012 SPETL----------------------------AKIIVHGKDRDDAISKLNQALEETKVYGCI-TNIDYLKSIITSDFFAKAKVST 1065

Figure 4.9 Amino acid sequence homology in the pyruvate-binding domain of yeast Pyc2

Shading indicates sequence identity with Pyc2, and clear boxes marked with asterisks show the amino acid differences present in the published Pyc2 sequence (Stucka *et al.*, 1991). The bent arrows surround the amino acids which align with the residues in the 5S subunit of transcarboxylase which were proposed to be involved in metal ion coordination (49 - 73) (Thornton *et al.*, 1993). The black circles numbered 4 - 7 indicate the position of cysteines "4" - "7" in Pyc1 that were changed to serine residues in the site-directed mutagenesis described in Chapter 7. The black boxing highlights potentially important residues or motifs including: the cysteine residues in each of the sequences, the tryptophan residues which correspond to the residue in the 5S subunit of transcarboxylase that was localised to the pyruvate binding site (W73) (Kumar *et al.*, 1988b), the histidine resides comprising the HXHXH motif which is may form part of the metal binding site (see section 4.3.4.3), some of the other residues which may be involved as ligands in the metal binding site of transcarboxylase in place of the first two histidine residues in the HXHXH motif (K200, D209, D211). The sequences shown were from the following sources: yeast Pyc2, the data presented in this thesis; yeast Pyc1, (Lim *et al.*, 1988); mosquito Pyc, Genbank; mouse Pyc, (Zhang *et al.*, 1993); rat Pyc, (Booker, 1990); human Pyc, (Walker *et al.*, 1995); *Klebsiella pneumoniae* Odc, (Schwarz *et al.*, 1988); *Salmonella typhimurium* Odc, (Woehlke *et al.*, 1992); 5S subunit of *Propionibacterium shermanii* Tc, (Thornton *et al.*, 1993). Abbreviations: Pyc, pyruvate carboxylase; Odc, oxaloacetate decarboxylase; Tc, transcarboxylase.

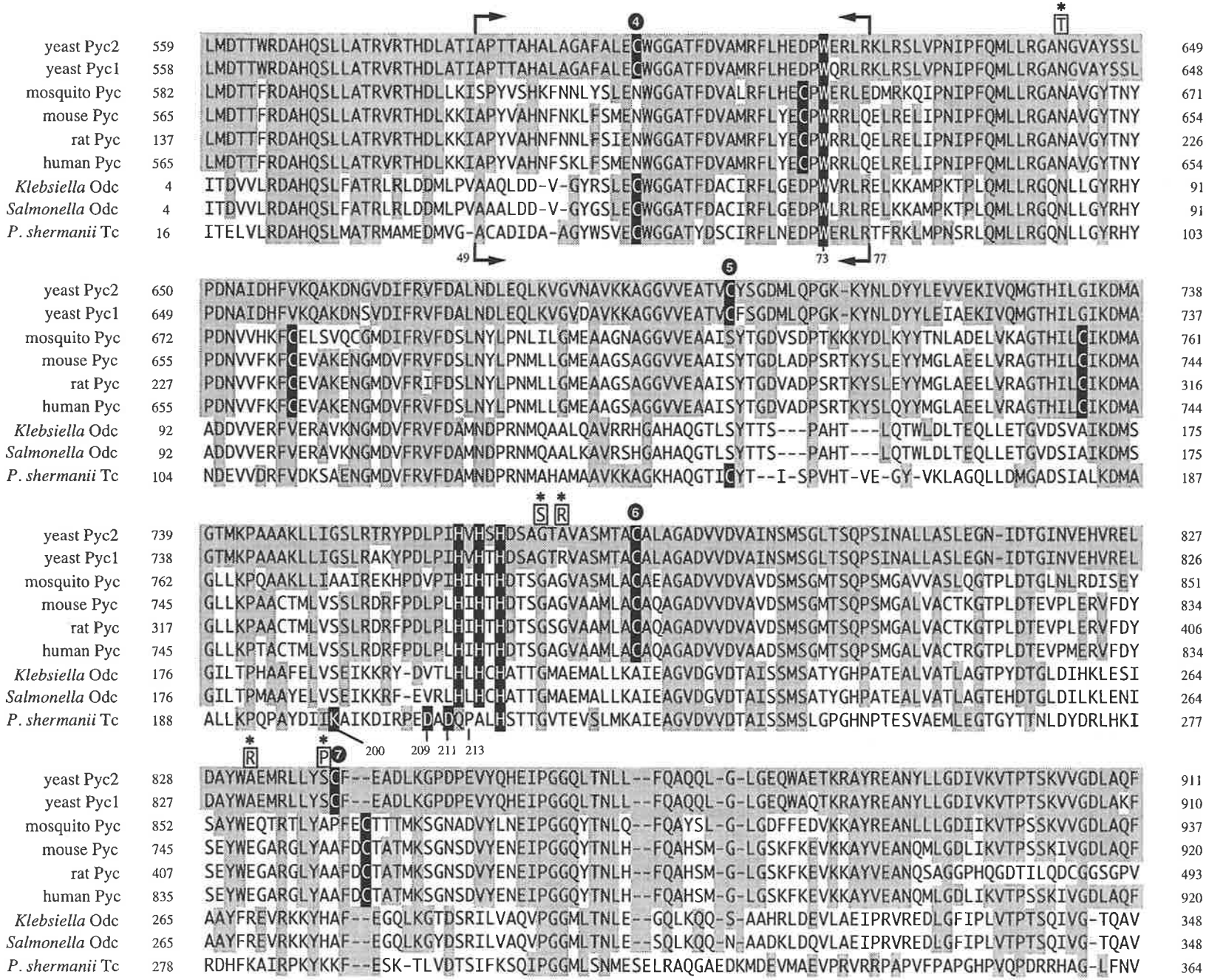
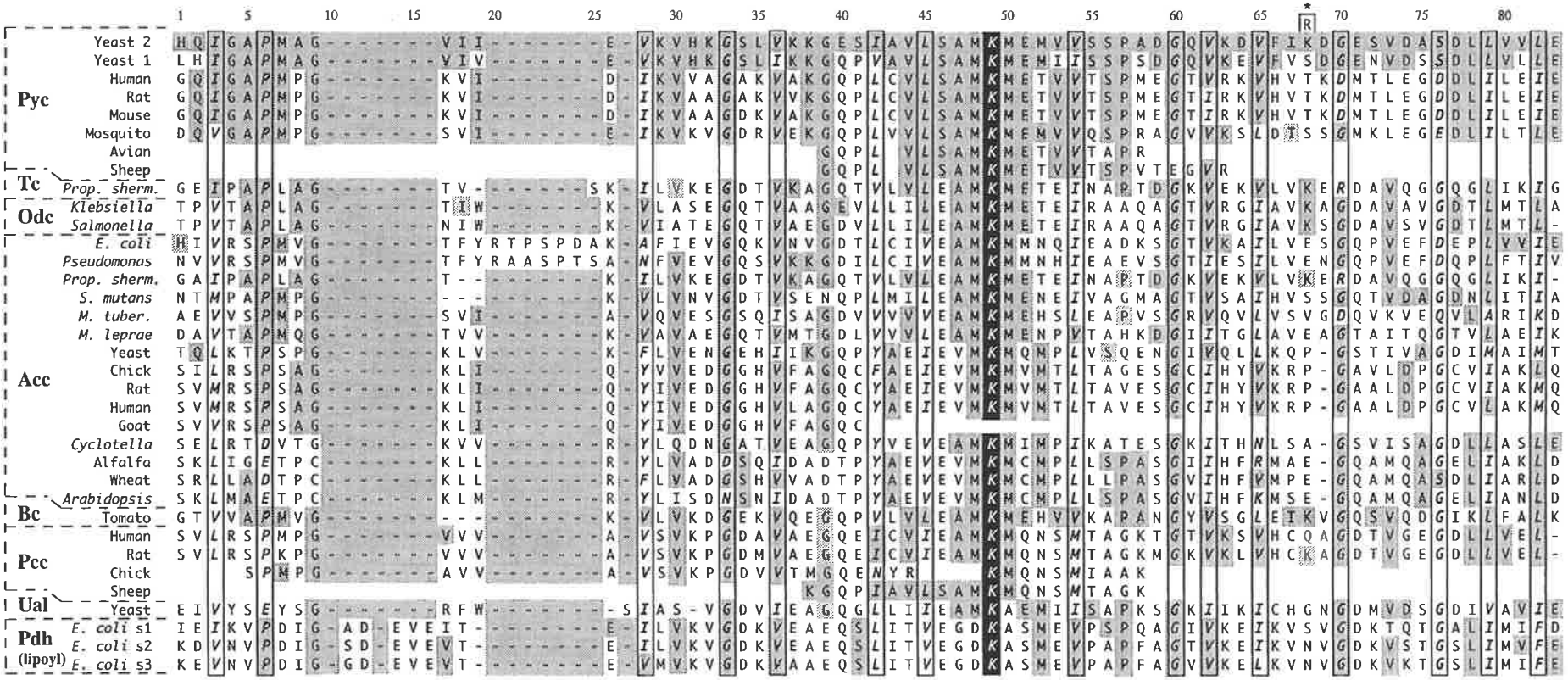

yeast Pyc2 559 LMDTTWRDAHQSLLATRVRTHDLATIAPTTAHALAGAFALECWGGATFDVAMRFLHEDPWERLRKLRSLVPNIPFQMLLRGANGVAYSSL 649
yeast Pyc1 558 LMDTTWRDAHQSLLATRVRTHDLATIAPTTAHALAGAFALECWGGATFDVAMRFLHEDPWQRLRKLRSLVPNIPFQMLLRGANGVAYSSL 648
mosquito Pyc 582 LMDTTFRDAHQSLLATRVRTHDLLKISPYVSHKFNNLYSLENWGGATFDVALRFLHECPWERLEDMRKQIPNIPFQMLLRGANAVGYTNY 671
mouse Pyc 565 LMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFNKLFSMENWGGATFDVAMRFLYECPWRRLQELRELIPNIPFQMLLRGANAVGYTNY 654
rat Pyc 137 LMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFNNLFSIENWGGATFDVAMRFLYECPWRRLQELRELIPNIPFQMLLRGANAVGYTNY 226
human Pyc 565 LMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFSKLFSMENWGGATFDVAMRFLYECPWRRLQELRELIPNIPFQMLLRGANAVGYTNY 654
Klebsiella Odc 4 ITDVVLRDAHQSLFATRLRLDDMLPVAAQLDD-V-GYRSLECWGGATFDACIRFLGEDPWVRLRELKKAMPKTPLQMLLRGQNLLGYRHY 91
Salmonella Odc 4 ITDVVLRDAHQSLFATRLRLDDMLPVAAALDD-V-GYGSLECWGGATFDACIRFLGEDPWLRLRELKKAMPKTPLQMLLRGQNLLGYRHY 91
P. shermanii Tc 16 ITELVLRDAHQSLMATRMAMEDMVG-ACADIDA-AGYWSVECWGGATYDSCIRFLNEDPWERLRTFRKLMPNSRLQMLLRGQNLLGYRHY 103
49
73
77
yeast Pyc2 650 PDNAIDHFVKQAKDNGVDIFRVFDALNDLEQLKVGVNAVKKAGGVVEATVCYSGDMLQPGK-KYNLDYYLEVVEKIVQMGTHILGIKDMA 738
yeast Pyc1 649 PDNAIDHFVKQAKDNSVDIFRVFDALNDLEQLKVGVDAVKKAGGVVEATVCFSGDMLQPGK-KYNLDYYLEIAEKIVQMGTHILGIKDMA 737
mosquito Pyc 672 PDNVVHKFCELSVQCGMDIFRVFDSLNYLPNLILGMEAAGNAGGVVEAAISYTGDVSDPTKKKYDLKYYTNLADELVKAGTHILCIKDMA 761
mouse Pyc 655 PDNVVFKFCEVAKENGMDVFRVFDSLNYLPNMLLGMEAAGSAGGVVEAAISYTGDLADPSRTKYSLEYYMGLAEELVRAGTHILCIKDMA 744
rat Pyc 227 PDNVVFKFCEVAKENGMDVFRIFDSLNYLPNMLLGMEAAGSAGGVVEAAISYTGDVADPSRTKYSLEYYMGLAEELVRAGTHILCIKDMA 316
human Pyc 655 PDNVVFKFCEVAKENGMDVFRVFDSLNYLPNMLLGMEAAGSAGGVVEAAISYTGDVADPSRTKYSLQYYMGLAEELVRAGTHILCIKDMA 744
Klebsiella Odc 92 ADDVVERFVERAVKNGMDVFRVFDAMNDPRNMQAALQAVRRHGAHAQGTLSYTTS---PAHT---LQTWLDLTEQLLETGVDSVAIKDMS 175
Salmonella Odc 92 ADDVVERFVERAVKNGMDVFRVFDAMNDPRNMQAALKAVRSHGAHAQGTLSYTTS---PAHT---LQTWLDLTEQLLETGVDSIAIKDMS 175
P. shermanii Tc 104 NDEVVDRFVDKSAENGMDVFRVFDAMNDPRNMAHAMAAVKKAGKHAQGTICYT--I-SPVHT-VE-GY-VKLAGQLLDMGADSIALKDMA 187
S
R
6
yeast Pyc2 739 GTMKPAAAKLLIGSLRTRYPDLPIHVHSHDSAGTAVASMTACALAGADVVDVAINSMSGLTSQPSINALLASLEGN-IDTGINVEHVREL 827
yeast Pyc1 738 GTMKPAAAKLLIGSLRAKYPDLPIHVHTHDSAGTRVASMTACALAGADVVDVAINSMSGLTSQPSINALLASLEGN-IDTGINVEHVREL 826
mosquito Pyc 762 GLLKPQAAKLLIAAIREKHPDVPIHIHTHDTSGAGVASMLACAEAGADVVDVAVDSMSGMTSQPSMGAVVASLQGTPLDTGLNLRDISEY 851
mouse Pyc 745 GLLKPAACTMLVSSLRDRFPDLPLHIHTHDTSGAGVAAMLACAQAGADVVDVAVDSMSGMTSQPSMGALVACTKGTPLDTEVPLERVFDY 834
rat Pyc 317 GLLKPAACTMLVSSLRDRFPDLPLHIHTHDTSGSGVAAMLACAQAGADVVDVAVDSMSGMTSQPSMGALVACTKGTPLDTEVPLERVFDY 406
human Pyc 745 GLLKPTACTMLVSSLRDRFPDLPLHIHTHDTSGAGVAAMLACAQAGADVVDVAADSMSGMTSQPSMGALVACTRGTPLDTEVPMERVFDY 834
Klebsiella Odc 176 GILTPHAAFELVSEIKKRY-DVTLHLHCHATTGMAEMALLKAIEAGVDGVDTAISSMSATYGHPATEALVATLAGTPYDTGLDIHKLESI 264
Salmonella Odc 176 GILTPMAAYELVSEIKKRF-EVRLHLHCHATTGMAEMALLKAIEAGVDGVDTAISSMSATYGHPATEALVATLAGTEHDTGLDILKLENI 264
P. shermanii Tc 188 ALLKPQPAYDIIKAIKDIRPEDADQPALHSTTGVTEVSLMKAIEAGVDVVDTAISSMSLGPGHNPTESVAEMLEGTGYTTNLDYDRLHKI 277
R
P
7
200
209
211
213
yeast Pyc2 828 DAYWAEMRLLYSCF--EADLKGPDPEVYQHEIPGGQLTNLL--FQAQQL-G-LGEQWAETKRAYREANYLLGDIVKVTPTSKVVGDLAQF 911
yeast Pyc1 827 DAYWAEMRLLYSCF--EADLKGPDPEVYQHEIPGGQLTNLL--FQAQQL-G-LGEQWAQTKRAYREANYLLGDIVKVTPTSKVVGDLAKF 910
mosquito Pyc 852 SAYWEQTRTLYAPFECTTTMKSGNADVYLNEIPGGQYTNLQ--FQAYSL-G-LGDFFEDVKKAYREANLLLGDIIKVTPSSKVVGDLAQF 937
mouse Pyc 745 SEYWEGARGLYAAFDCTATMKSGNSDVYENEIPGGQYTNLH--FQAHSM-G-LGSKFKEVKKAYVEANQMLGDLIKVTPSSKIVGDLAQF 920
rat Pyc 407 SEYWEGARGLYAAFDCTATMKSGNSDVYENEIPGGQYTNLH--FQAHSM-G-LGSKFKEVKKAYVEANQSAGGPHQGDTILQDCGGSGPV 493
human Pyc 835 SEYWEGARGLYAAFDCTATMKSGNSDVYENEIPGGQYTNLH--FQAHSM-G-LGSKFKEVKKAYVEANQMLGDLIKVTPSSKIVGDLAQF 920
Klebsiella Odc 265 AAYFREVRKKYHAF--EGQLKGTDSRILVAQVPGGMLTNLE--GQLKQQ-S-AAHRLDEVLAEIPRVREDLGFIPLVTPTSQIVG-TQAV 348
Salmonella Odc 265 AAYFREVRKKYHAF--EGQLKGYDSRILVAQVPGGMLTNLE--SQLKQQ-N-AADKLDQVLAEIPRVREDLGFIPLVTPTSQIVG-TQAV 348
P. shermanii Tc 278 RDHFKAIRPKYKKF--ESK-TLVDTSIFKSQIPGGMLSNMESELRAQGAEDKMDEVMAEVPRVRRPAPVFPAPGHPVQPDRRHAG-LFNV 364

Figure 4.10 Amino acid sequence homology in the biotin domain of yeast Pyc2

Shading indicates sequence identity with Pyc2, and clear boxes marked with asterisks show the amino acid differences present in the published Pyc2 sequence (Stucka *et al.*, 1991). The black boxing highlights the position of the biotinylated or lipoated (in the case of the Pdh lipoyl domains) lysine residue. Dashes within the sequences show spaces added to improve the alignments. The vertical dashes show the N- and C-terminal ends of the structured domain evident from NMR structures of biotinyl and lipoyl domains, and the larger lettering indicates those hydrophobic residues which have been identified as key folding residues in these structures (Brocklehurst and Perham, 1993; Brocklehurst *et al.*, 1995). The sequences shown were from the following sources: yeast Pyc2, the data presented in this thesis; yeast Pyc1, (Lim *et al.*, 1988); human Pyc, (Walker *et al.*, 1995); mouse Pyc, (Zhang *et al.*, 1993); mosquito Pyc, Genbank; avian (chicken and turkey) and sheep, (Rylatt *et al.*, 1977); *E. coli* Pdh1, 2 and 3, (Guest *et al.*, 1985); the 5S subunit of *Propionibacterium shermanii* Tc, (Maloy *et al.*, 1979; Murtif *et al.*, 1985) *Klebsiella pneumoniae* Odc, (Schwarz *et al.*, 1988); *Salmonella typhimurium* Odc, (Woehlke *et al.*, 1992); *E. coli* Acc, (Kondo *et al.*, 1991); *Pseudomonas aerogenosa* Acc, (Best and Knauf, 1993); *Propionibacterium shermanii* Acc, (Swiss-Prot database); *S. mutans* Acc, (Wang *et al.*, 1993); *M. tuber* Acc, (Norman *et al.*, 1994); *M. leprae* Acc, (Norman *et al.*, 1994); *A. thiana*, Acc, (Genbank); *Anabaena* sp. Acc, strain PCC 7120 (Gornicki *et al.*, 1993); *S. hirsuta* Acc, (Le Gouill *et al.*, 1993); yeast Acc, (Al-Feel *et al.*, 1992); chicken Acc, (Takai *et el.*, 1988); rat Acc, (Lopez-Casillas *et al.*, 1988); human Acc, (HA *et al.*, 1994); goat Acc, Genbank; *Cyclotella cryptica* Acc, (Roessler and Ohlrogge, 1993); alfalfa Acc, (Shorrosh *et al.*, 1994); wheat Acc, (Gornicki *et al*, 1994), *Arabadopsis* Acc, (Roesler *et al.*, 1994), tomatoa Bc, (Hoffman *et al.*, 1987); human Pcc, (Lamhonwah *et al.* 1989); rat Pcc, (Browner *et al.*, 1989); chicken Pcc, (Whittle, 1986); sheep Pcc, (Whittle, 1986). Abbreviations: Pyc, pyruvate carboxylase; Pdh, pyruvate dehydrogenase; Tc, transcarboxylase; Odc, oxaloacetate decarboxylase; Acc, acetyl-CoA carboxylase; Pcc, propionyl-CoA carboxylase.


1
5
10
15
20
25
30
35
40
45
50
55
60
65
*
R
70
75
80
Pyc
Yeast 2
Yeast 1
Human
Rat
Mouse
Mosquito
Avian
Sheep
Tc
Prop. sherm.
Odc
Klebsiella
Salmonella
E. coli
Pseudomonas
Prop. sherm.
S. mutans
M. tuber.
M. leprae
Acc
Yeast
Chick
Rat
Human
Goat
Cyclotella
Alfalfa
Wheat
Arabidopsis
Bc
Tomato
Pcc
Human
Rat
Chick
Sheep
Ual
Yeast
Pdh
(lipoyl)
E. coli s1
E. coli s2
E. coli s3
1170
1168

homologous region, and these workers suggested that this may be a consequence of a duplication event.

Despite this low level of identity among biotin domains, Brocklehurst and Perham (1993) recognised that yeast Pyc1 has an identical pattern of the hydrophobic residues shown to have a key structural role in lipoyl domains (See Fig. 4.10). On this basis they calculated that yeast Pyc1 has a structure very similar to the Pdh lipoyl domains of *E. coli* and *B. stearothermophilus*..

The remaining 42 amino acid differences between the two isozymes are primarily clustered near the N-terminus, or in the "spacer" region towards the C-terminus between the pyruvate and biotin domains.

Given the number of differences that occur between Pyc1 and Pyc2, it would be difficult to speculate as to possible effect of any of the individual differences, even allowing for the fact that many of them are conservative changes and thus may not need to be considered. Furthermore at this stage the only enzymatic difference that has been observed between the two isozymes is a difference in specific activity, ie Pyc1 appears to have a specific activity approximately twice that of Pyc2 (M. Walker, pers. comm). Obviously then, before meaningful predictions can be made regarding the effect of any of the sequence differences a detailed comparative study on the enzymatic properties of the two isozymes would need to be performed .

4.3.3 Predictions based on sequence comparisons

4.3.4.1 ATP binding site

The universal importance of ATP throughout nature is evident from the occurrence of ATP binding sites in a multitude of proteins with widely different functions. Three dimensional structures have been determined for many of these ATP binding proteins, and numerous genes encoding proteins with ATP binding sites have been sequenced. Comparison of the protein sequences has revealed that a common feature of ATP and GTP binding sites is a short glycine rich sequence involved in the formation of a phosphate binding "P-loop" structure (Saraste *et al.*, 1990), originally referred to as "motif A" (Walker *et al.*, 1982) (Fig. 4.11). In fact this glycine rich sequence is also present in NAD and FAD

binding proteins in the form of a **GXGXXG** sequence (X represents any residue) (Sternberg and Taylor, 1984), while the actual P-loop sequence motif common to many families of ATP binding proteins is **GXXXXGK**[T / S]. However it should be noted that not all ATP binding proteins contain this motif, and the presence of this motif does not necessarily indicate that these residues are involved in ATP binding (Saraste *et al.*, 1990).

Figure 4.8a reveals that the majority of ATP binding biotin carboxylases contain the sequence **GahP**[ah]$_3$**KAXXGGGG**[R / K]**GahR** (where ah represents an aliphatic hydrophobic residue), while most Cps enzymes have the similar **GahPX**[ah]$_2$**RXX**[Y / F]**XahGGXG**[S / G]**G** sequence. Both these consensus sequences contain the **GXXXXG** sequence, but they lack the final K and [T / S] residues present in the general P-loop motif.

The presence of the **GXXXXG** motif in yeast Cps led Lusty *et al.* (1983) to propose that this sequence comprises part of the ATP binding site. On the basis of sequence comparisons to ATP binding proteins with known crystal structures, these workers proposed that the residues from 327 - 355 in the yeast enzyme may similarly be involved in ATP binding (see Fig. 4.8c). Post and co-workers (1990) have investigated the role of the **GXXXXG** motif in *E. coli* Cps by the use of site-directed mutagenesis. Upon changing either **G**$_{176}$ or **G**$_{180}$ of the **GG**$_{176}$**XG**[S / G]**G**$_{180}$ putative P-loop motif to more bulky isoleucine residues, they found that the V_{max} of the bicarbonate-dependent ATPase reaction was reduced by >90 %, suggesting that this is in fact the ATP binding site. On the basis of these results Kondo *et al.* (1991) and Li and Cronan (1992) proposed that the corresponding region in biotin carboxylases is probably also the site of ATP binding.

Recently Waldrop and co-worker have solved the crystal structure of the *E. coli* biotin carboxylase (Bc) subunit of Acc to a resolution of 2.4 angstroms (Waldrop *et al.*, 1994 a & b). Surprisingly the crystal structure revealed that biotin carboxylase does not appear to contain the "P-loop" ATP binding motif. Instead the above mentioned glycine rich motif is part of a small domain of unknown function ("B-domain") which protrudes out of the main body of the protein, and is ill-defined due conformational flexibility.

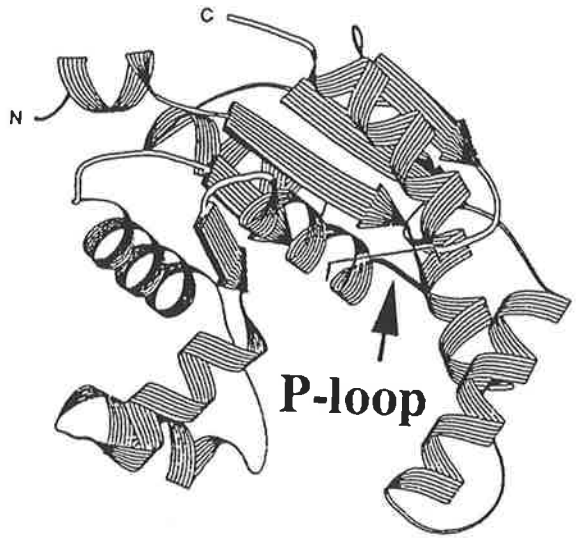


Figure 4.11 Schematic diagram of the "P-loop" nucleotide fold (adapted from Saraste *et al.*, 1990)

A further unexpected characteristic was the finding that the N-terminal 103 residues are folded in the form of a di-nucleotide binding motif with the structural architecture of the class typical of NAD^+ -dependent dehydrogenases and UDP-galactose 4-epimerase, rather than that of ATP binding proteins. As it is outside of the "ATP domain" sequence defined by sequence homology (which starts from V128, see Fig. 4.8) this raises the possibility that this dinucleotide domain may be unique to Acc enzymes, ie all Acc enzymes also have high homology in this region (Gornicki *et al.*, 1994). Alternatively it may be present in most or all of the enzymes included in figure 4.8, but the similarity of all the enzymes in this region was not evident on the basis of sequence identity. Similarly, the structural similarities of biotinyl and lipoyl domains would probably have gone unnoticed were it not for the fact that they both contain a similarly placed lysine residue to which their respective prosthetic groups are attached.

Although at this stage the resolution of the structure and the lack of detailed substrate analogue binding studies preclude an accurate description of the interactions between the substrates and the active site, Waldrop *et al.* (1994b) were able to locate an area of electron

density which appeared to match that expected for a phosphate ion (the crystals were grown in K_3HPO_4). This "phosphate" ion was surrounded by K238, R292, Q294, V295, and R338 (black boxed residues in Fig. 4.8b & c). It is of interest to note that the sequence comparisons shown in Fig. 4.8b & c reveal that of the 5 residues which allegedly surround the "phosphate" ion, only R292 is present in Cps enzymes as well as the biotin carboxylases. In fact 3 of these residues occur in a QVEH motif which is present in all the biotin carboxylase, but is absent from Cps enzymes which all have in its place an SRSS sequence (apart from human Cps which allegedly has SPNS, although on this basis it would seem possible this is actually due to a sequencing error).

4.3.4.2 Other residues in the 1st partial reaction site

Upon comparing their sequence for the *E. coli* Bc subunit of Acc with the rat Acc and *E. coli* Cps sequences, Li and Cronan (1992) noticed that downstream of the G-rich putative "P-loop" sequence there was a stretch of 16 amino acids from R228 to A243 (in *E. coli* Bc) which appeared to be unique to biotin carboxylases. Although the alignment shown in figure 4.9b between the Cps and biotin carboxylase sequences in this region differs from that presented in the report by Li and Cronan (1992) (due to the inclusion of many more sequences), the dissimilarity of the Cps enzymes is still apparent.

As the major difference between the ATP / bicarbonate reaction catalysed by Cps enzymes is the fate of the carboxy-phosphate intermediate, these workers went on to propose that this stretch of residues in *E. coli* Bc may be involved in the carboxylation of biotin by carboxy-phosphate. The other piece of data which led them to this proposal was the fact that *E. coli* Bc and rat Acc both contained a cysteine residue within this "biotin carboxylation motif". This agreed with the evidence from kinetic pH profiles and deuterium fractionation factor effects with *E. coli* Bc, suggesting that a thiol group is probably involved as a proton donor / acceptor in the carboxylation of biotin by carboxy-phosphate (Tipton and Cleland, 1988 a & b) (see Chapter 1, 1.4.1.3). Indeed the extensive sequence comparisons of the ATP domain sequences shown in figure 4.9 reveals that all Class I biotin enzymes contain this cysteine residue, adding further weight to this proposal. Similarly, cysteine 360 (Pyc1) would also appear to be an attractive candidate for the catalytic cysteine as it is contained in

all Pyc, all Pcc, and 2 / 3 of the Acc enzymes sequenced to date (see Fig. 4.8). Furthermore, the corresponding cysteine in Bc (237) is adjacent to R238, which is one of the residues believed to surround the "active site" phosphate ion in crystal structure of Bc (Waldrop, *et al.*, 1994b).

Actually the kinetic data of Tipton and Cleland (1988 a & b) led them to propose that a cysteine residue was involved in catalysing the enolisation of biotin (see Chapter 1, 1.4.1.3) in the form of a base-sulfhydryl ion pair. To this end figure 4.8b reveals that within this "biotin carboxylation motif" there are 5 residues (R, K, or H) which could be considered as candidates for the base involved in this proposed ion pair with a cysteine group. However, there is also some evidence to suggest that the biotin carboxylation reaction may not involve an enzymic cysteine / cysteine-lysine ion pair. For example, Ahmad *et al.* (1984) found that the substrates for the biotin carboxylation reaction were unable to convey protection against inactivation of rat Acc by the sulfhydryl modifying reagent *p*-hydroxymercuribenzoate (PHMB). Also, contrary to what was previously commonly believed, it has recently been noted that deuterium fractionation factors less than 0.5 are not necessarily diagnostic of the involvement of an enzymic sulfhydryl group. Low-barrier hydrogen bonds which form between two groups with similar pK's (eg two carboxylate or two phenolic groups) can also exhibit fractionation factors less than 0.5.

Further inspection of the ATP domain sequences reveals that all Class I biotin enzymes contain a QVEH motif (residues 315 - 318 in Pyc2) which is unique to these enzymes, as Cps enzymes have an SRSS sequence in its place. Hence, in addition to those mentioned above, these residues may also play important catalytic or substrate binding roles in the chemistry of the biotin carboxylation reaction. Indeed the preliminary active site localisation experiments recently performed by Waldrop *et al.* (1994b) tend to confirm these sequence based predictions.

Waldrop *et al.* (1994b) determined the area of electron density difference (3 Angstroms resolution) between the *E. coli* Bc native crystal, and the crystal soaked in a solution of Ag^+ / biotin complex. They calculated that this area could easily accommodate a biotin group, which in entering the active site appeared to have displaced two water molecules and the phosphate ion. The amino acid residues in *E. coli* Bc which surround this

area of electron density difference (putative biotin group) are H236, K238, E276, E288, N290, R292, Q294 and E296 (boxed residues in Fig. 4.8b). Thus these results suggest that the unique QVEH motif in Class I biotin carboxylases may indeed be playing an important role in the chemistry of the reactions at the 1st partial reaction site which is unique to biotin enzymes, ie the carboxylation of biotin by the carboxyl-phosphate intermediate.

4.3.4.2 The 2nd partial reaction site

Figure 4.10 reveals that the sequence identity between the pyruvate domain of Pyc2 and the corresponding domains of other biotin enzymes which carry out the same transcarboxylation reaction (or the reverse reaction) is remarkably high throughout essentially the whole domain. Therefore, on the basis of sequence comparisons alone it would seem almost impossible to make any meaningful predictions about the relative importance of particular residues within this domain with regard to their role in the transcarboxylation reaction.

The most significant study with regard to localising the active site residue(s) within the pyruvate domain came from work with the enzyme transcarboxylase (Tc). Kumar *et al.* (1988a) found that pyruvate reduced the intrinsic tryptophanyl fluorescence quenching of the 5S subunit of Tc, suggesting that certain tryptophan residues may be involved in pyruvate binding. Furthermore, upon characterising the tryptic peptides produced after modification with the tryptophan specific reagent 2,4-dinitrophenylsulfenyl chloride (DNPS-Cl) in the presence or absence of pyruvate, they found that W73 was protected by pyruvate binding. They concluded that this residue is at or near the pyruvate binding site, and suggested that it may be close to the Co^{2+} / Zn^{2+} binding site. On the basis of sequence comparisons between the 5S subunit of Tc and other "pyruvate domain" enzymes, workers from the same group further proposed that in addition to W73, the surrounding residues from 49 -77 (see Fig. 4.9) may be involved in metal ion coordination (Thornton *et al.*, 1993).

4.3.4.3 The metal binding site

As discussed in Chapter 1 section 1.5.2, it is well known that yeast Pyc(s) contains a tightly bound Zn^{2+} ion, while the vertebrate enzymes contain a bound Mn^{2+} ion. The metal

ion in Pyc is believed to have a structural role, being located close to the pyruvate binding site but too far away to be directly involved in the reaction (reviewed in Scrutton *et al.*, 1973b). Recently Dimroth and Thomer (1992) have shown that oxaloacetate decarboxylase (Odc) from *Klebsiella pneumoniae* also contains a bound Zn^{2+} ion. In addition, on the basis of its extremely high sequence homology with the *Klebsiella* enzyme (Woehlke *et al.*, 1992), it would be expected that Odc from *Salmonella typhimurium* probably also contains a bound Zn^{2+} ion. Furthermore, given the high level of sequence similarity between Odc and Pyc enzymes within the pyruvate domain, the Zn^{2+} ion in both these enzymes probably has an analogous structural role, and is similarly located within the domain.

The enzyme transcarboxylase differs from the above mentioned enzymes in that it contains a mixture of both Co^{2+} and Zn^{2+} ions (Northrop and Wood, 1969). Ahmad *et al.* (1972) demonstrated by a variety of experiments that both metals are tightly bound to the $5S_E$ component, which consists of a homodimer of the 60 kDa pyruvate binding / transcarboxylation subunits. Furthermore, from the ratio of components present in the subunit structure revealed by electron microscopy, Green *et al.* (1972) calculated that there were 2 metals per $5S_E$ component, leading Ahmad *et al.* (1972) to conclude that there is one Co^{2+} or Zn^{2+} per pyruvate binding subunit. These workers also reported that the ratio of Co^{2+} to Zn^{2+} ions in the enzyme varied widely depending on the composition of the growth media, suggesting that the metal binding site has a similar affinity for these two ions.

Figure 4.9 shows a sequence comparison for the pyruvate domains of the Pyc, Odc and Tc enzymes. Inspection of this comparison reveals that the Pyc and Odc enzymes all have a HXHXH motif (residues 763, 765, 767 in Pyc2; black boxed residues), while the Tc enzyme only has the last H residue of this motif. In the light of the ability of histidine residues to be coordinated to metal ions (discussed below) it is tempting to speculate that this HXHXH motif may form part of the metal ion binding site in the Pyc and Odc enzymes. In the case of the Tc enzyme, other suitable neighbouring residues could be providing the necessary coordination of the metal ion in the binding site, instead of the first two histidine residues in this motif. This may explain the different properties of the metal ion binding site of the Tc enzyme, ie the greater flexibility in the metal ion content (Co^{2+} or Zn^{2+} ions).

Clearly this proposal is different from the proposal of Thornton *et al.* (1993) (see section 4.3.4.2), as they suggested that residues from 49 - 77 in the Tc enzyme may be involved in metal ion coordination (see Fig. 4.9). These workers made this proposal on the basis that a) W73 was shown to be near the pyruvate binding site in Tc (Kumar *et al.*, 1988a), b) the metal ion in Tc has also been shown to be near the pyruvate binding site (Northrop and Wood, 1969; Fung *et al.*, 1974) and c) there is a high level of sequence identity between all the enzymes in this region. However, while reasons a) and b) are consistent with this proposal, they in no way constitute a proof and thus do not exclude other possibilities. Furthermore, I would suggest that reason c) could more correctly be interpreted as evidence against, rather than for this proposal. In other words, it is would seem unlikely that the metal ion binding site would be located in a stretch of amino acids which is almost totally identical in all the enzymes, given the differences in metal ion specificity between the various enzymes and species.

With regard to distinguishing between the Mn^{2+} binding site of vertebrate Pycs and the Zn^{2+} site of the yeast enzymes, comparatively little is known about the structure of Mn^{2+} binding sites. Apart from vertebrate Pycs the other well known metallo-enzymes which contain a bound manganese ion are bacterial superoxide dismutases. The Mn^{3+} ion in the enzyme from *Thermus thermophilus* is coordinated in a trigonal bi-pyramidal arrangement to three histidine residues, an aspartic acid and a water molecule (Stallings *et al.*, 1985). However the coordination in vertebrate Pycs would clearly be different as the manganese ion is instead in the +2 oxidation state, and has a structural rather than catalytic role. D-Xylose isomerase, although generally regarded as magnesium metalloprotein, could also be considered as a manganese metalloprotein as its active site can be occupied by two metals from the group (Mg^{2+}, Mn^{2+} or Co^{2+}) (Glusker, 1991). The structure of the Mn^{2+} coordination sites in this enzyme is shown in figure 4.12c.

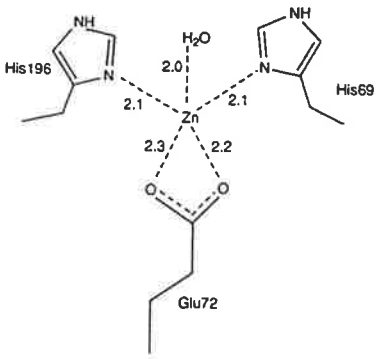
Regarding Zn^{2+} binding sites, it is known that Zn coordination polyhedra exhibit variations in ligand number, charge, structure and amino acid composition depending on the particular role of the metal ion (reviewed in Vallee and Auld, 1990). From the 12 zinc enzymes with known crystal structures (at the time of this review) it has been found that Zn^{2+} ions with catalytic roles all have the Zn^{2+} ion coordinated to a water molecule and three

amino acids which are either histidine, cysteine, aspartic acid or glutamic acid residues. Examples of these arrangements are shown in figures 4.12a & b. On the other hand, the structural Zn^{2+} ion in alcohol dehydrogenase and aspartate transcarbamylase are both coordinated by 4 cysteine residues each separated by from 2 - 22 amino acids (Vallee and Auld, 1990).

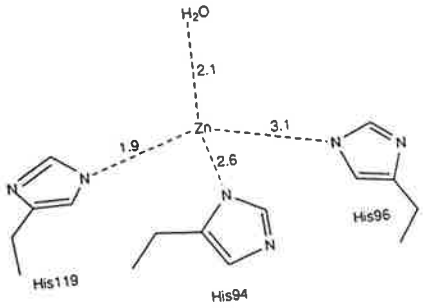
On the basis of the 4 x cysteine coordination of the two characterised structural zinc binding sites, the HXHXH motif may be considered as a poor candidate for a structural zinc binding site. However, the HXHXH arrangement of ligands would appear to be compatible with the geometry required for Zn^{2+} coordination, as carbonic anhydrases all contain the HXH motif as part of the metal liganding site (Vallee and Auld, 1990), and metalloendopeptidases all have the HEXXH or HXXEH motif (Jiang and Bond, 1992). Furthermore Mozier *et al.*, (1991) recently reported their findings that protein kinase C inhibitor-1 (PKCI-1) contains what they describe as a "novel zinc binding site" involving the HVHLH sequence. Although they have not established the role of the HXHXH motif of PKCI-1 *in vivo*, a 24 amino acid tryptic peptide of PKCI-1 containing the HVHLH, and a number of synthetic peptides containing various length segments of the PKCI-1 sequence spanning the HVHLH site were all found to bind a single $^{65}Zn^{2+}$ ion. Taken together these data would seem to constitute sufficient evidence upon which to base the hypothesis that the HXHXH sequence in yeast Pyc2, and indeed all Pyc and Odc enzymes may constitute part of the metal binding site.

If this were the case the different metal ion specificity of the vertebrate Pycs could be explained by differences in conformation, or the type of remaining ligand(s) making up this site. In the case of transcarboxylase, as the last H residue in this motif is also present in this enzyme, it may be that this H residue along with other neighbouring residues unique to Tc (possibly residues 200, 209, and 211; black boxed residues in Fig. 4.9) are involved as metal binding ligands in the Zn^{2+} / Co^{2+} site.

a)



b)



c)

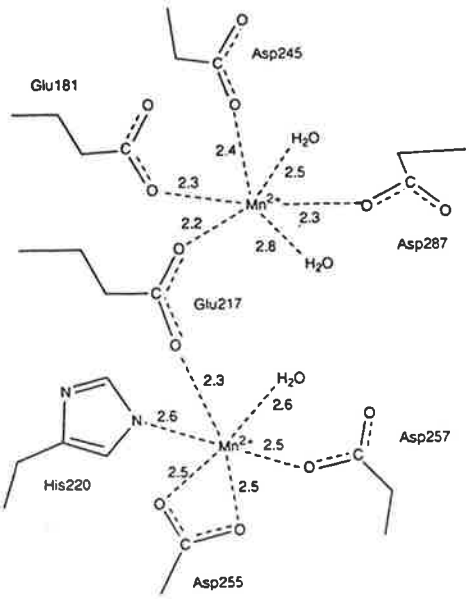


Figure 4.12 Structures of typical Zn(II) and Mn(II) binding sites, from Glusker (1991).

Atomic arrangements around the: **(a)** catalytic Zn(II) ion in carboxypeptidase, **(b)** catalytic Zn(II) ion in carbonate dehydratase, and **(c)** catalytic Mn(II) ions in D-xylose isomerase.

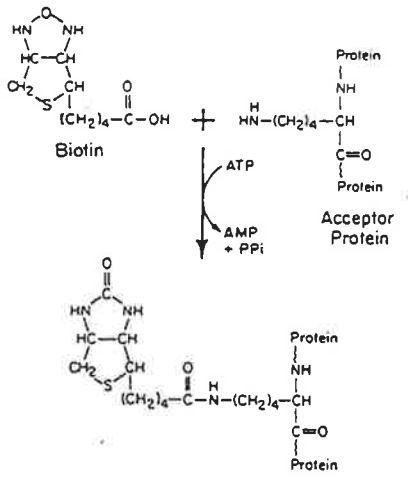
CHAPTER 5

EXPRESSION BIOTINYLATION AND PURIFICATION OF Pyc BIOTIN DOMAIN PEPTIDES

5.1 INTRODUCTION

Biotin (vitamin H) is synthesised endogenously by plants, most bacteria, and some fungi (Ogata, 1970), while other organisms rely on exogenous sources of this vitamin. Yeast on the other hand are unusual in that they are only able to carry out the last two steps in the biotin biosynthesis pathway (Ogata, 1970; Zhang *et al.*, 1994). Despite these differences, the essential role played by biotin in all cells is in each case dependent on the specific covalent attachment of this co-factor to a few key enzymes which utilise biotin as a mobile carboxyl carrier.

The specific attachment of biotin to the acceptor proteins is carried out by biotin ligase enzymes (also called holocarboxylase synthetases), and this reaction involves an ATP dependent formation of an amide bond between the carboxyl group of biotin, and the ε-amino group of a specific lysine in the biotin domain of the acceptor protein (Lane *et al.*, 1964) as outlined in Scheme VI.



Scheme VI The biotinylation reaction, from Cronan (1989).

This process is known to be a post-translational event, requiring correctly folded domains as substrates (Wood *et al.*, 1980; Reed and Cronan, 1991). Furthermore, biotin ligase enzymes from various species have been found to recognise and biotinylate acceptor proteins from very different species both *in vivo* (Cronan, 1990) and *in vitro* (McAllister and Coon, 1966). This is presumably because biotin domains all have similar structures. Indeed, upon

recognising that yeast Pyc1 had the same pattern of the key hydrophobic folding residues identified in the structures of lipoyl domains, Brocklehurst and Perham (1993) predicted a structure for yeast Pyc1 which was very similar to the structures of lipoyl domains.

As yet however, the specific structural features of biotin domains that are responsible for their precise recognition and biotinylation have not yet been determined, despite the fact that the *E. coli* biotin ligase enzyme (Wilson *et al.*, 1992) and BCCP biotin domain (Brocklehurst *et al.*, 1995) structures have recently been solved.

As previously outlined, the comparison between the published *PYC2* sequence and the sequence of the *PYC2* gene described in Chapter 4 indicates that the published sequence differs in the region encoding the biotin domain (3' end) by having a transversion resulting in a conservative amino acid change (K1155R), and a single base insertion at +3533. This insertion produces the most significant difference between the two Pyc2 sequences, as it results in a frame shift changing Q1178 into a proline, and extending the coding sequence by 5 amino acids (VIFTR extension) (Fig. 4.6).

Clearly, one way to confirm the predicted amino acid differences in the Pyc2 biotin domain is to obtain an accurate molecular weight for a purified C-terminal peptide containing the biotin domain. Then, by expressing the other forms of the Pyc2 biotin domain with respect to the two sequence differences found in this region, the effects of these differences on the degree of biotinylation can be determined. This may also extend our knowledge of the structural features which are involved in recognition and post-translational modification by biotin ligase enzymes. Having determined an efficient method for expressing and purifying a Pyc2 biotin domain, it should be possible to scale up the procedure to obtain sufficient material for 3D structure determination by NMR. This would be important as it would represent the first 3D structural data for a pyruvate carboxylase enzyme, and only the second biotin domain structure known to date. Furthermore, an additional solved biotin domain structure to compare with the known BCCP structure (Brocklehurst *et al.*, 1995) could be extremely valuable in identifying the common structural features which allow these substrates to act as substrates for biotinylation.

This chapter describes the:

(i) expression, biotinylation, and purification of the biotin domain encoded by the *PYC2* gene described in Chapter 4

(ii) N-terminal sequencing and mass spectrometric analysis of the purified Pyc2 biotin domain

(ii) expression and comparative biotinylation *in vivo* of various Pyc biotin domain peptides by *E. coli* biotin ligase

5.2 RESULTS

5.2.1 Expression and biotinylation of the Pyc2 biotin domain peptide

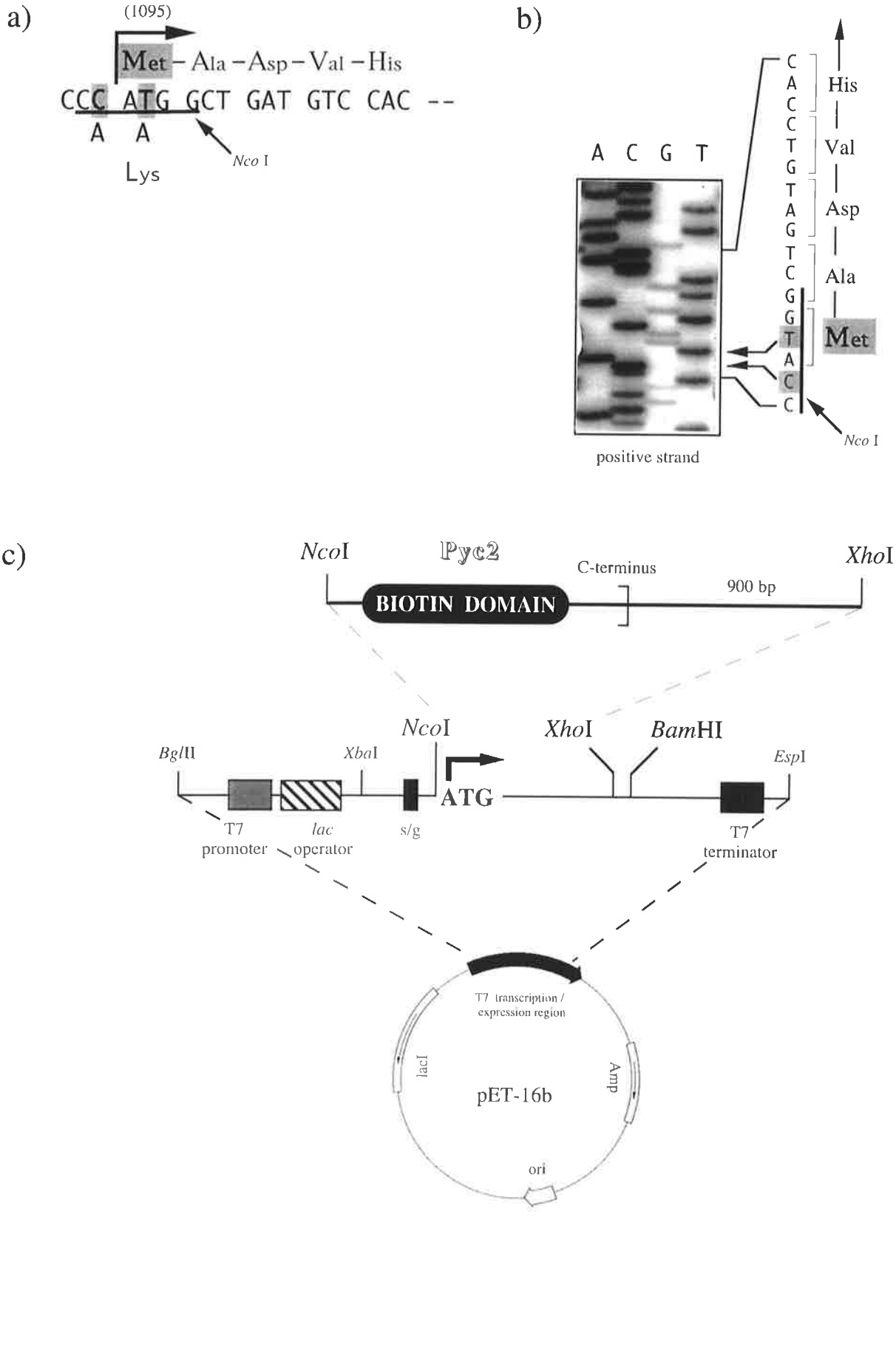
As Dr A Chapman-Smith in our laboratory had recently found that the *E. coli* T7 pET expression system (pET-16B; from Novagen Inc.) was able to produce high levels of expression of the biotin domain from the BCCP subunit of *E. coli* Acc (Chapman-Smith *et al.*, 1994), a construct was engineered to express the Pyc2 biotin domain in *E. coli* using the same system. Subcloning and subsequent expression of the Pyc2 biotin domain peptide using the pET expression system required the introduction of an initiating methionine (K1095M) and *Nco*I cloning site 5' of the start of the biotin domain (at +3281 in *PYC2* ; Fig. 5.1a). This was performed by site directed mutagenesis using a modification of the method of Kunkel *et al.* (1987) (using unmodified T7 DNA polymerase instead of Klenow DNA polymerase I) with the oligonucleotide O3671, and clone p58 DNA (see Chapter 4). The DNA sequence showing the introduced *Nco*I site is shown in figure 5.1b.

The construct required to express the Pyc2 biotin domain peptide corresponding to the *PYC2* DNA sequence reported in Chapter 4 was then obtained by subcloning the 3' end of *PYC2* as a 900 bp *Nco*I-*Xho*I fragment into the pET-16b expression vector (see Fig. 5.1c), thus allowing the C-terminal end of Pyc2 containing the entire biotin domain (c86K; Fig. 5.6a) to be expressed as a recombinant peptide from an engineered methionine residue (K1095M) to the native C-terminus.

Expression was induced with IPTG in the presence of excess biotin (10 mM) and biotin ligase (*E. coli bir*A gene product). These expression conditions were used in order to

Figure 5.1 Construct to express the C-terminal Pyc2 biotin domain peptide

a) Schematic representation of the sequence changes required to introduce an *Nco*I site and initiating methionine upstream of the start of the Pyc2 biotin domain. The wild type nucleotide and amino acid sequences are shown below. **(b)** Diagram showing the mutagenic changes in the nucleotide sequence of the DNA used as a template to express the C-terminal Pyc2 biotin domain peptide. **(c)** Schematic diagram of the T7 expression construct engineered to express the biotin domain of Pyc2. Arrows and underlining indicate the position of the introduced *Nco*I site, shading shows the introduced sequence changes, and S/G indicates the position of the Shine-Dalgarno sequence.


a)
(1095)
Met – Ala – Asp – Val – His
CCC ATG GCT GAT GTC CAC --
A A
Lys
Nco I
b)
A C G T
C A C C T G T A G T C G G T A C C
His
Val
Asp
Ala
Met
Nco I
positive strand
c)
NcoI
Pyc2
C-terminus
XhoI
BIOTIN DOMAIN
900 bp
NcoI
XhoI
BamHI
BglII
XbaI
EspI
ATG
T7 promoter
lac operator
s/g
T7 terminator
T7 transcription / expression region
lacI
Amp
pET-16b
ori

ensure that the extent of biotinylation would not be limited by the quantity of either biotin, or the biotin ligase enzyme. IPTG-induced the T7 expression of the Pyc2 domain by inducing the genomic expression of T7 RNA polymerase from the *lac*UV5 promoter, while the excess biotin ligase was supplied from a *araBAD* promoter on the plasmid pCY216 (Chapman-Smith *et al.*, 1994) in the presence of arabinose. A schematic representation of the events involved in the expression and *in vivo* biotinylation of the Pyc2 biotin domain peptide is presented in figure 5.2.

The SDS-PAGE gel shown in figure 5.3a (Lane 1) shows that the Pyc2 biotin domain peptide was indeed present in the induced lysate, and avidin alkaline phosphatase detection clearly shows that this peptide had been biotinylated *in vivo* by *E. coli* biotin ligase (Lane 1, Fig. 5.3b).

When expressing the BCCP subunit of *E. coli* Acc using the same system, Dr Chapman-Smith found that significant amounts of the expressed peptide remained as apo-protein even in the presence of excess biotin and biotin ligase, apparently due to insufficient ATP being available for the ligase reaction (Chapman-Smith *et al.*, 1994). Hence, by adding excess ATP it was possible to complete the biotinylation of the Pyc2 biotin domain *in vitro*, and by adding ^{3}H-biotin the Pyc2 peptide could be labelled, thus facilitating its detection during the purification.

5.2.2 Purification of the Pyc2 biotin domain peptide

In order to obtain sufficient Pyc2 biotin domain peptide material for N-terminal sequencing and molecular weight determination by electrospray mass spectrometry it was necessary to develop a method for the purification of this peptide.

In 1980 Rock and Cronan reported a simple method for purifying the *E. coli* acyl carrier protein (ACP) involving an isopropanol precipitation step. Essentially they were able to achieve a one step purification of ACP, as this small acidic protein was one of only a few of *E. coli* proteins that remained soluble in 60% isopropanol. More recently Chapman-Smith *et al.* (1994) found that the similarly small and acidic (cal. pKi = 4.35) 87 amino acid BCCP peptide also remained soluble in 60% isopropanol, thus greatly simplifying its purification.

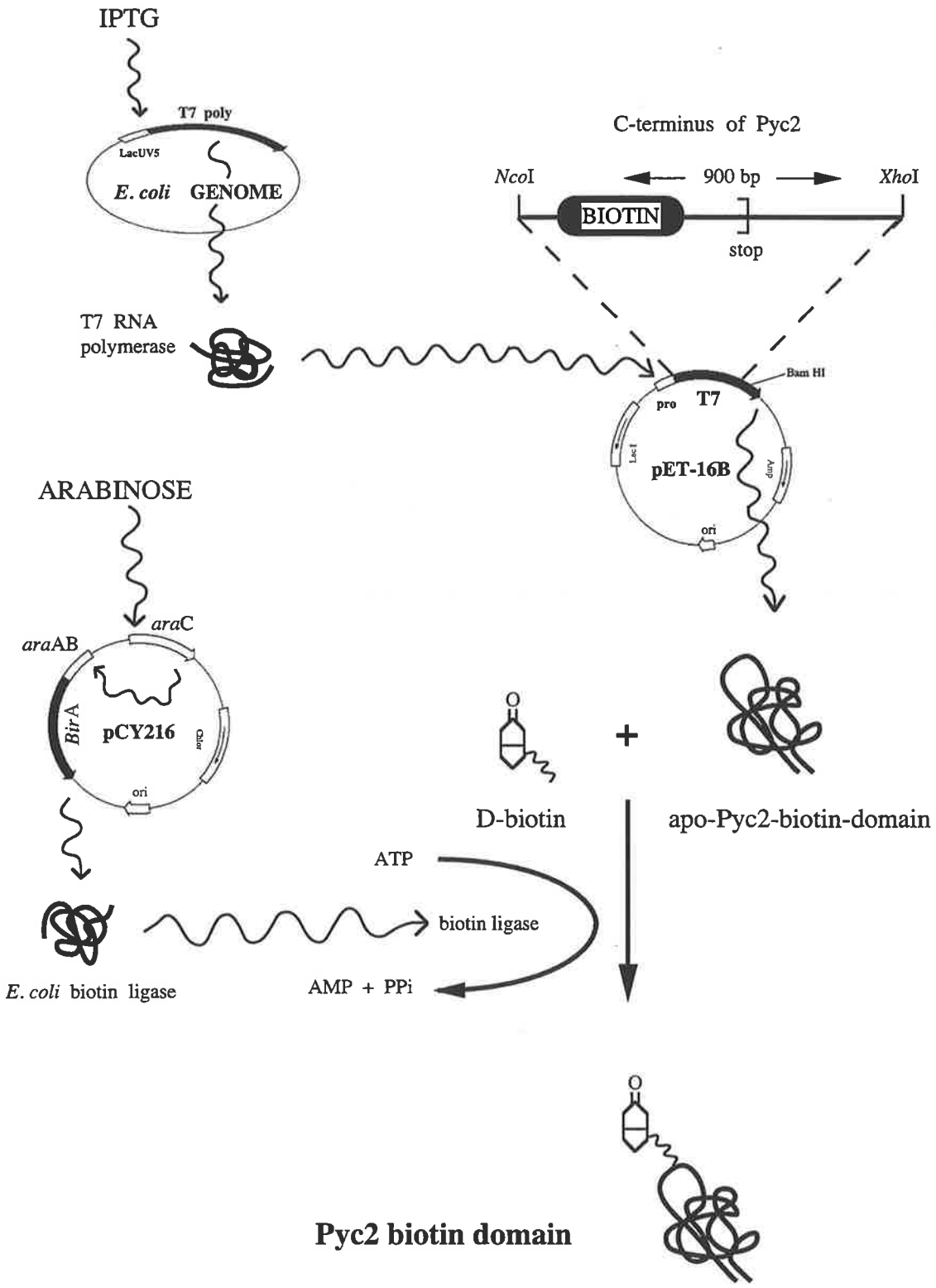
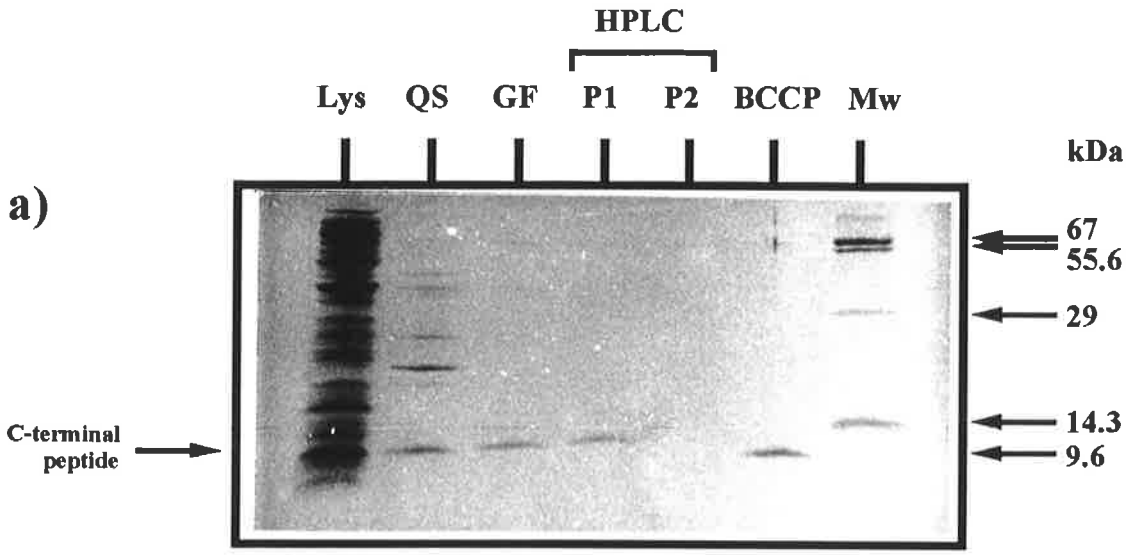


Figure 5.2 **Outline of the expression and *in vivo* biotinylation of the C-terminal Pyc2 biotin domain peptide**

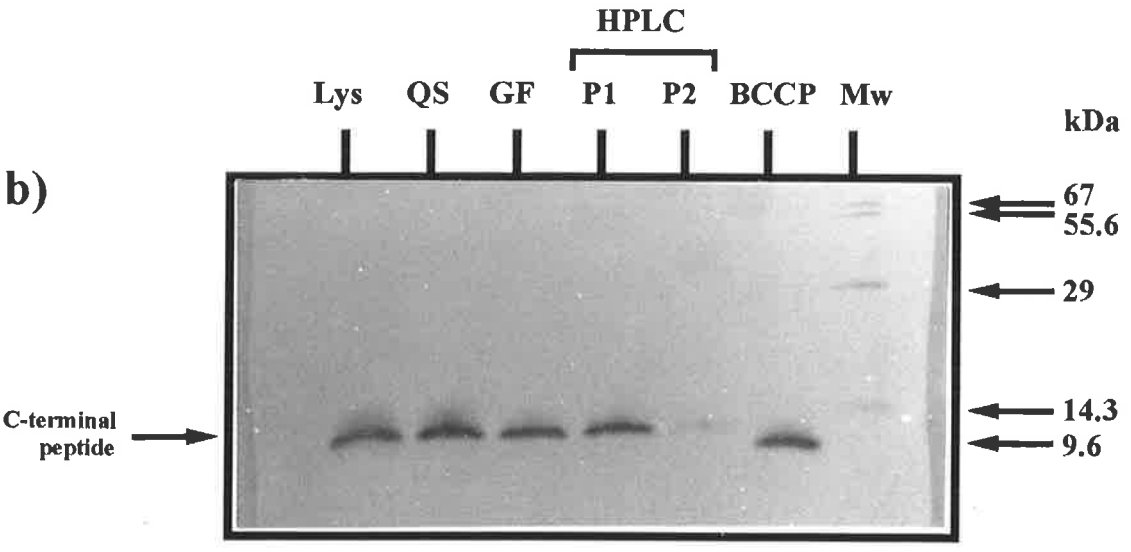
Figure 5.3. SDS-PAGE and Western analysis of the material obtained from each stage of the purification of the Pyc2 C-terminal peptide.

Duplicate protein samples were prepared from an aliquot taken at each stage of the purification of the C-terminal Pyc2 peptide. Samples were analysed by running duplicate 12% Tris-tricine SDS-PAGE gels. One gel **(a)** was Coomassie stained for total proteins, and the other gel **(b)** used for the detection of biotinylated proteins by Western blotting and avidin alkaline phosphatase detection as outlined in the Chapter 2. **Lanes**: **Lys**, induced lysate; **QS**, anion exchange step (Q-Sepharose); **GF**, gel filtration (Sephacryl-S200) step; **P1** & **P2**, HPLC peaks 1 & 2; **BCCP**, purified and biotinylated 87 amino acid C-terminal biotin domain peptide from the biotin carboxyl carrier protein of *E. coli* acetyl Co-A carboxylase (Chapman-Smith, *et al.*, 1994); **Mw**, biotinylated molecular weight markers of the sizes indicated.

Coomassie Stain



Western Blot



Being of a similar size and isoelectric point (cal. 4.67) as the BCCP peptide, it was expected that the Pyc2 biotin domain peptide would also be soluble in 60% isopropanol, especially as sequence comparisons suggested that all biotin domains probably have similar structures. Moreover, the yeast Pyc2 and *E. coli* BCCP biotin domains had been found to be sufficiently similar in structure for the *E coli* biotin ligase to biotinylate both of these molecules (see section 5.2.1). However, ^{3}H-biotin labelled Pyc2 biotin domain peptide was found to be insoluble in 60, 50 and even 40% isopropanol (data not shown).

Experiments were then set up to determine whether altering the pH, or increasing the NaCl concentration, would have any significant affect on the solubility of the Pyc2 peptide in isopropanol. Altering the pH from the usual value of 7.2 either up or down within the range of 4.6 to 9.0 had no significant affect on the solubility, and although increasing the salt concentration did result in a moderate increase in solubility, it was not of sufficient magnitude to enable this to be a useful purification step (data not shown).

Upon finding that the isopropanol precipitation step was ineffective, a more standard approach was tried, selecting chromatographic steps to purify this small acidic protein from first principles. The Pyc2 biotin domain peptide was subjected to anion exchange chromatography under acidic conditions (Bis-Tris pH 6.1), so as to limit the number other proteins which would be negatively charged and thus capable of binding to the anion exchange resin (Q-Sepharose), followed by gel filtration chromatography. The low molecular weight fractions were then be pooled and concentrated by a further anion exchange step, followed by reverse phase HPLC. The results presented in figures 5.3a & b revealed that this purification strategy worked remarkably well.

The Pyc2 peptide was found to elute from the Q-Sepharose anion exchange column at approximately 0.2 M NaCl (Fig 5.4). Analysis of the eluate by SDS-PAGE (Lane 2, Fig. 5.3a) and Western blotting (Lane 2, Fig. 5.3b) showed that the Q-Sepharose step had removed all the other small molecular weight proteins (below about 17 kDa), and the majority of the larger proteins as well. After the Sephacryl -S200 gel filtration step, the Pyc2 biotin domain peptide was found to be reasonably pure (Lane 3, Fig. 5.3a). Pure material was ultimately obtained as a single peak on C4 reverse phase HPLC (P1; Lane 4 of Fig. 5.3a & b) .

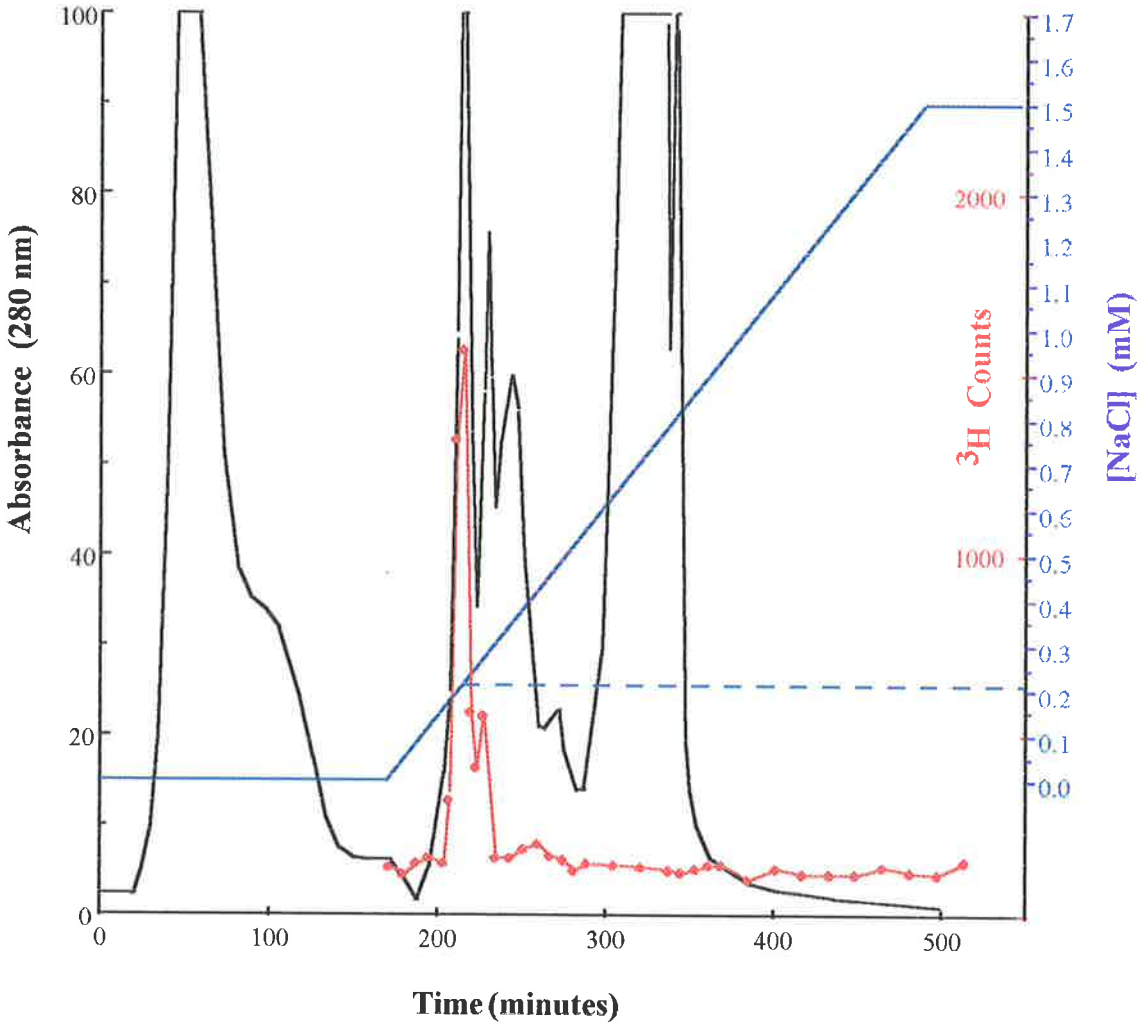


Figure 5.4 **Q-Sepharose anion exchange chromatography of the C-terminal Pyc2 biotin domain peptide**

After the overnight *in vitro* biotinylation step (see Chapter 2 sec. 2.2.2.6) the diluted crude lysate sample containing the Pyc2 peptide was loaded onto a Q-Sepharose column equilibrated in 20 mM Bis-Tris pH 6.1. The unbound proteins were removed by washing the column with 20 mM Bis-Tris pH 6.1 until the absorbance at 280 nm returned to background levels, and the C-terminal Pyc2 biotin domain peptide was eluted using a 0 to 1.5 M NaCl gradient while monitoring the absorbance at 280 nm. Ten mL fractions were collected and the quantity of Pyc2 peptide in each fraction was determined by ^{3}H scintillation counting in HiSafe OptiScint fluid. The dotted blue line shows the approximate NaCl concentration at which the Pyc2 C-terminal biotin domain peptide eluted from the column.

Absorbance (280 nm)

[NaCl] (mM)

^{3}H Counts

a) ADVHDTHQIG

b)

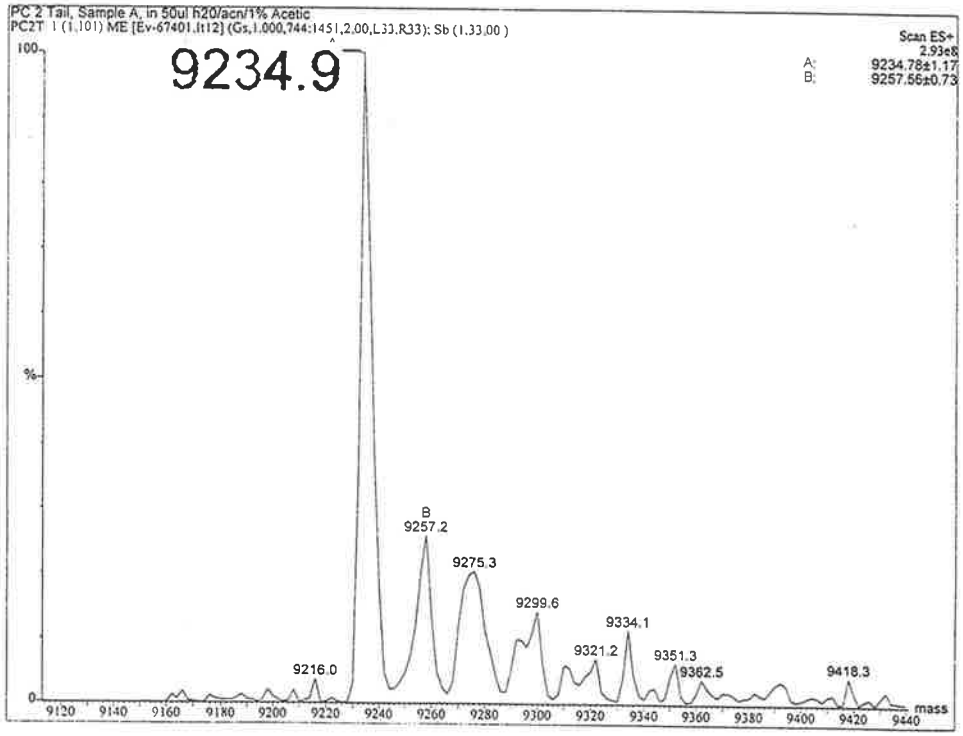


Figure 5.5 N-terminal sequencing and mass spectrometric analysis of the purified C-terminal Pyc2 peptide

a) N-terminal sequencing Sequence of the first 10 N-terminal residues obtained by sequencing 100 pmol of pure Pyc2 peptide material from the P1 HPLC peak using an Applied Biosystems 470A Sequencer.

b) Mass spectrometric analysis Summary in graph form of the data calculated from electrospray mass spectrometry of approximately 200 pmol of the purified Pyc2 C-terminal peptide using a VG Quattro mass spectrometer.

5.2.3 N-terminal sequencing and mass spectrometric analysis of the Pyc2 biotin domain peptide

Figure 5.5a shows the N-terminal sequencing results for the first 10 residues of the purified material obtained upon from expression of the Pyc2 biotin domain peptide encoded by the C-terminus of the *PYC2* gene reported in Chapter 4. Allowing for the expected removal of the N-terminal methionine by *E. col.* methionyl-aminopeptidase (Hirel *et al.*, 1989), these results confirmed that the N-terminus of the expressed Pyc2 biotin domain peptide corresponded with the predicted protein sequence. Furthermore, electrospray mass spectrometry of the pure material gave a calculated Mw of 9234.9 (Fig. 5.5b), which is within 0.01% of the theoretical value of 9235.7 for the peptide predicted from the sequence reported in Chapter 4 (c86K; see Fig 5.6a), and 620.6 atomic mass units below the theoretical value for the peptide which corresponds to the published sequence (c91R; see Fig. 5.6a). Taken together these data provide convincing evidence for the amino acid sequence predicted for the C-terminus of Pyc2, and hence the DNA sequence determined for this region of the *PYC2* gene I have isolated (Fig. 4.4).

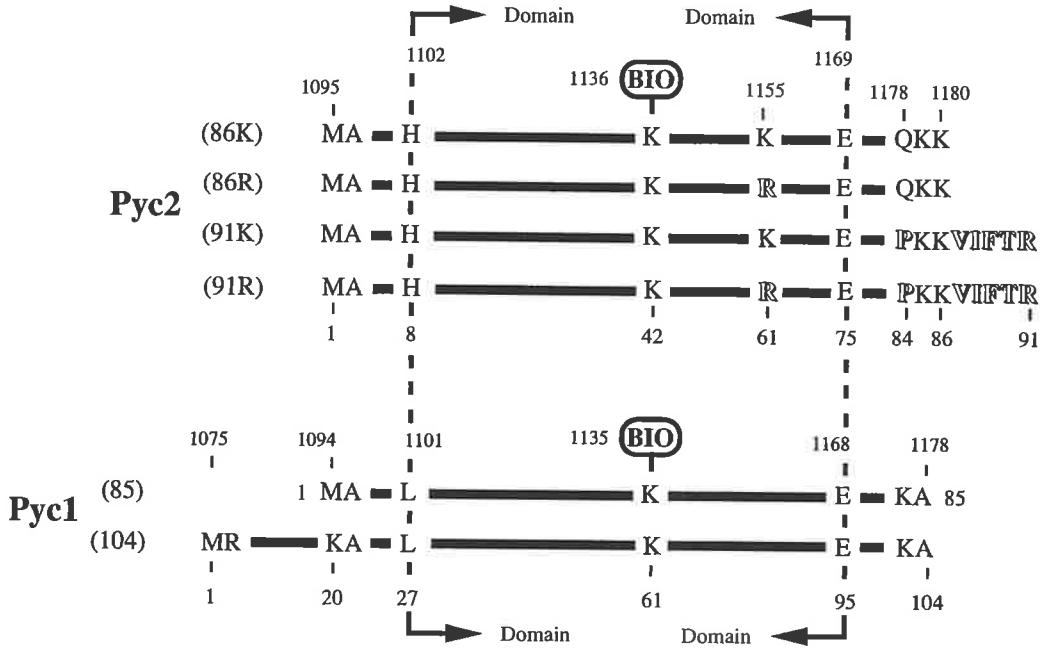
5.2.4 Comparative *in vivo* biotinylation of various Pyc biotin domain peptides

In order to determine whether the two sequence differences in the Pyc2 biotin domain affect the degree of *in vivo* biotinylation, constructs were engineered to express the 3 remaining peptides required so that a peptide with each of the possible combinations of these two differences had been produced. As a control for the subsequent comparative biotinylation experiments with these peptides, the analogous native Pyc1 biotin domain peptide starting from the corresponding N-terminal residue was also expressed (c85, Fig. 5.6a). In addition a longer native Pyc1 biotin domain peptide was also produced (c104, Fig. 5.6a). The reasons for the particular length of the longer peptide will not be discussed in detail here, as it is outside the scope of this chapter, being related to the work of Dr Chapman-Smith in investigating potential residues which may be involved in subunit-subunit interactions between Pyc1 monomers.

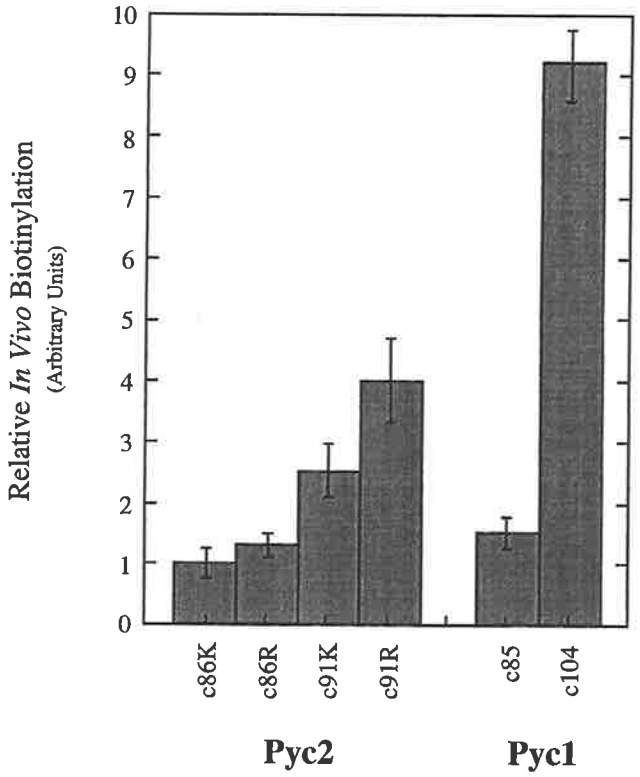
Figure 5.6 *In vivo* biotinylation of the Pyc2 and Pyc1 biotin domain peptides.

a) Schematic diagram of the Pyc biotin domain peptides. The numbers above each group of peptides indicate the amino acid positions in the native Pyc proteins, while the numbers below indicate the positions relative to the N-terminus of each expressed peptide. The dotted lines indicate the position of the residues in the Pyc sequences that correspond to first and last "structured" residues (shown in capital lettering) in the three-dimensional structure determined for the biotin domain of the BCCP subunit from *E. coli*. Acc (Brocklehurst *et al.*, 1995). The key internal residues distinguishing the various peptide constructs are shown in capital letters, while the shadowed capitals indicate the residues in Pyc2 predicted by the published sequence (Stucka *et al.*, 1991). The name used to refer to each peptide is given in brackets on the left hand side.

b) Comparative *in vivo* biotinylation of the Pyc peptides. Lysates were prepared from clones expressing each of the Pyc peptides at an appropriate level for measuring the degree of *in vivo* biotinylation (see section 5.2.4). Samples of each lysate containing approximately 10 μg of protein were run on 12% Tris-tricine gels in duplicate, with one gel Coomassie stained for total proteins and the other used for Western blotting and avidin alkaline phosphatase detection (see Chapter 2). The graph presents the average results (arbitrary units) obtained from triplicate gels and Westerns, by quantitating and comparing the amount of Coomassie stained protein with the amount of biotinylated material detected on the Western blots as outlined in Chapter 2.


Domain
Domain
1102
1169
1095
1136
BIO
1155
1178
1180
(86K) MA H K K E QKK
(86R) MA H K R E QKK
(91K) MA H K K E PKKVIIFTR
(91R) MA H K R E PKKVIIFTR
Pyc2
1 8 42 61 75 84 86 91
1075
1094
1101
1135
BIO
1168
1178
(85) 1 MA L K E KA 85
(104) MR KA L K E KA
Pyc1
1 20 27 61 95 104
Domain
Domain

a)


Relative *In Vivo* Biotinylation
(Arbitrary Units)
0
1
2
3
4
5
6
7
8
9
10
c86K
c86R
c91K
c91R
c85
c104
Pyc2
Pyc1

b)

To prepare the 3 additional Pyc2 constructs (c86R, c91K, & c91R) the mutations encoding the Q1178P / VIFTR extension, and the K1155R transversion were both introduced into the 3' *PYC2* clone (p58) by PCR with the oligonucleotides O6612 and O6613 respectively, as follows. A 460 bp 3' *PYC2* fragment containing the "O6613 mutation" was obtained by PCR using the external primers (O3671 and PCR*Bam*HI) with a mixture of the "wt" 349 bp and "mut" 209 bp overlapping fragments shown in figure 5.7. The 460 bp fragment thus obtained was then blunt ended with Klenow DNA polymerase I, digested with *Nco*I & *Bam*HI, and subcloned into pET-16b. Recombinant clones containing inserts were identified by the "Lid Lysate" technique (see Chapter 2), and the correct recombinants containing the desired "mutant" construct required to express the c86R peptide were identified by DNA sequencing. The construct required to express the peptide with the "O6612 mutation" (c91K) was obtained as outlined above using the O6612 "mut" 275 bp fragment (Fig. 5.7) in place of the "mut" 209 bp fragment, and the construct expressing the peptide with both mutations (c91R) was similarly obtained by PCR with the external primers from a mixture of the two 460 bp "single mut" fragments followed by subcloning into pET16b. Figure 5.8 shows the original wild type DNA sequence, and the resultant mutant DNA sequences (O6612 & O6613), confirming the introduced sequence changes at each of these sites.

Introduction of the restriction sites and initiating methionine (where required) for the subcloning and subsequent expression of the c85 and c104 Pyc1 peptides was performed by single stranded uDNA site directed mutagenesis method of Kunkel *et al.* (1987). To construct the plasmid for expressing the c85 peptide an ATG (K1094M) and *Nco*I site was introduced into *PYC1* at +3278 (Fig. 5.9a) using the oligo O5012. The 3' end of *PYC1* containing this "O5012 mutation" was then subcloned into pET16b as a 920 bp *Nco*I-*Bam*HI fragment. To express the c104 peptide a *Bsp*HI site was introduced at the native ATG located at +3221 in *PYC1* (Fig. 5.9b) using the oligo O5011. The 3' end of *PYC1* containing this "O5011 mutation" was then subcloned into pET16b as a 980 bp *Bsp*HI-*Bam*HI fragment. Both the mutations were confirmed by DNA sequencing (data not shown).

Induced and uninduced lysates were prepared from cells containing each of the Pyc2 and Pyc1 constructs, and duplicate samples of each lysate were run on separate SDS-PAGE

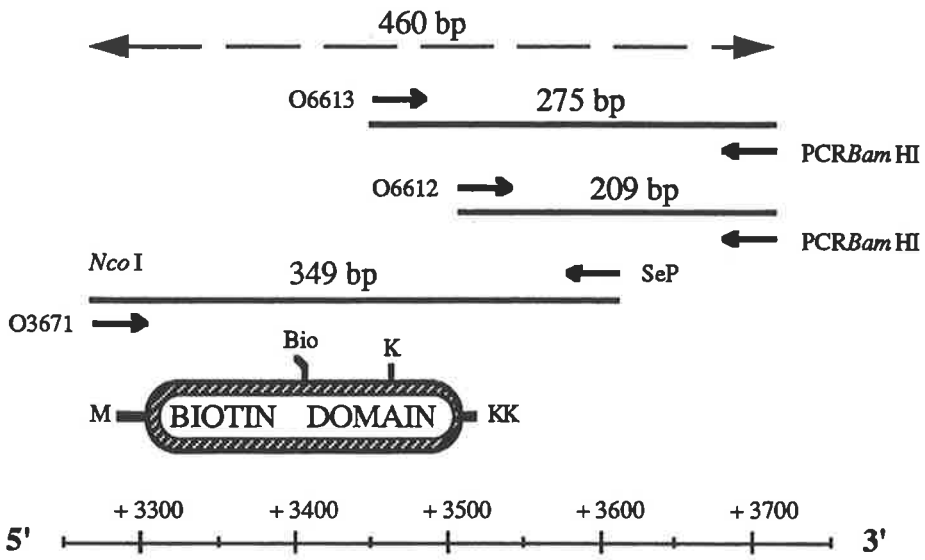


Figure 5.7 PCR mutagenesis strategy.

Short blunt arrows (not to scale) indicate the position of the mutagenic / PCR primers, while the dashed arrow indicates the size of the full length PCR product obtained using the external O3671 and PCR*Bam* HI mutagenic primers. The solid lines show the size and position of the PCR fragments that were combined in the PCR reactions used to prepare the various "mutant" full length products, and the scale below shows the nucleotide positions relative to the first ATG in *PYC2*. The central schematic diagram shows the position, length, and important features of the biotin domain expressed from the *PYC2* sequence reported in Chapter 4.

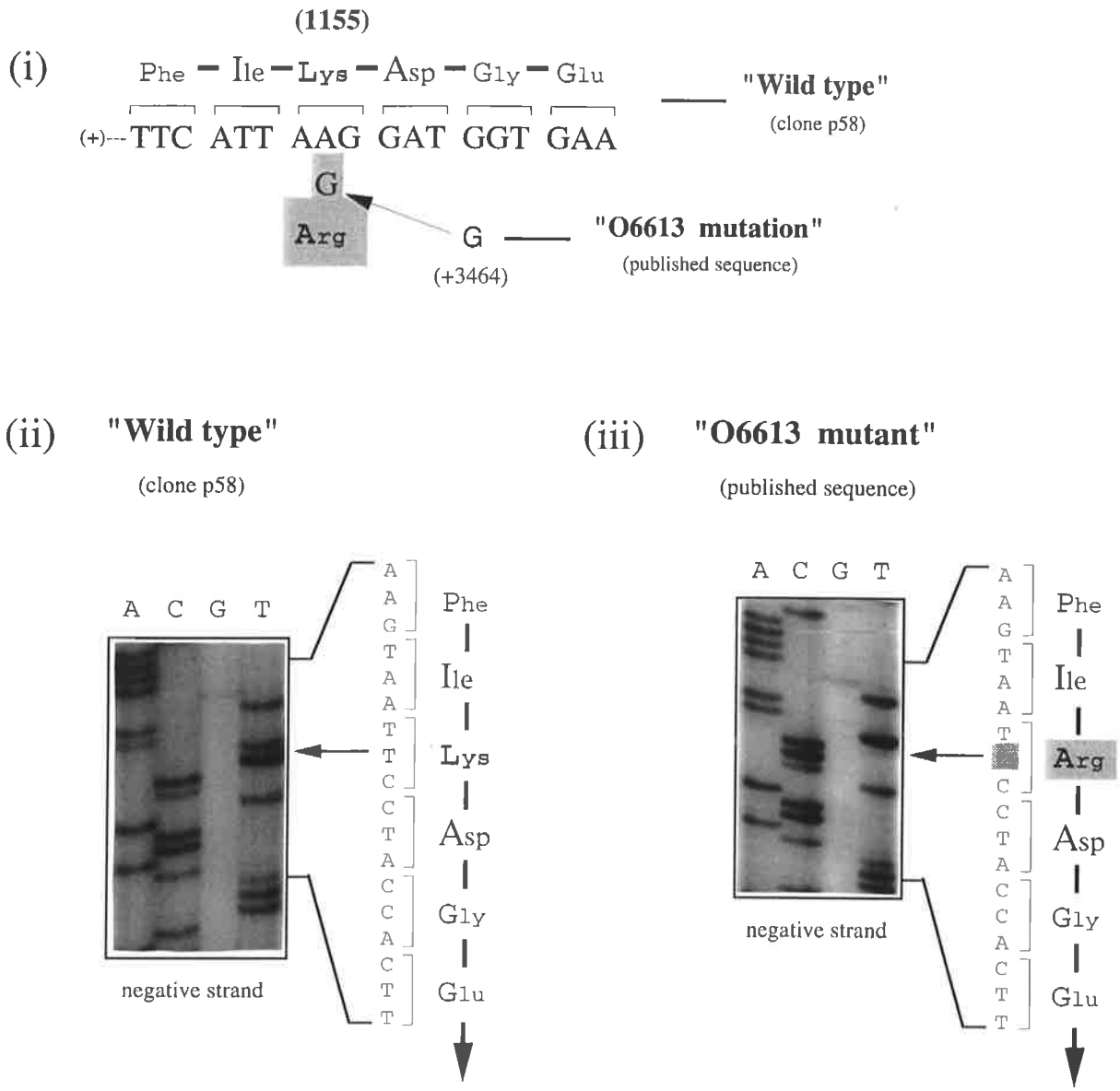


Figure 5.8 Sequence differences in the Pyc2 biotin domain.

a) A/G transversion. (i) Schematic diagram showing the differences in the DNA(+ strand) and amino acid sequences between clone p58 and the published sequence, at position +3464. (ii) DNA and amino acid sequences of the "wild type" *PYC2* gene (clone p58). (iii) DNA and amino acid sequences of the "O6613 mutant" (corresponding to the published data). Fine arrows indicate the site of the A/G transversion in the"O6613 mutant". Bold arrows indicate the direction of translation. Shading shows the introduced sequence changes corresponding to the published sequence.

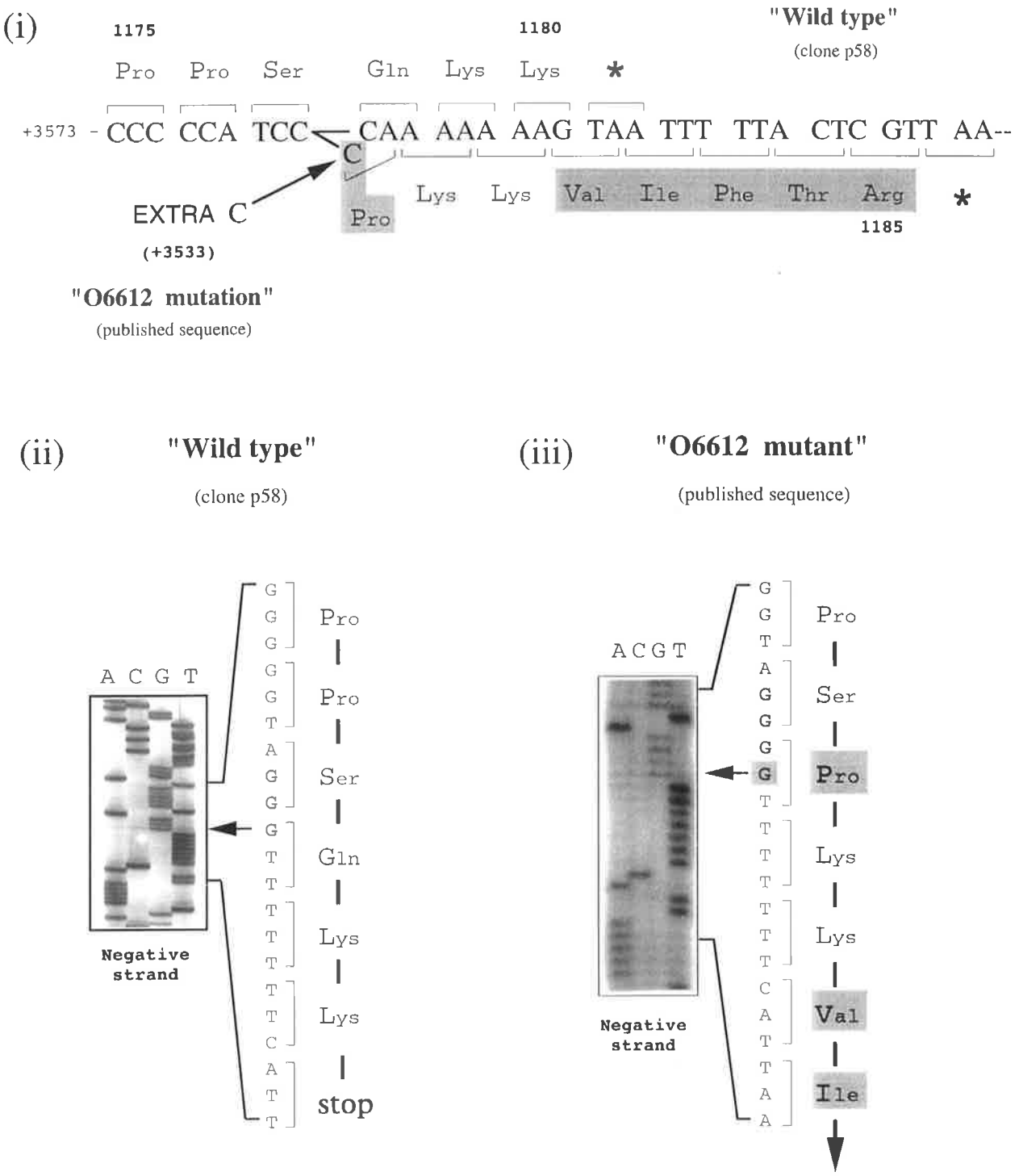


Figure 5.8 **Sequence differences in the Pyc2 biotin domain.**

b) G/C insertion. (i) Schematic diagram showing the differences in the DNA(+ strand) and amino acid sequences between clone p58 and the published sequence, at position +3533. (ii) DNA and amino acid sequences of the "wild type" *PYC2* gene (clone p58). (iii) DNA and amino acid sequences of the "O6612 mutant" (same as published data). Fine arrows indicate the site of the single base (G/C) insertion in the"O6612 mutant". Bold arrows indicate the direction of translation. Shading shows the introduced sequence changes corresponding to the published sequence.

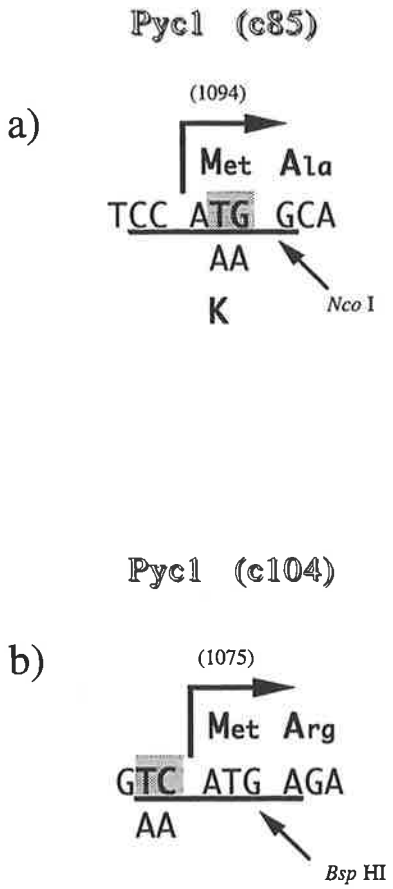


Figure 5.9 **Mutagenic changes introduced to express the 85 and 104 amino acid Pyc1 biotin domain peptides.**

Schematic diagram showing the nucleotide changes that were introduced in order to generate cloning sites and an initiating methionine for expression of the **a)** c85 peptide, and **b)** c104 peptide from the pET-16b system. The mutated DNA and amino acid sequences are shown above, with the original wild type sequence shown below. Shading indicates the residues changed by mutagenesis, underlining shows the introduced cloning site, and bold arrows indicate the direction of translation.

gels, one gel was Coomassie stained, and the other was used for Western blotting and avidin alkaline phosphatase detection of the biotinylated proteins. The relative extent of *in vivo* biotinylation of each peptide was then determined by quantitating and comparing the amount of Coomassie stained expressed protein with the quantity of biotinylated material detected by avidin alkaline phosphatase. To ensure that meaningful comparisons were made, the constructs being compared were all analysed on the same gels and Westerns, and the experiment was repeated 2-3 times.

As previously outlined (section 5.2.1), experiments with the BCCP and the Pyc2 c86K peptide suggested that with T7-expressed biotin domains some of the biotin domain material remains unbiotinylated unless the reaction is forced to completion *in vitro* by the addition of excess ATP. However, analysis of samples from a number of the Pyc constructs revealed that the extent of biotinylation for a given peptide remained relatively constant with increasing expression provided that the amount of expressed peptide was less than 20% of the total cellular protein (data not shown). This suggests that within this range the observed extent of biotinylation is a measure of the particular peptide's inherent ability to act as a substrate in the biotin ligase reaction. Consequently, calculations of the extent of *in vivo* biotinylation for each of the peptides were made using lysates from cultures in which the cells were expressing a quantity of peptide within this range.

The comparative extent of biotinylation calculated from a number of such gels and Western blots is represented graphically in figure 5.6b. This suggests that the various Pyc peptides differ quite considerably in their abilities to act as substrates for *in vivo* biotinylation by the *E. coli* biotin ligase. Comparing the extent of biotinylation of the four Pyc2 constructs, it appeared that the peptide containing the biotin domain sequence from the gene reported here (c86K) was the most poorly biotinylated. Both of the sequence differences in the biotin domain corresponding to the DNA sequence reported by Stucka *et al.* (1991) improved the *in vivo* biotinylation of the Pyc2 biotin domain. The K1155R change in c86R produced a slight increase in biotinylation on its own, while the Q1178P / VIFTR extension in c91K produced a 2 fold increase in biotinylation. When both of these changes were combined in the one peptide (c91R; the peptide corresponding to the DNA sequence of Stucka *et al.*, 1991), the resultant peptide had a 4 fold increase in biotinylation.

Comparison of the extent of biotinylation of the c85 and c104 Pyc1 peptides suggests that the 19 amino acid N-terminal extension present in c104 produced a 6 fold increase in biotinylation.

5.3 DISCUSSION

5.3.1 Purification properties of the Pyc2 biotin domain peptide

The most likely explanation for the observed differences in the isopropanol solubility between the Pyc2 and BCCP biotin domains is that these two proteins have a considerable number of different surface, solvent exposed residues, on an otherwise very similar three-dimensional structure. It would seem less likely that it is due to any major differences in their tertiary structures, as all biotin domains are believed to share a common structure. For example, the recently determined structure for the BCCP biotin domain (Brocklehurst *et al.*, 1995) was found to be very similar to the predicted structure of Pyc1, obtained on the basis of it sharing certain key hydrophobic residues known to be important in the solved structure of lipoyl domains (Brocklehurst and Perham, 1993). Indeed figure 4.10 shows that all biotin domains, including Pyc2, share this same pattern of key hydrophobic residues, further implying that they all share the same basic structure. This is presumably why biotin ligase enzymes can recognise biotin domains from widely different species (see section 5.1).

Assuming that the key residues involved in the precise recognition of biotin domains by ligase enzymes are on the surface of the molecule, the isopropanol solubility result suggests that it should be possible to eliminate a considerable number of the surface residues as candidates in this recognition. Only those surface residues which are shared between the Pyc2 and BCCP biotin domain structures could be involved in recognition. For this and other reasons it would be expected that a detailed comparison of the 3 dimensional structures of Pyc2 and BCCP could be very useful in furthering our understanding of the structure-function relationships in biotin domains. Consequently Dr Chapman-Smith is presently continuing the work on the purification of the Pyc biotin domains in our laboratory with the aim of obtaining enough material to enable structure determination by NMR.

5.3.2 Analysis of the C-terminus of Pyc2

The most significant difference between the inferred Pyc2 protein sequence reported by Stucka *et al.* (1991) and the sequence presented in Chapter 4 is the Q1178P / 5 amino acid (VIFTR) extension at the C-terminus of the published sequence. This difference was sufficiently large to enable mass stectrometry to be used to determine whether the C-terminus of the Pyc2 protein expressed from the gene I isolated actually lacks these amino acids, as is predicted on the basis of the DNA sequence (see Fig. 4.4 & 4.6). By obtaining an accurate molecular weight of a recombinant C-terminal peptide it was possible to confirm that the C-terminus of the Pyc2 protein expressed from the gene isolated in this report did in fact correspond to the predicted amino acid sequence presented in Chapter 4, rather than the published sequence.

5.3.3 Comparative *in vivo* biotinylation

As it appeared quite possible that some or all of the sequence differences in the published Pyc2 sequence may be due to real polymorphisms (see Chapter 4) it was of interest to determine what affects if any the Q1178P / 5 amino acid (VIFTR) extension, and the K1155R "change" would have on *in vivo* biotinylation. Although these sequence differences were not investigated at the level of the whole Pyc2 protein, the effects on *in vivo* biotinylation observed with isolated biotin domains can be expected to reflect those that would be observed with the whole protein, as biotin domains have been shown to be stable independently folding domains (Fall and Vagelos, 1973; Cronan, 1990; Reed and Cronan, 1991).

Ideally, to get a true measure of the relative biotinylation of the Pyc2 peptides *in vivo*, they should have been expressed in the presence of excess biotin and the yeast biotin ligase, rather than the *E. coli* biotin ligase. However, the ability of ligase enzymes to biotinylate the domains from a wide range of even very different species is well known (McAllister and Coon, 1966; Cronan, 1990). Also, previous attempts to express Pyc1 biotin domain peptides in yeast have resulted in low yields, and problems due to proteolysis (presumably down to the core domain) and aggregation (Dr. A Chapman-Smith, unpublished results). Consequently, the peptides were expressed in *E. coli* to avoid the

proteolysis problems, and in the presence of excess *E. coli* biotin ligase as the yeast gene although cloned is still being characterised (Dr. J. E. Cronan, pers comm.). By expressing the peptides under these conditions it was possible to make the following observations.

Firstly, the results showed that biotin domain peptides from the yeast Pyc1 and Pyc2 proteins are able to be recognised and biotinylated *in vivo* by the *E. coli* biotin ligase, providing further evidence that biotin domains can be biotinylated efficiently by ligase enzymes from different species. Secondly, the biotin domains expressed from the two "polymorphic" forms of Pyc2 differed in their ability to be biotinylated *in vivo* (see Fig. 5.6b). Hence the Pyc2 enzymes containing these domains would also be expected to show differences in *in vivo* biotinylation and thus enzymatic activity, unless the yeast biotin ligase behaves quite differently.

More specifically, the 5 amino acid extension at the C- terminus of Pyc2 produced a noticeable increase in biotinylation. This was unexpected as the C-terminus of Pyc2 expressed from the gene we isolated already extends some 10 residues past the aligned C-terminus of the BCCP biotinyl domain (Brocklehurst *et al.*, 1995; see Fig. 5.6a & Fig. 4.10). In the case of Pyc1, the 19 amino acid N-terminal extension also produced an increase in biotinylation of this domain. This was a similarly unexpected result, since the extension was to an N-terminus already 8 residues upstream of the expected beginning of the biotin domain structure (see Fig. 5.6a). However, the N- and C-termini of the Pyc1 (and presumably Pyc2) biotin domain are both adjacent in the predicted 3D structure (Brocklehurst and Perham, 1993), as is the case with the structure of the BCCP biotin domain (Brocklehurst *et al.*, 1995). This juxtaposition in the 3D structure may explain the similar effects on biotinylation observed with both the N- and C-terminal extensions.

Although all known biotin enzymes have methionine residues flanking the biotinylated lysine, these methionine residues are not required for biotinylation. Instead, mutagenesis of these residues affects the ability of the biotin moiety to participate in the carboxyl transfer reactions (Kondo *et al.*, 1984; Shenoy *et al.*, 1992). Similarly, changing either of the flanking residues in the analogous DKA sequence present in all lipoyl enzymes was shown to affect the reductive acylation of the lipoyl moiety, rather than lipoylation itself (Wallis and Perham, 1994). Changing the DKA in the Pdh lipoyl domain to MKM did not

result in biotinylation instead of lipoylation, but displacing the K from its position at the tip of the β-hairpin loop by one residue in either direction destroyed lipoylation (Wallis and Perham, 1994). These results suggest that, in addition to the exposed position of the recipient lysine at the tip of the β-hairpin loop (see Fig. 5.10), other features of the domains are recognised by the respective ligase enzymes .

The fact that N- and C-terminal extensions to the Pyc biotin domains both improve biotinylation suggests this region of the molecule may in some way be involved with recognition by biotin ligase. This proposal is in agreement with the effects observed upon modifying the C-terminus of the transcarboxylase (Tc) enzyme from *P. shermanii*. Murtif and Samols (1987) found that truncation of the C-terminus of Tc prior to I122, or changing this penultimate residue from a hydrophobic residue (to G, but not to V or L; see Fig 5.10), destroys the biotinylation. In fact this penultimate hydrophobic residue has since been identified to be one of the key folding residues involved in forming the hydrophobic core of lipoyl and biotinyl domains (Brocklehurst and Perham, 1993; Brocklehurst *et al.*, 1995). In addition, a G70N mutation in the bovine H protein located at the first residue of the no. 6 β-strand (Fig. 5.10) adjacent to the C-terminus of the protein (Pares *et al.*, 1994), is known to destroy the lipoylation of this domain (Fujiwara *et al.*, 1991). Taken together these results all suggest that the "tail ends" of both biotin and lipoyl domains are playing some role in the correct recognition and post-translational modification by their respective ligases.

With respect to the extensions to the N and C-termini of the core biotin domains from Pyc1 and Pyc2, as they are located at the opposite end of the molecule to the recipient lysine residue, the improved ability of the *E. coli* biotin ligase enzyme to biotinylate these domains is probably due to the extensions producing a stabilisation of the protein folding, rather than any direct interactions with the ligase enzyme.

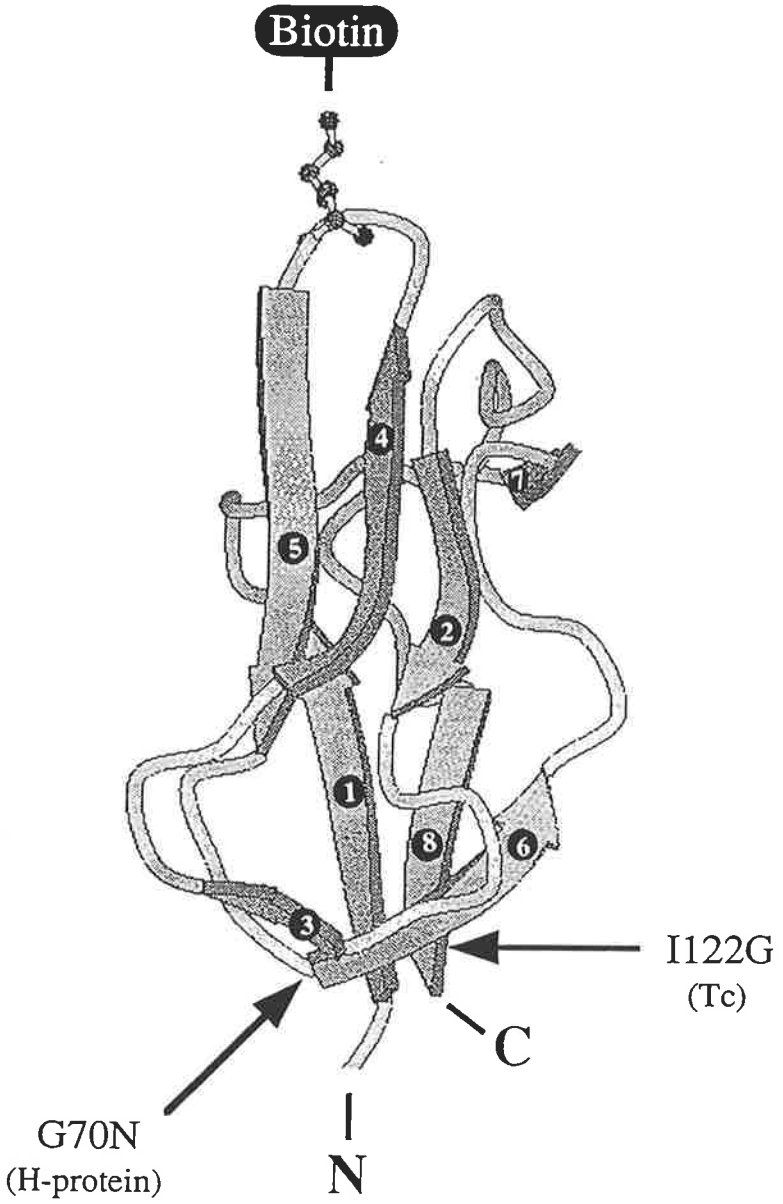


Figure 5.10 Schematic drawing of the three-dimensional structure of the E. coli BCCP biotin domain, adapted from Brocklehurst *et al.* (1995)

The eight β-sheet strands making up the bulk of secondary structure of the biotin domain are numbered from the N-terminus, and the N and C-termini of the structured biotin domain are as shown. The carbon chain of the biotinylated lysine residue is shown in "ball and stick" format, and the arrows indicate the sites corresponding to the mutations found to destroy the biotinylation of the 1.3S transcarboxylase (Tc) subunit from *P. shermanii* (Murtif and Samols (1987), and lipoylation of the bovine H-protein (Fujiwara *et al.*, 1991).

CHAPTER 6

CONSTRUCTION OF *PYC2* NULL MUTANTS

6.1 INTRODUCTION

Probably the most powerful and widely used modern technique employed in the study of protein structure / function is site-directed mutagenesis. This technique enables the researcher to introduce specific changes to one or more amino acids at specific sites within a protein sequence. The altered gene encoding this mutant protein can then be expressed, and the properties of the mutant protein can be used to deduce the role of the altered amino acid(s) in the particular protein function being studied.

Clearly, regardless of the type of protein expression system being used, or the particular assay employed, a correct interpretation of the assay results for mutant proteins generated by site-directed mutagenesis requires that their properties can be unambiguously distinguished from any contaminating background activities. When mutant proteins are being assayed *in vivo*, this implies that the expression hosts should ideally be devoid of any of the endogenous native protein (or its homologue if the host is a different species), and other related proteins which could interfere with the assay. Hence, the generation of a suitable strain of the host organism chosen for the expression of the mutant protein(s) is frequently a prerequisite in structure / function studies. The primary aim of the work described in this chapter was to generate a strain of *Saccharomyces cerevisiae* which was devoid of any pyruvate carboxylase activity to use as a host for the expression of various yeast Pyc constructs.

The technique which is probably the method of choice for generating specific null mutants is gene disruption by homologous recombination. This technique has been widely used with *Saccharomyces cerevisiae*, and the method chosen for the homologous recombination work described in this chapter was the one-step gene disruption method of Rothstein (1983). Basically this involves the transformation of a haploid yeast strain with a linear DNA construct containing the gene of interest, a central portion of which has been replaced by a selectable marker gene. Upon entering the cells, a small percentage of the gene disruption constructs recombine with the endogenous gene via their homologous ends, thus rendering the gene non-functional by replacing the central native gene fragment with a non-homologous selectable marker gene.

In the case of pyruvate carboxylase from *Saccharomyces cerevisiae*, generating a null mutant strain which is devoid of all Pyc activity requires the disruption of both *PYC* genes. Prior to the work described in this chapter, Dr M. E. Walker had constructed a strain of yeast in which the *PYC1* gene had been disrupted with the *LEU2* marker gene (MW21.3) (Walker *et al.*, 1991). Indeed, it was as a result of studying the properties of strain MW21.3 that the *PYC2* gene was discovered (see Chapter 1, 1.8). Hence, the construction of a double null mutant strain of *Saccharomyces cerevisiae* could be achieved by disrupting the *PYC2* gene in MW21.3.

In addition to their use as hosts for the expression of various mutant protein constructs, null mutant strains themselves are often very useful in providing information on protein function. For example, in the case of organisms which contain multiple isoforms of a particular protein, much can be learnt about the structure / function relationships of the protein, and the individual metabolic roles of the different isoforms, by comparing the properties of the wild type strain with the multiple and single null mutant strains.

This chapter describes the:

(i) preparation of a *PYC2* gene disruption construct using the *TRP1* marker gene

(ii) isolation and characterisation of *pyc2*::*TRP1* null mutant strains

(iii) preparation of a *PYC2* gene disruption construct with the *HIS3* marker gene

(iv) isolation and characterisation of *pyc2*::*HIS3* null mutant strains

6.2 RESULTS AND DISCUSSION

6.2.1 Isolation of *pyc2* null mutants

6.2.1.1 Preparation of a *PYC2* gene disruption construct using the *TRP1* marker gene

The wild type laboratory strains of *Saccharomyces cerevisiae* that had been used by Dr M. E. Walker to produce the *PYC1* null mutant strain MW21.3 (MAT α), and its

opposite mating type MW31 (MAT a), were DBY746 and DBY747 respectively. These parental strains are identical, apart from their mating type, having the genotype (*his3*, *trp1*, *ura3*, *leu2*, *PYC1*, *PYC2*). As the *LEU2* gene had been integrated into the genome of these two parental strains in the construction of the above mentioned *pyc1* null mutant strains, this left the *HIS3*, *TRP1*, and *URA3* genes as potential auxotrophic marker genes in the preparation of gene disruption constructs. However, the *URA3* gene was deemed unsuitable as it is commonly used as a selection marker on yeast plasmids, leaving either the *HIS3* or *TRP1* genes as useful marker genes. Of these, the *TRP1* gene was chosen due to its ready availability in our laboratory.

As the *TRP1* gene was contained on a 800 bp *Eco*RI-*Bgl*II fragment (from the YRp7 plasmid, Struhl *et al.*, 1979; Stinchcomb *et al.*, 1979), the strategy employed to prepare the *pyc2*::*TRP1* gene disruption construct (T3) essentially involved cloning this fragment into the 5' end of *PYC2* in place of the 777 bp 5' *Bgl*II-*Eco*RI, and 362 bp *Eco*RI-*Eco*RI fragments. Basically this was achieved by digesting the pBDV2 construct with *Eco*RI and *Bgl*II, and re-ligating the resultant construct with the original 576 bp *PYC2* *Eco*RI-*Bgl*II fragment, and the 800 bp *Eco*RI-*Bgl*II *TRP1* gene fragment, as outlined in figure 6.1a.

The correct ligation of these fragments was then confirmed by restriction mapping and Southern hybridisation to the *TRP1* gene (data not shown).

6.2.1.2 Isolation of *pyc2*::*TRP1* null mutants

The approach taken to isolate *pyc2*::*TRP1* null mutants involved the following steps:

1) *Bam*HI digestion of the T3 recombination construct so as to produce a linear DNA fragment containing the *TRP1* gene flanked on each end by a few hundred base pairs of *PYC2* sequence (see Fig. 6.1a),

2) transformation of wild type (DBY746 and DBY747) and *pyc1*::*LEU2* strains (MW21.3 and MW31) with the digested T3 DNA by the Li^+ method (see Chapter 2),

3) selection of TRP^+ clones on glucose minimal media plates containing 8 mM L-aspartate,

4) preparation of chromosomal DNA from TRP^+ clones,

5) screening for TRP+ clones which contain the T3 construct correctly integrated into the genome (*pyc2::TRP1* clones) by Southern hybridisation of *Eco*RI digested chromosomal DNA, using the 362 bp *Eco*RI-*Eco*RI *PYC2* fragment as a probe. The correct *pyc2::TRP1* clones could then be obtained as TRP+ colonies which lack the 362 bp *Eco*RI-*Eco*RI *PYC2* fragment (see Fig. 6.1a).

In the first two transformation experiments each parental strain was transformed with approximately 5 µg of *Bam*HI digested T3 DNA, and 1 - 20 TRP+ yeast colonies appeared on the +DNA transformation plates from each strain. However, there were an equal number of TRP + colonies growing on the " minus DNA" transformation control plates. This result suggested that the transformation had not been very successful. Hence the putative "TRP+ " colonies on the +DNA plates were not subjected to the remaining steps in the screening procedure. Furthermore, the numbers of TRP+ colonies resulting from "- DNA" control transformations suggested that for some reason these yeast strains contained a high background of cells which display a TRP+ phenotype despite the absence of a functional *TRP1* gene.

It was expected that increasing the quantity of digested DNA in the transformations should increase the ratio of correct T3 integrants to background colonies displaying the TRP+ phenotype. Hence, in the subsequent transformation experiments the above mentioned wild type and *pyc1::LEU2* strains were each transformed (as previously described) with 20 - 30 µg of *Bam*HI digested T3 DNA. The transformation results from this experiment are presented in TABLE 6.1 below.

Strains	DBY746	DBY747	MW21	MW31
- DNA	3	5	0	8
+DNA	25	23	2	14

TABLE 6.1 **Transformation of yeast with the T3 construct.**
The numbers represent the amount of TRP+ colonies obtained upon transformation by the Li+ method (see Chapter 2).

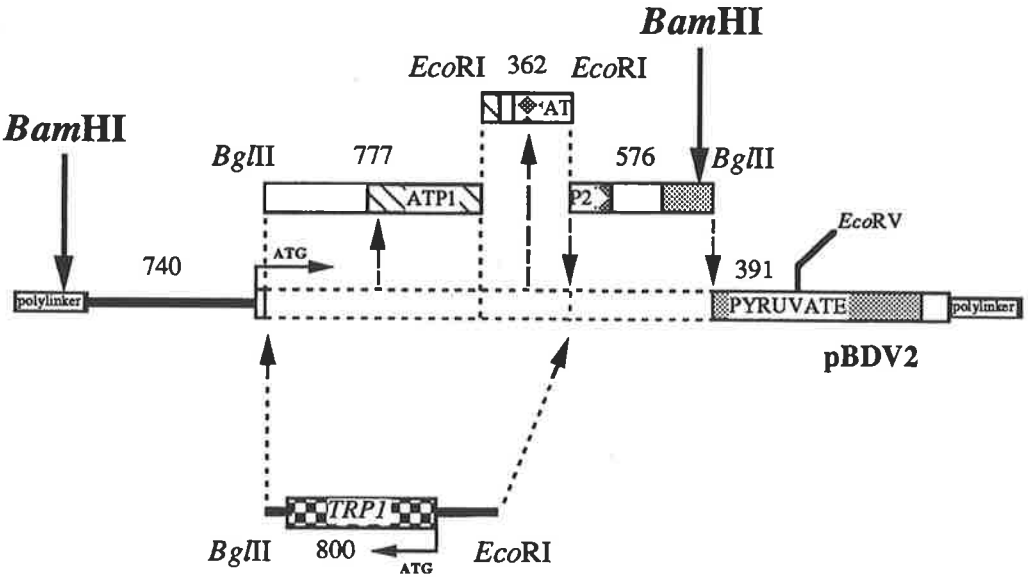
Figure 6.1 Schematic diagram of the strategy used to prepare the *PYC2* gene disruption constructs

a) Preparation of the T3 construct The T3 *PYC2* gene disruption construct was prepared by replacing the 777 bp *Bgl*II-*Eco*RI and 362 bp *Eco*RI-*Eco*RI *PYC2* fragments from pBDV2 with an 800 bp fragment containing the *TRP1* gene. The linear *PYC2-TRP1* recombinant fragment used to transform the yeast strains in gene disruption experiments was then obtained by digesting the T3 construct with *Bam*HI.

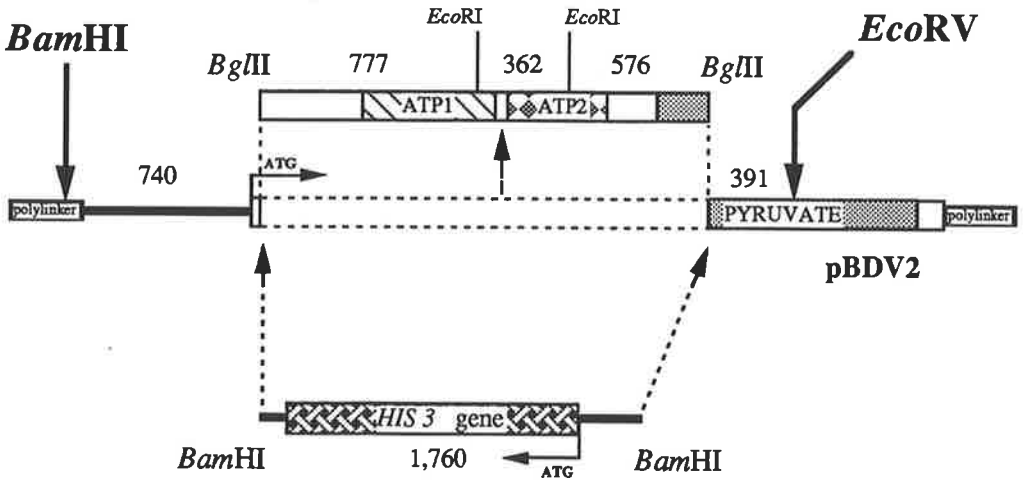
b) Preparation of the H11 construct The H11 *PYC2* gene disruption construct was prepared by replacing the *PYC2 Bgl*II fragment from pBDV2 with a 1.76 kbp *Bam*HI fragment containing the *HIS3* gene. The linear *PYC2-HIS3* recombinant fragment used to transform the yeast strains in the gene disruption experiments was then obtained by digesting the H11 construct with *Bam*HI and *Eco*RI.

The dashed arrows show the removal and insertion of the fragments performed during the preparation of each of the constructs, the bold plain arrows show the positions of the flanking restriction sites used to linearise the construct prior to use in gene disruption experiments, and the numbers represent fragment sizes in base pairs.

a)


BamHI
EcoRI 362 EcoRI
BamHI
BglII 777 576 BglII
ATP1
P2
EcoRV
740
391
ATG
polylinker
PYRUVATE
polylinker
pBDV2
TRP1
BglII 800 EcoRI
ATG

b)


BamHI
EcoRI EcoRI
EcoRV
BglII 777 362 576 BglII
ATP1
ATP2
740
391
ATG
polylinker
PYRUVATE
polylinker
pBDV2
HIS 3 gene
BamHI 1,760 BamHI
ATG

Although the total numbers of TRP+ colonies obtained were once again relatively low, the numbers of colonies resulting from the +DNA transformations were in each case higher than the "- DNA" controls. Hence, 12 +DNA TRP+ colonies from each of the wild type (DBY) strains, and all the TRP+ colonies from both the *pyc1*::*LEU2* strains were picked and screened using the approach described above.

Figure 6.2 shows the Southern hybridisation results after probing the transferred, *Eco*RI digested chromosomal DNA from each of these colonies, with the 362 bp *Eco*RI-*Eco*RI fragment of *PYC2*. These results revealed that none of the TRP+ colonies from either of the two *pyc1*::*LEU2* strains (MW21.3 and MW31) were correct T3 integrants, as they all contained the 362 bp *Eco*RI-*Eco*RI fragment of *PYC2*. However, clones 2, 6, 8, 9, and 12 from DBY746, and clones 7 and 11 from DBY747 all lacked the *Eco*RI-*Eco*RI fragment of *PYC2*, indicating that all these TRP+ colonies were in fact *pyc2* null mutants containing the T3 construct integrated at the *PYC2* locus, ie having the genotype (*his3*, *ura3*, *leu2*, *PYC1*, *pyc2*::*TRP1*). Two of the *pyc2*::*TRP1* clones from the parental strain DBY746 (2 and 6), and two from the parental strain DBY747 (7 and 11) were chosen for further characterisation. In future these clones are referred to as DV6.2, DV6.6, DV7.7 and DV7.11 respectively. Furthermore, the construction of strain DV6.2 has previously been reported in the literature (Brewster *et al.* 1994).

As none of the TRP+ clones from the parental strains MW21.3 or MW31 were correct T3 integrants, it was clear that further experiments would be required in order to obtain a *pyc1*/*pyc2* double null mutant strain. Several unsuccessful homologous recombination experiments were performed using the approach outlined above. In addition, the following modifications were also tried, without success:

1) increasing the quantity and / or purity of the digested T3 DNA, (gel filtration through Superose 6 column) (see Chapter 2)
2) transformation of the cells by electroporation (Becker and Guarente, 1991) rather than the Li+ method, (Ito *et al.*, 1983) and
3) including 10 nM D-biotin and Triton X100 in addition to the 8 mM L-aspartate as extra nutritional supplements in the agar plates.

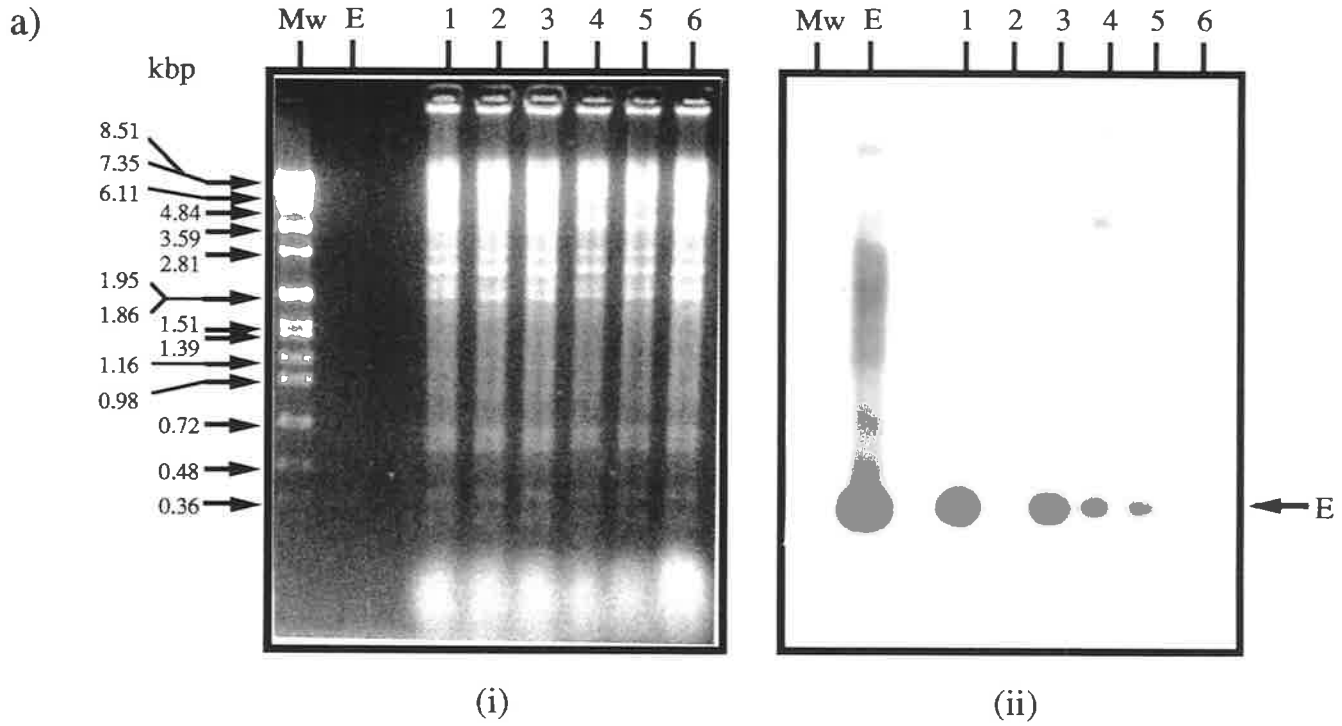


Figure 6.2 a) Screening the TRP$^+$ clones from strain DBY746 for *pyc2* null mutants.

Approximately 5µg of *Eco*RI digested chromosomal DNA from 6 of the TRP$^+$ clones obtained by transformation of DBY746 with the *Bam*HI digested T3 gene disruption construct were analysed by (i) 0.8% agarose gel electrophoresis, followed by (ii) Southern transfer to nitrocellulose and probing with the 362 bp *PYC2 Eco*RI-*Eco*RI fragment (E). Lane Mw refers to molecular weight markers, and lane E contained ~10ng of the 362 bp *PYC2 Eco*RI-*Eco*RI fragment. The arrow indicates the position of the 362 bp *PYC2* (E) fragment in the digests from the clones containing the wild-type *PYC2* gene.

(b)

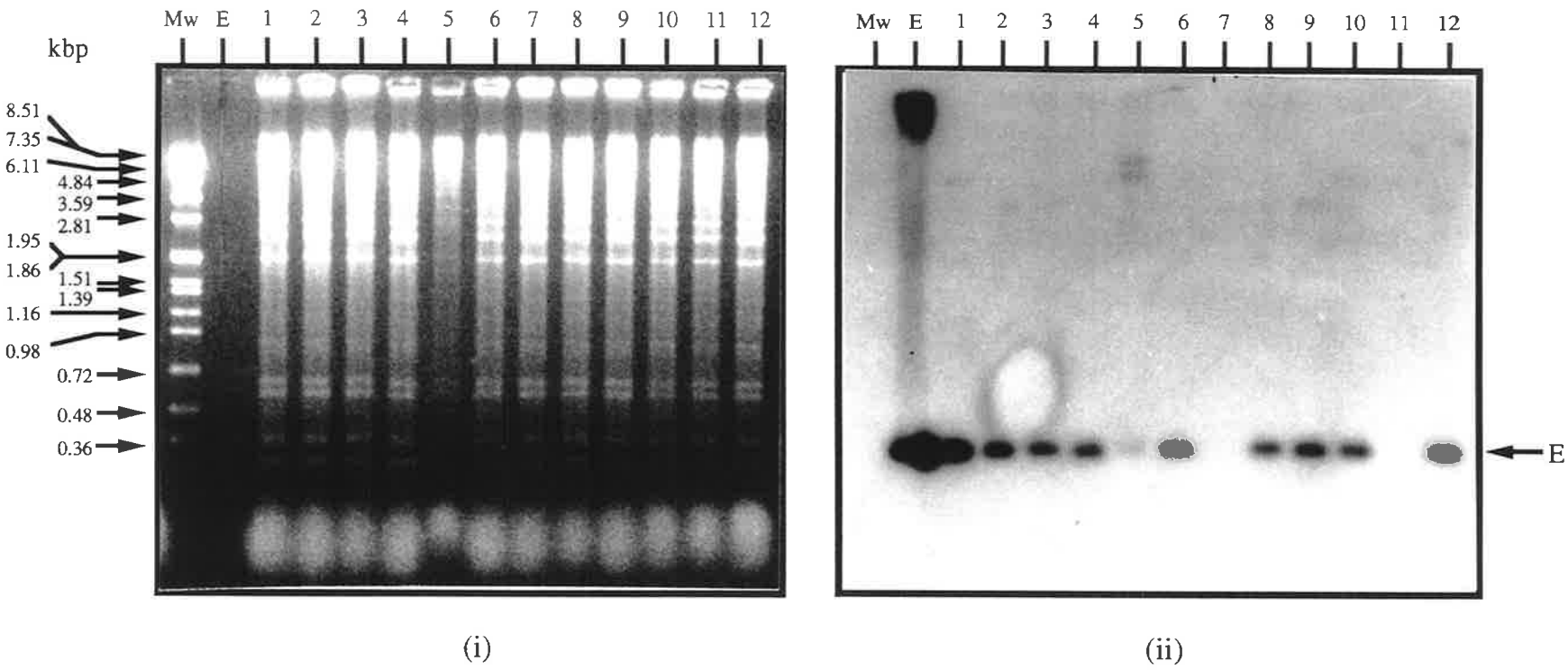


Figure 6.2 b) Screening the TRP$^+$ clones from strain DBY747 for *pyc2* null mutants.

Approximately 5µg of *Eco*RI digested chromosomal DNA from 12 of the TRP$^+$ clones obtained by transformation of DBY747 with the *Bam*HI digested T3 gene disruption construct were analysed by (i) 0.8% agarose gel electrophoresis, followed by (ii) Southern transfer to nitrocellulose and probing with the 362 bp *PYC2 Eco*RI-*Eco*RI fragment (E). Lane Mw refers to molecular weight markers, and lane E contained ~10ng of the 362 bp *PYC2 Eco*RI-*Eco*RI fragment. The arrow indicates the position of the 362 bp *PYC2* (E) fragment in the digests from the clones containing the wild-type *PYC2* gene.

In each case a high background of TRP1+ colonies were obtained on the "- DNA" control plates, and it was this interference that was believed to be the probable cause of the lack of detection of any *pyc* double mutants, rather than any inherent flaws in the technique itself. This high background of non-integrant colonies exhibiting the TRP+ phenotype was presumed to be a problem that would be peculiar to the selection of this particular phenotype. Furthermore, the *pyc* double mutants were not expected to be lethal, as the inclusion of 8 mM aspartate in the growth medium was expected to adequately compensate for the lack of oxaloacetate which is normally supplied to the TCA cycle in cells containing Pyc. For these reasons it was decided to persist with attempts to isolate *pyc* double null mutants, but instead to try the *HIS3* gene as the selection marker in the gene knockout experiments.

6.2.1.3 Preparation of a *PYC2* gene disruption construct using the *HIS3* marker gene

Figure 6.1b shows a schematic representation of the strategy used to prepare the *pyc2*::*HIS3* gene disruption construct (H11). This involved digesting the 5' *PYC2* clone (pBDV2) with *Bgl*II, and cloning the 1.76 kb *Bam*HI-*Bam*HI fragment containing the *HIS3* gene from the YEp6 vector (Struhl *et al.*, 1979) between the two extreme *Bgl*II sites.

As was the case with the *TRP1* construct, the correct ligation of these fragments was then confirmed by restriction mapping and Southern hybridisation to the *HIS3* gene (data not shown).

6.2.1.4 Isolation of *pyc2*::*HIS3* null mutants

Having prepared the H11 construct, the approach taken to isolate the *pyc2*::*HIS3* null mutants using this construct was essentially identical to the approach which had been used to isolate the *pyc2*::*TRP1* null mutants using the T3 construct, apart from the initial restriction enzyme digestion step (step 1). This needed to be different as the 576 bp *Eco*RI-*Bgl*II fragment of *PYC2* containing the 3' *Bam*HI site had been replaced in this *pyc2*::*HIS3* construct (H11). Hence, linearisation of the H11 construct so as to produce a fragment containing the *HIS3* gene flanked at each end by *PYC2* sequences was instead achieved by

digestion with *Bam*HI and *Eco*RV (see Fig. 6.1b). Also, the transformed yeast cells were of course plated onto minimal plates to select for clones with a HIS^+ rather than TRP^+ phenotype.

The transformation results from electroporation of strains DBY746, DBY747, and MW21.3, with a range of quantities of *Bam*HI-*Eco*RV digested, ethanol precipitated H11 DNA are presented in TABLE 6.2 below.

DNA (µg)	DBY746	DBY747	MW21.3
0	0	0	0
~ 3	29	42	~ 750
~ 15	~ 400	34	~ 2,500
~ 60	--	--	~ 6,000

TABLE 6.2 Transformation of yeast with the H11 construct.
The numbers represent the amount of HIS^+ colonies obtained upon transformation with the approximate (~) quantities of digested H11 DNA as shown, using the electroporation method (see Chapter 2).

The results presented in TABLE 6.2 clearly show that the transformations had worked well. Furthermore, the problem of high numbers of background colonies present on the -DNA control plates that had been encountered with the *TRP1* marker gene, had not occurred when selecting for this *HIS3* gene.

On the basis of these promising transformation results, 9 clones derived from strain MW21.3, and 6 clones from each of the two wild type (DBY) strains were picked and subjected to the remaining steps (4 and 5) in the isolation / screening approach outlined in section 6.2.1.2 above. The Southern hybridisation experiments revealed that clones 2, 4, 5, and 6 from strain DBY746, and clones 3 and 5 from DBY747 all lacked the *Eco*RI-*Eco*RI fragment of *PYC2* (Fig. 6.3). These results indicated that all these HIS^+ colonies were in fact *pyc2* null mutants containing the H11 construct integrated at the *PYC2* locus, ie having the genotype (*ura3*, *trp1*, *leu2*, *PYC1*, *pyc2*::*HIS3*). Two of these *pyc2*::*HIS3* clones from the parental strain DBY746 (2 and 4), and the two from the parental strain DBY747 (3 and 5)

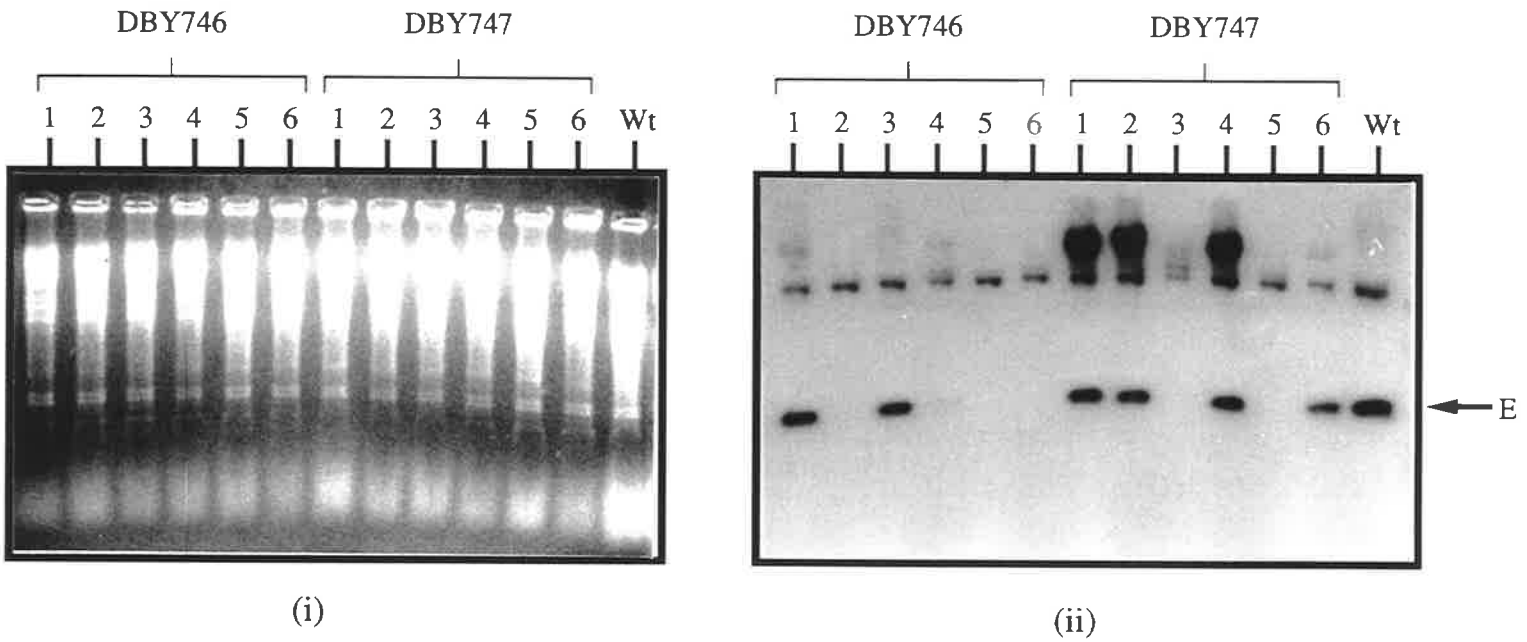


Figure 6.3 Screening the HIS3$^+$ clones from strains DBY746 and DBY747 for *pyc2* null mutants.

Approximately 5µg of *Eco*RI digested chromosomal DNA from 6 of the HIS3$^+$ clones obtained by transformation of DBY746 or DBY747 with the *Bam*HI-*Eco*RV digested H11 gene disruption construct were analysed by (i) 0.8% agarose gel electrophoresis, followed by (ii) Southern transfer to nitrocellulose and probing with the 362 bp *PYC2* *Eco*RI-*Eco*RI fragment (E). The arrow indicates the position of the 362 bp *PYC2* (E) fragment in the digests from the clones with the wild-type *PYC2* gene.

were then chosen for further characterisation. In future these clones are referred to as DV6H.2, DV6H.4, DV7H.3 and DV7H.5 respectively.

Although the H11 construct had successfully disrupted the *PYC2* gene in the wild type strains, the results with MW21.3 were less promising. All of the 9 chosen HIS$^+$ clones from MW21.3 were found to contain the *Eco*RI-*Eco*RI fragment of *PYC2* (data not shown), indicating that they did not contain the H11 construct correctly integrated at the *PYC2* locus. However, as there were large numbers of HIS$^+$ colonies on the plates from MW21.3, it still seemed quite possible that there may be some correct clones on these plates.

Upon close inspection of the MW21.3 plates after the additional growth which had occurred in the interim period during the screening procedure, it was noticed that amongst the numerous colonies on the plates there were a small percentage of slower growing colonies. These smaller colonies seemed like good candidates for *pyc1*/*pyc2* double mutants. Furthermore, it seemed unlikely that any of these smaller colonies would have been large enough to have been picked in the previous screening experiment. Therefore, 100 of these smaller colonies were picked, restreaked onto fresh glucose minimal media plates, and incubated (30 ^{o}C) for 4 days. The 48 slowest growing colonies were then re-picked and grown for a further 4 days in 3 mL of YPD liquid media in the presence of 8 mM L-aspartate. These 48 clones were then screened for the presence of the correctly integrated H11 construct as previously described, except that the chromosomal DNA from these clones was not digested with *Eco*RI prior to electrophoresis and Southern transfer. This step was omitted, as earlier experiments had revealed that hybridisation occurred equally well with the undigested yeast chromosomal DNA, and thus the *Eco*RI digestion step was in fact unnecessary (data not shown).

The Southern hybridisation experiments revealed that clones 14, 18, 22, 24, 26, 39, and 43 all lacked the *Eco*RI-*Eco*RI fragment of *PYC2*, and thus contained the correctly integrated H11 construct (Fig. 6.4). Thus, all these HIS$^+$ clones were in fact *pyc2* null mutants, and having been derived from MW21.3 they were also *pyc1* null mutants. Therefore these clones would each have the genotype (*ura3*, *trp1*, *pyc1*::*LEU2*, *pyc2*::*HIS3*), and as they would contain no functional Pyc proteins they should be devoid of any Pyc enzymatic activity. Two of these double mutant clones were then selected for further

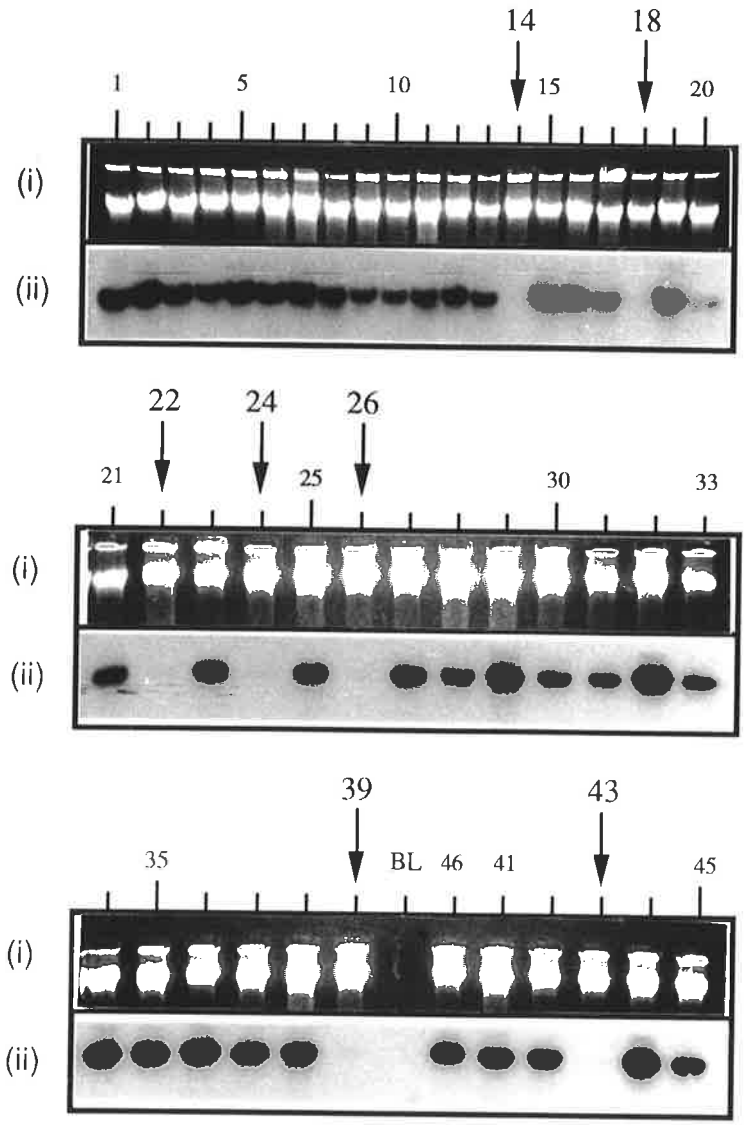


Figure 6.4 **Screening the slow growing HIS3$^+$ clones from strain MW21.3 for *pyc2* null mutants.**

Approximately 5µg of undigested chromosomal DNA from 46 of the slowest growing HIS3$^+$ clones obtained by transformation of MW21.3 with the *Bam*HI-*Eco*RV digested H11 gene disruption construct were analysed by (i) 0.8% agarose gel electrophoresis, followed by (ii) Southern transfer to nitrocellulose and probing with the 362 bp *PYC2* *Eco*RI-*Eco*RI fragment. The arrows indicate the "putative" *pyc2* null mutant clones, and BL indicates the position of the blank lane.

characterisation (clones 18 and 22), and in future these clones are referred to as DM18 and DM22 respectively. The construction of the double mutant strain DM18 has also been reported previously in the literature (Brewster *et al.* 1994).

6.2.2 Confirmation of gene knockout in the *PYC2* "null mutants"

The Southern hybridisation evidence showing the loss of a 362 bp *Eco*RI-*Eco*RI fragment within the coding sequence of *PYC2*, and the phenotypic data which implied that the marker gene had been integrated into the genome, together form convincing evidence of *PYC2* gene knockout in the isolated putative "null mutant strains". However, final confirmation of the gene knockout event requires the demonstration that *PYC2* gene no longer produces a functional enzyme product. In the case of pyruvate carboxylase there are at least three lines of evidence which can be used to demonstrate the absence of a functional gene product.

Firstly, as the gene disruption should have occurred 5' to the biotin domain (see Fig. 6.1), the absence of a ~130 kDa biotinylated protein band corresponding to Pyc2 should be evident on Western blots by avidin alkaline phosphatase detection. Secondly, the absence of a functional Pyc2 protein should be detectable by enzyme assays, and thirdly, this should also presumably be detected by an altered growth phenotype.

6.2.2.1 Western analysis of the *PYC2* "null mutants"

Each of the "null mutant strains" which were selected for further characterisation (see section 6.2.1) were grown up to mid-log phase in YPD rich media, and cell lysates were prepared. Duplicate lysate samples from each strain were run on two identical 10% Tris-glycine SDS-PAGE gels, with one gel being Coomassie stained for total protein, and the other gel used for Western blotting. The SDS-PAGE and Western blotting results for the "null mutant" strains derived from strain DBY746 are shown in figure 6.5a, while the corresponding results for the strains derived from DBY747 are shown in figure 6.5b.

As expected the Western blotting technique revealed the presence of a ~130 kDa biotinylated protein corresponding to pyruvate carboxylase in both the control *pyc1 / PYC2* null mutant strains (MW21.3 and MW31), ie the Pyc2 protein, and in all the putative *pyc2 /*

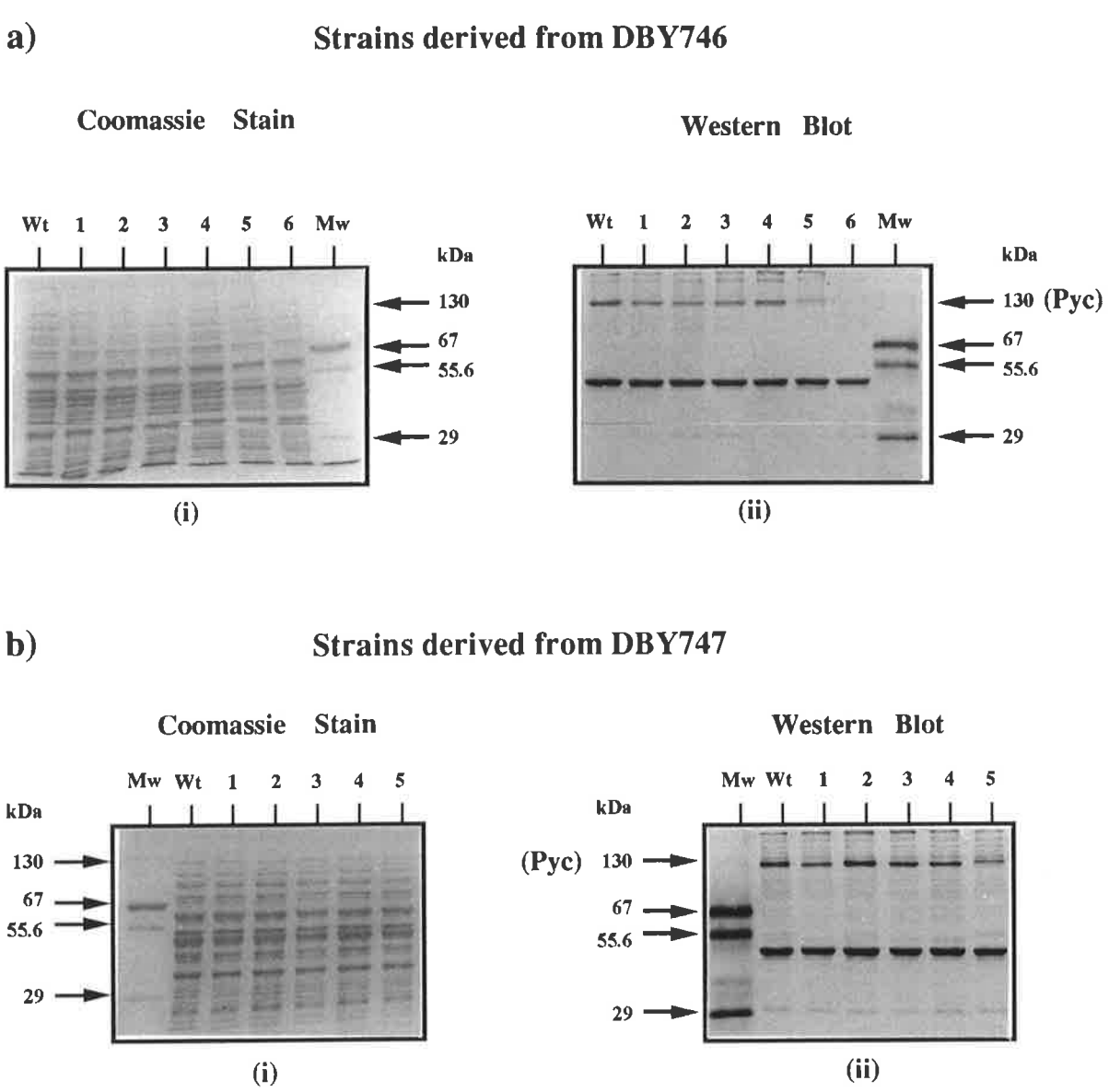


Figure 6.5 **Western analysis of the *pyc2* "null mutant" strains.**

Duplicate protein samples (~10 µg) were prepared from crude cell lysates obtained from each of the parental DBY strains and their derivative *pyc1* null and "putative" *pyc2* null mutant strains, grown to mid-log phase in YPD rich media. Samples were analysed by running duplicate 10% Tris-glycine SDS-PAGE gels, and all the samples from the strains derived from each parental DBY strain were run on the one gel, (a) **DBY746 derived strains,** (b) **DBY747 derived strains.** One gel **(i)** was Coomassie stained for total proteins, and the other gel **(ii)** was used for the detection of biotinylated proteins by Western blotting with avidin alkaline phosphatase (as outlined in Chapter 2). DBY746 Lanes: **Wt**, DBY746; **Lane 1**, DV6.2; **Lane 2**, DV6.6; **Lane 3**, DV6H.2; **Lane 4**, DV6H.4; **Lane 5**, MW21.3; **Lane 6**, DM18. DBY747 Lanes: **Wt**, DBY747; **Lane 1**, DV7.7; **Lane 2**, DV7.11; **Lane 3**, DV7H.3; **Lane 4**, DV7H.5; **Lane 5**, MW31. **Mw** refers to biotinylated molecular weight markers of the sizes indicated.

PYC1 single "null mutant strains", ie the Pyc1 protein. In the case of the *pyc1* null mutant controls, the amount of ~130 kDa biotinylated protein detected was markedly less than the total Pyc present in the parental wild type strains, indicating that Pyc1 probably accounts for a significant proportion of the total Pyc in the yeast cell . This was most obvious in the case of strain MW21.3. On the other hand, there was no obvious decrease in the amount of ~130 kDa biotinylated protein detected in the *PYC1 / pyc2* null mutant strains when compared with their parental wild type strains.

While this result means that the Western blotting technique failed to give any real confirmation of gene knockout in the putative *pyc2* single null mutants, it is not necessarily inconsistent with *pyc2* gene knockout having occurred. Single null mutants containing only a functional *PYC1* gene may still contain essentially the same levels of Pyc protein if the Pyc2 protein normally makes up only a minor proportion of the total amount of Pyc protein in the cell. Alternatively, the yeast cells may compensate for the loss of the Pyc2 protein by up regulating the expression of the Pyc1. However, this second option can be ruled out on the basis of the gene regulation studies conducted out in our laboratory by N. K. Brewster using strains DBY747, MW21.3, and DV6.2, (Brewster *et al.*, 1994). RNase transcription assays and Northern blots revealed that *PYC1* and *PYC2* are differentially regulated, such that during fermentative growth the lack of either of the *PYC* genes had no effect on the level and pattern of expression of the other *PYC* gene. In addition, the RNase transcription assays on samples from DV6.2 confirmed that *pyc2* gene knockout had been knocked out in this strain.

In the case of the putative double null mutant strain DM18, figure 6.5a clearly shows that the cell lysate from this strain did not contain any biotinylated protein material in the molecular weight range which would correspond to a pyruvate carboxylase enzyme (~130 kDa). This is consistent with the absence of a functional *PYC* gene is this strain (ie, it is a double null mutant), thus indicating that the *PYC2* gene had been successfully knocked out.

6.2.2.2 Enzyme activity of the *PYC2* "null mutants"

In order to further investigate the characteristics of each of the above mentioned putative *pyc2* "null mutants" strains, pyruvate carboxylase assays were performed using separate aliquots of the same mid-log lysates that were used for the Coomassie stained and

Western transferred SDS-PAGE gels presented in figures 6.5a and b. The results for these *pyc2* "null mutant" strains, and the various parental strain controls are shown in TABLE 6.3 below.

	Strains	Enzyme Activity (mU/mg)		Strains	Enzyme Activity (mU/mg)
Parental	DBY746	4.62 ± 0.20	Parental	DBY747	4.88 ± 0.15
pyc1 null	MW21.3	1.15 ± 0.11	*pyc1* null	MW31	2.62 ± 0.20
pyc2 nulls	DV6.2	1.80 ± 0.11	*pyc2* nulls	DV7.7	1.93 ± 0.20
	DV6.6	2.17 ± 0.11		DV7.11	4.20 ± 0.41
	DV6H.2	2.29 ± 0.33		DV7H.3	2.78 ± 0.07
	DV6H.4	1.63 ± 0.01		DV7H.5	2.87 ± 0.17
pyc1/pyc2 double null	DM18	0.00			

TABLE 6.3 Pyruvate carboxylase activity in the *pyc2* null mutants.
Values given are the mean of the avidin-sensitive mU / mg of activity (+/- the standard error of the mean) calculated from triplicate assays. Protein concentrations of the lysates were determined by the method of Bradford (1976), and the pyruvate carboxylase activity was measured using the $^{14}CO_2$ incorporation assay (Lim *et al.*, 1987). One U of Pyc activity is defined as the µmol / min of $^{14}CO_2$ converted into oxaloacetate.

From the pyruvate carboxylase assay results presented in TABLE 6.3 it can be seen that the level of enzyme activity in all the putative *PYC1 / pyc2* single "null mutant" strains was in each case less than the parental wild type strains. This is consistent with the *PYC2* gene having been knocked out in these strains. The only strain for which this conclusion would seem less definite is DV7.11, which had an enzyme activity only marginally less than the parental wild type strain DBY747.

In the case of the *pyc2* single null strains derived from DBY746, the enzyme activity of these strains were noticeably higher than MW21.3, consistent with the higher levels of Pyc protein detected on the Western blots by avidin alkaline phosphatase (see Fig. 6.5a). Hence these results infer that Pyc1 is the predominant Pyc isozyme. However, the differences in activity between the *pyc1* and *pyc2* null mutant strains derived from DBY747 were less noticeable. Three of the 4 *pyc2* null mutant strains did have higher levels of

enzyme activity than the *pyc1* null mutant (MW31), but apart from DV7.11 (which had an activity similar to the parental strain), the differences were only minimal. This was primarily due to the higher level of enzyme activity in MW31, when compared to MW21.3, the isogenic strain of the opposite mating type. This difference in activity between MW21.3 and MW31 is in agreement with observations made by other members of our laboratory (Dr. M. E. Walker, pers. comm.), but as yet the reasons for this difference have not been determined.

The pyruvate carboxylase assay results for strain DM18 clearly indicate that this strain is devoid of pyruvate carboxylase activity, in agreement with the Western transfer results which failed to detect any biotinylated Pyc protein in this strain (see Fig. 6.5a). Hence, along with the earlier Southern hybridisation results (see section 6.2.1.4), these results form convincing evidence that DM18 is indeed a *pyc1/pyc2* double mutant strain devoid of Pyc protein and enzyme activity. Therefore strain DM18 should be a suitable strain for use as a host in which to express various mutant yeast Pyc proteins generated by site-directed mutagenesis.

6.2.2.3 Growth phenotypes of the *PYC2* "null mutants"

In view of the important gluconeogenic and anaplerotic roles of pyruvate carboxylase (see Chapter 1, sec. 1.2), it might be expected that strains with null mutations in either or both of the *PYC* genes may display differences in growth phenotype. Indeed, during the course of the isolation of DM18 (and the other "double mutant" strains implicated by Southern hybridisation) it was found that this strain displays a total dependence on aspartate for growth, even in YPD rich media. This dependence on aspartate can be explained by the fact that, upon deamination within the cell, aspartate from the growth media is able to provide oxaloacetate in place of that which would normally be supplied by the pyruvate carboxylase reaction. Hence, this result implies that under normal growth conditions the carboxylation of pyruvate by Pyc is an essential reaction for carbon metabolism in yeast.

MW21.3, the *pyc1* null parental strain of DM18, similarly displayed dependence on aspartate in order to achieve normal rates of growth on glucose minimal media, but not on YPD rich media. On the other hand all of the *pyc2* null strains discussed in the preceeding

sections displayed no dependence on aspartate for growth in YPD or glucose minimal media (data not shown).

As N. K. Brewster in our laboratory was investigating the regulation of expression and the metabolic roles of each of the Pyc isozymes as a part of his PhD project, a detailed examination of the growth phenotypes of the various *pyc2* null mutants discussed in this chapter was considered to be outside the scope of my research project. These investigations have instead been carried out by N. K. Brewster (Brewster *et al.*, 1994) and are briefly summarised below.

TABLE 6.4 below shows the results of an examination of the growth rate dependence on aspartate concentration for strains DBY746, MW21.3, DV6.2, and DM18 in glucose minimal media.

	L-Aspartate (mM)			
Strains	0	0.5	2.5	12.5
DBY746	3.69	3.44	3.18	3.57
MW21.3	11.60	6.93	4.98	4.66
DV6.2	4.38	4.00	3.89	3.84
DM18	DNG	DNG	10.81	6.52

TABLE 6.4 Effect of *pyc* null mutations on aspartate growth dependence.

Adapted from Brewster *et al.*, (1994). Values refer to the mean generation times in hours. All 4 strains were grown in YPD media to late log phase, harvested, washed, then used to inoculate glucose minimal media containing the appropriate supplements to an initial density of 0.175 OD_{600} units. DNG, did not grow.

These results show that disruption of the *PYC1* gene had the most significant effect on aspartate dependence during fermentative growth on glucose. Disruption of the *PYC2* gene on the other hand had very little effect on the ability grow on glucose. Hence Pyc1 appears to be play the most important role in supplying the TCA cycle with C4 intermediates during fermentative growth on glucose. Similarly, N. K. Brewster found that fermentative growth

on ethanol is also primarily dependent on Pyc1 activity, as the MW21.3 and DM18 have an absolute requirement of aspartate for fermentative growth on ethanol, while DV6.2 shows no aspartate dependence (Brewster *et al.*, 1991).

These observed differences between the Pyc1 and Pyc2 isozymes contrast with the results of Stucka *et al.* (1991). During the course of the studies presented in this chapter these workers reported the partial characterisation of *pyc1* and *pyc2* single and double null strains that they had constructed from a similar laboratory wild type strain to DBY746 (W303-1A). They found little difference between the activities of the *pyc1* and *pyc2* single null mutant strains, reporting that both strains had similar generation times, and were not dependent on aspartate for growth in YPD or glucose minimal media. Only their double null mutant strain was dependent of aspartate for growth. Furthermore, their preliminary kinetic characterisations of the partially purified Pyc proteins from their two single null strains suggested that Pyc1 and Pyc2 have essentially the same K_m values for pyruvate and ATP.

Clearly, before the reasons for the reported differences between the corresponding *pyc* null mutant strains from our group and those of Stucka *et al.* (1991) can be understood with any certainty, a detailed comparison between these strains will need to be conducted by the one researcher. However, in the absence of such a comparison, some confirmation of the aspartate dependence of MW21.3 is provided by the similar phenotype observed with strain MW17, a *pyc* mutant strain obtained from the wild type strain S288C by ethane methyl sulphonate mutagenesis (Walker and Wallace, 1991).

CHAPTER 7

MUTAGENESIS OF THE CYSTEINE RESIDUES IN THE PYRUVATE DOMAIN

7.1 INTRODUCTION

Since the initial report by Keech and Utter (1963) showing that various mercury salts were powerful inhibitors of chicken liver Pyc (CLPyc), there have been numerous studies implicating free sulfhydryl (-SH) groups as being essential for the activity of pyruvate carboxylase. In 1967 Cazzulo and Stoppani reported that purified yeast Pyc (possibly a mixture of Pyc1 and Pyc2) was inhibited by the -SH reagents *p*-chloromercuribenzoate (pCMB), *o*-iodosobenzoic acid (IOB), and N-ethylmaleimide (NEM).

Palacian and Neet (1970, 1972) subsequently carried out further investigations with the CLPyc enzyme. They found that a range of -SH reagents had profound effects on both the activity and the quaternary structure of this enzyme. Of these, cystine and NEM were found to inactivate CLPyc in a bi-phasic manner. The first phase involved the modification of 4 to 32 of the estimated 55 free -SH groups in the enzyme (Scrutton and Utter, 1965), depending upon the reagent used and buffer anions present. This was accompanied by a partial loss in enzyme activity in the case of NEM modification, or no loss of activity in the case of cystine. Also, neither reagent produced any gross changes in the quaternary structure during this phase, ie it remained as a tetramer.

The second phase involved a modification of 16 - 39 additional -SH groups, resulting in complete inactivation, and dissociation to monomers with cystine, or the formation of high molecular weight aggregates in the case of NEM modification. However, on the basis of their experiments it was unclear as to whether the inactivation occurred prior to, or as a result of the dissociation / aggregation events.

To determine whether the inactivation was associated with the modification of any active site residues, Palacian and Neet (1972) went on to study the effects of the substrates of the Pyc reaction (ATP, pyruvate, oxaloacetate), and acetyl-CoA, on the rates of inactivation by cystine and NEM. They found that these molecules had a diverse range of effects on the rate of inactivation depending on the -SH reagent used, buffer anion present, and the phase of inactivation. As a result of this diversity of effects, they concluded that it was unlikely that the -SH group(s) responsible for the inactivation were located in the active site.

Upon purifying and characterising the sheep kidney enzyme (SKPyc) in our laboratory, Bais (1974) found that the -SH reagent 5,5'-dithiobis(2-nitrobenzoate) (DTNB) reacted with approximately 2 -SH groups / monomer of the native enzyme. This modification of only ~2 cysteine residues of SKPyc with DTNB, and ~1 cysteine of SLPyc with NEM (Nielsen, 1970) led to the sheep enzyme being chosen instead of the chicken enzyme for the subsequent chemical modification studies carried out in our laboratory.

Using the sheep liver enzyme, Hudson *et al.* (1975) investigated the possibility that the -SH modifying reagent bromopyruvate (BP) may act as a specific affinity label for the 2nd partial reaction site, reasoning that if an -SH group were in close proximity to pyruvate in this reaction site, it may well react with the pyruvate analogue BP. They found that BP produced complete inactivation of SLPyc by modifying 1 residue / Pyc monomer. Furthermore, by using high voltage paper electrophoresis to resolve the mixture of single amino acids resulting from complete pronase digestion of bromo-[2^{14}C]-pyruvate labelled SLPyc, they were able to show that the labelled residue was indeed a cysteine.

To determine whether this labelled cysteine was actually in the pyruvate binding site Hudson *et al.* (1975) evaluated the effectiveness of the substrates of the two Pyc half reactions, and acetyl-CoA, in protecting against inactivation by BP. They found that ATP and HCO_3^- provided no protection, while acetyl-CoA gave 73% protection, oxaloacetate gave 52-57%, acetyl-CoA + oxaloacetate gave ~88%, and pyruvate gave <15% protection of total Pyc activity (numerical figures are from Hudson, 1974). Similar patterns of protection against inactivation were also observed with these reagents when the 2nd half reaction was studied in isolation (Hudson, 1974).

On the basis of their experiments with bromopyruvate Hudson *et al.* (1975) concluded that this reagent was in fact an affinity label for the pyruvate binding site, specifically labelling one cysteine residue in this site. They suggested that the lower levels of protection provided by pyruvate than oxaloacetate may be explained by an earlier report with the chicken liver enzyme, showing that in the absence of the substrates of the 1st sub-site, pyruvate binds poorly to Pyc (Mildvan *et al.* 1966). Also, the protection by acetyl-CoA was believed to result from the conformational changes it induces, rather than any direct shielding of a cysteine residue in the acetyl-CoA binding site. This was concluded on the basis that it

had been previously shown that both the acetyl-CoA dependent and independent activities of SLPyc were equally affected by the reaction with BP (Nielsen, 1970).

Goss (1978) further investigated the sheep liver enzyme by using the -SH specific reagent DTNB. He found that this reagent totally inactivated the enzyme, modifying approximately 1.3 residues per subunit. Oxaloacetate and acetyl-CoA provided total protection, suggesting that the -SH group that was causing the inactivation was in or close to the 2nd partial reaction site. However, he found that replacement of the 2-nitrobenzoic acid group on the modified -SH, with a cyano group, resulted in recovery of enzymatic activity. This suggested that the modified -SH group could not be acting as a catalytic base in the 2nd half reaction. Furthermore, bromopyruvate, which had the characteristics of an affinity label for the pyruvate binding site in the case of SLPyc (Hudson, 1974), appeared to act as a non-specific alkylating agent with the transcarboxylase enzyme (Goss, 1978). On the basis of these data Goss (1978) concluded that the modified -SH group which was responsible for the inactivation of SLPyc was most likely located close to, but not actually in the pyruvate binding site.

In contrast, upon repeating and extending the DTNB modification studies with the SLPyc enzyme, Bagley (1982) found that the DTNB inactivation was actually reversible by both thiolysis (with DTE), and replacement of the S-(2-nitrobenzoic acid) group with S-cyano or S-fluoro derivatives. Bagley also found that SLPyc was inactivated by three reagents considered diagnostic for conformationally vicinal cysteines (*o*-iodosobenzoate, arsenite and cadmium). In view of these results, and the complete protection against DTNB inactivation by oxaloacetate + acetyl-CoA, Bagley proposed that there is at least one catalytically important cysteine residue in the 2nd partial reaction (transcarboxylation) subsite.

From the preceding discussions it can be seen that there are conflicting data and views regarding the presence of a catalytically important cysteine residue in the transcarboxylation site of pyruvate carboxylase. These differences are further reflected in the mechanisms which have been proposed for the transcarboxylation reactions of Pyc, and all biotin carboxylase enzymes (see Chapter 1, 1.4.3). In view of this uncertainty, the aim of the work presented in this chapter was to use a site-directed mutagenesis approach to

determine whether there is actually a catalytically essential cysteine residue in the pyruvate binding site of Pyc.

Having successfully constructed the double null mutant strain DM18, mutant Pyc enzymes generated by site-directed mutagenesis could be expressed in this yeast strain which contains no background Pyc activity. The effect of the single amino acid substitutions could then be determined from the strains expressing the mutant proteins by directly assaying the Pyc activity in the lysates without having to purify the mutant proteins. Furthermore, as DM18 is an aspartate auxotroph, the effects of the specific amino acid substitutions could be assessed *in vivo* by the ability of the mutant Pyc enzymes to complement this phenotype.

Hence the approach taken in the work presented in this chapter was to separately change each of the 4 cysteine residues in the yeast Pyc1 pyruvate domain (see Fig 4.9) to a similarly shaped residue with different chemical properties, ie serine. Then if one (or more) of these residues is "essential" for the transcarboxylation reaction to proceed, changing this residue should totally destroy all measurable Pyc activity, and the ability of the expressed Pyc enzyme to complement the aspartate auxotrophy of the host strain.

This chapter describes the:

(i) introduction of the cysteine to serine substitutions in the pyruvate domain of yeast Pyc1

(ii) expression of the mutant Pyc1 enzymes in DM18 from single-copy and multi-copy yeast vectors

(iii) effect of the cysteine substitutions on the activity of Pyc and its ability to complement the aspartate auxotrophy of the DM18 host.

7.2 RESULTS

7.2.1 The mutagenesis approach

Figure 4.9 showed that both yeast Pyc isozymes contained the same four cysteine residues in the pyruvate domain. All other species for which the Pyc sequence has been determined thus far have 5 cysteine residues in the pyruvate domain, and only cysteine "6"

(residue 779 in Pyc1, see Fig. 4.9) is present in the same position in the amino acid sequences in all the Pyc enzymes. Hence this residue appeared to be the most likely candidate for the "putative" catalytically-essential cysteine involved in the transcarboxylation reaction.

However, the possibility that one of the 3 other cysteine residues in the pyruvate domain of the yeast enzymes could actually be the essential cysteine could not be entirely ruled out. For example, one of these 3 "non-conserved" residues could possibly occupy a very similar position in the tertiary structure as one of the non-aligned cysteines present in Pyc from the higher organisms (see Fig. 4.9), thus allowing them to share the same functional role. Consequently it was decided to separately mutate each of the 4 cysteine residues present in the yeast enzymes instead of just cysteine "6".

Yeast Pyc1 was chosen as the enzyme in which to conduct the mutagenesis in preference to Pyc2, due to the higher enzymatic activity of this isozyme. Consequently, with Pyc1 there would be a greater difference in measurable enzymatic activity between the wild-type and non-functional enzymes. In addition, single-copy expression of wild-type Pyc1, but not Pyc2, was known to be sufficient for yeast to grow in the absence of aspartate (see Chapter 6). Thus, the mutagenesis of a catalytically essential cysteine residue in the pyruvate domain of Pyc1 was expected to be clearly discernible by the inability of the mutant Pyc1 enzyme to complement the aspartate auxotrophy of the host strain DM18.

7.2.2 Introduction of the cysteine to serine substitutions

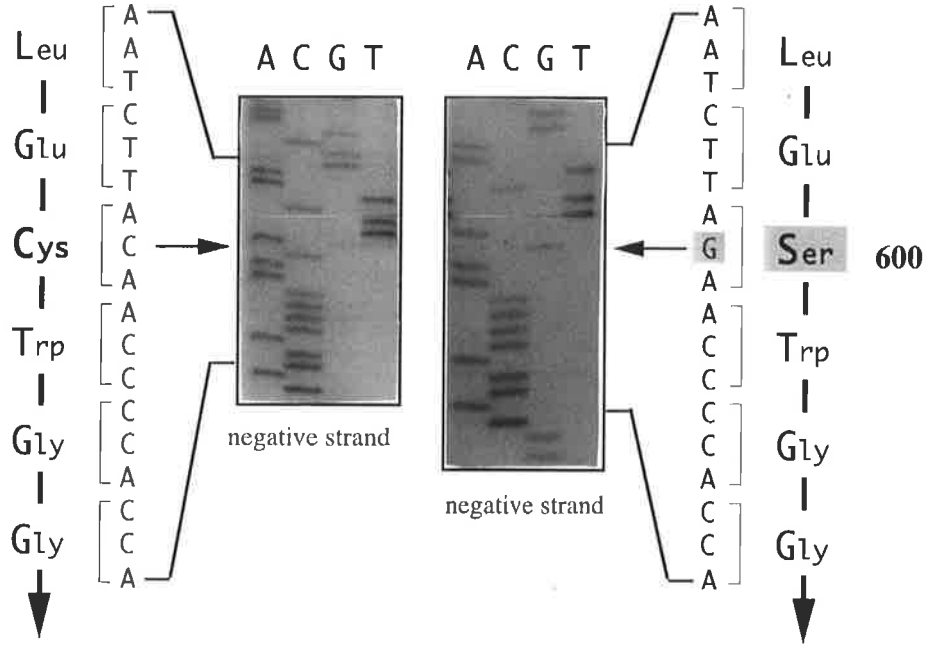
Site-directed mutagenesis of each of the 4 cysteine residues in the Pyc1 pyruvate domain was performed by the method of Kunkel *et al.* (1987) with full length *PYC1* DNA (4.69 kb *Eco*RV-*Dra*I fragment in KS+ bluescript) using the oligonucleotides YPC04 to YPC07 (see Chapter 2) for cysteine residues "4" to "7" respectively (see Fig. 4.9). Each C to S amino acid substitution involved a single base mutation so as to change the codon TGT to TCT. Figure 7.1 shows the DNA sequence of the negative or non-coding strand for the wild-type and C-S mutants at the sites encoding each of the four cysteine residues, confirming that the correct mutations had been introduced.

Figure 7.1 Nucleotide changes introduced to produce the cysteine to serine substitutions in the Pyc1 pyruvate domain

The parental wild-type DNA sequence **(i)**, and the mutant sequence produced by site-directed mutagenesis with the oligonucleotides YPC04 to YCP07 **(ii)** for the sites encoding cysteines "4", "5", "6", and "7" in the Pyc1 pyruvate domain (see Fig. 4.9) are shown together in parts **(a)**, **(b)**, **(c)**, and **(d)** respectively. Shading indicates the residues changed by mutagenesis, horizontal arrows show the positions of the nucleotide changes in the DNA sequence, bold vertical arrows indicate the direction of translation, and the numbers refer to the residue positions in Pyc1.

a) **YPC04**

(i) Wild-type **(ii) Mutant**



b) **YPC05**

(i) Wild-type **(ii) Mutant**

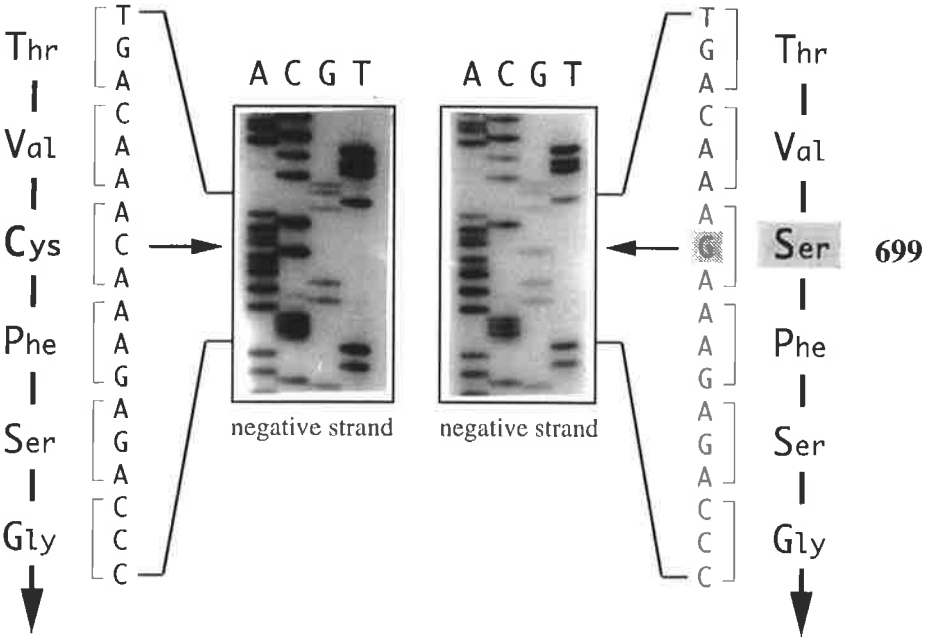
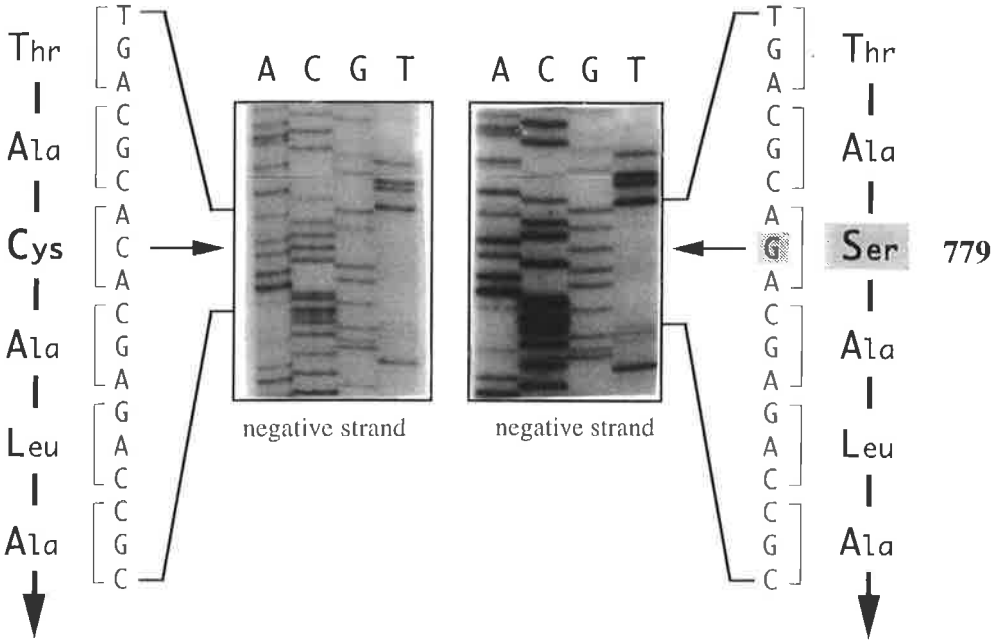


Figure 7.1 (continued)

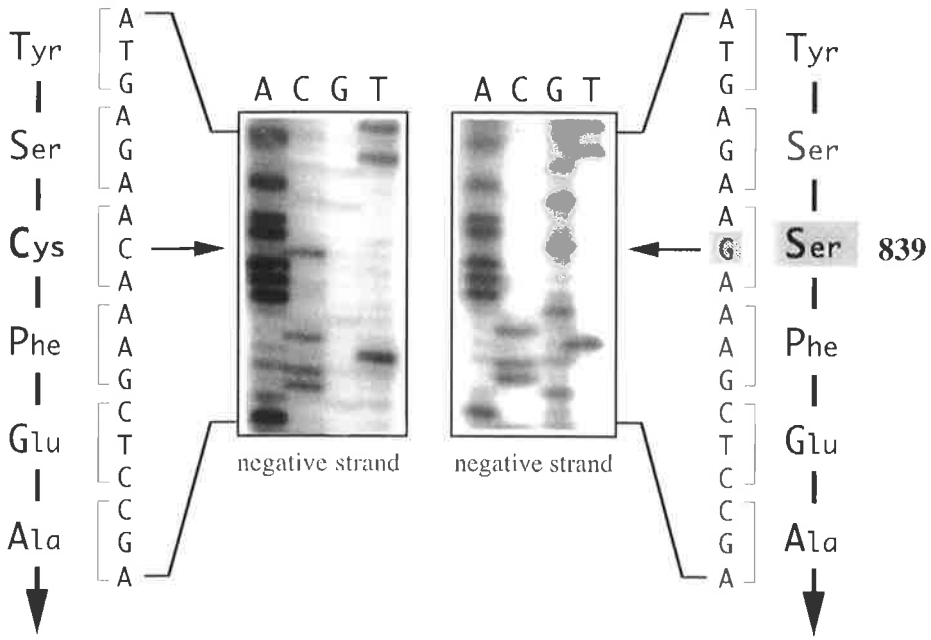
c) **YPC06**

(i) Wild-type **(i) Mutant**



d) **YPC07**

(i) Wild-type **(ii) Mutant**



Ideally, to ensure that the sequence is wild-type apart from the intended mutagenic change, mutagenesis should be performed on a template DNA fragment small enough to be completely sequenced after mutagenesis. With the C to S mutations in the pyruvate domain of Pyc1 this was not possible due to the lack of nearby suitable restriction sites. However this was not expected to be a problem as: a) it was considered unlikely that a misplaced mutation at a totally different site in the primary sequence (resulting from the incorrect annealing of the mutagenic oligonucleotide) could compensate for the loss of activity caused by the mutagenesis of a cysteine residue which is essential for catalysis, b) sequence comparisons revealed that the only regions of *PYC1* with significant homology to each of the mutagenic oligonucleotides were the intended sites, and c) DNA sequencing with the mutagenic oligonucleotides showed that they were each only priming from the desired site.

7.2.3 Preparation of the expression constructs

7.2.3.1 The multi-copy episomal expression constructs

The vector chosen for high level expression of the cysteine (CYS) mutants in the yeast strain DM18 was the multi-copy vector pVT100-U (Vernet *et al.*, 1987; see Fig. 7.2b). This vector was chosen for the following reasons: a) the level of transcription from the *ADH1* promoter had been estimated to be at least 1% of the poly(A)$^+$ RNA (Ammerer, 1983), and thus there should be sufficient Pyc expressed to facilitate subsequent purification if required, b) the promoter had been modified to provide constitutive expression (Vernet *et al.*, 1987), c) there were convenient restriction sites in the poly-linker to allow cloning of the *PYC1* gene as a single DNA fragment, and d) it contained a suitable marker gene for selection in strain DM18 (*URA3*).

In order to express the wild-type and CYS mutant Pyc1 constructs downstream of the *ADH1* promoter, a *Xho*I site was introduced 10 bp 5' of the initiating ATG in Pyc1(see Fig. 7.2a), using the method of Kunkel *et al.* (1987) with the oligonucleotide O2021 (see Chapter 2). This mutagenic step was performed on the full length *PYC1* gene fragment (4.69 kbp *Eco*RV-*Dra*I fragment in KS+ bluescript) prior to the mutagenesis introducing the C to S substitutions described above, and the correct introduction of this site was verified by DNA sequencing and restriction analysis (data not shown).

Figure 7.2 Constructs to express the Pyc1 cysteine mutants

(a) Introduction of the 5' *Xho*I site The shading highlights the mutagenic changes that were made to introduce a 5' *Xho*I cloning site 10 bp upstream of the first ATG in the Pyc1 open reading frame. The original wild-type sequence is shown below and the underlining indicates the position of the introduced *Xho*I site.

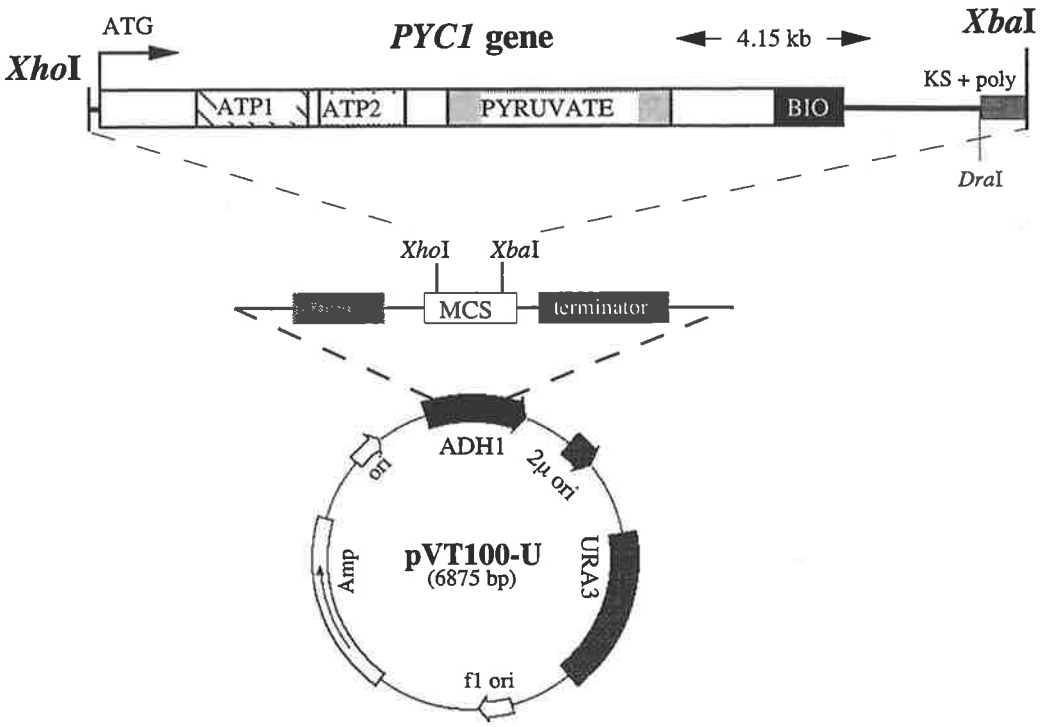
(b) Schematic diagram of the multi-copy pVT constructs The wild-type and each of the 4 cysteine mutant *PYC1* coding regions were cloned into the multiple cloning site downstream of the *ADH1* promoter present in the pVT100-U yeast expression vector (Vernet *et al.*, 1987). In each case the coding region was cloned into the vector as a 4.15 kbp fragment extending from the 5' introduced *Xho*I site in *PYC1* shown in part (a), to an *Xba*I site in a flanking KS+ polylinker sequence downstream of the Pyc2 termination sites.

(c) Schematic diagram of the single-copy YCpLAC33 vector (Geitz and Sugino, 1988) The cloning sites shown refer to those used to insert the 4.15 kbp *Xho*I-*Xba*I *PYC1* coding sequence fragment shown in part (b).

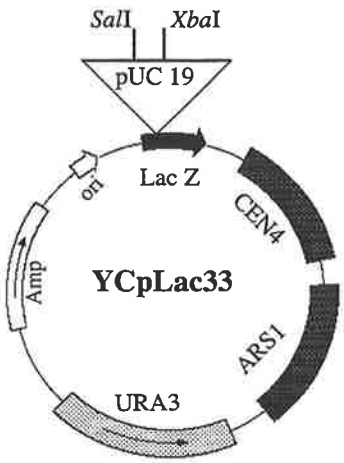
a)


XhoI site
M S Pyc1
5'-- CCTCGAGGAACA ATG TCG --3'
-10
TTA

b)


ATG
XhoI
PYC1 gene
4.15 kb
XbaI
KS + poly
ATP1
ATP2
PYRUVATE
BIO
DraI
XhoI
XbaI
MCS
terminator
ADH1
2μ ori
ori
Amp
pVT100-U
(6875 bp)
URA3
f1 ori

c)


SalI
XbaI
pUC 19
Lac Z
ori
CEN4
Amp
YCpLac33
ARS1
URA3

The wild-type, and each of the CYS mutant *PYC* constructs were then subsequently cloned into the pVT100-U expression vector as 4.15 kbp *Xho*I-*Xba*I fragments (Fig 7.2b). Correct recombinants were obtained by screening for plasmid size by the lid lysate technique (Hoekstra, 1988), and were confirmed by restriction analysis. The wild-type, and each of the CYS mutant constructs will be referred to as follows: wild-type Pyc1, pVT/13; cysteine "4" mutant, pVT/4.49; cysteine "5" mutant, pVT/5.3; cysteine "6" mutant, pVT/6.2; cysteine "7" mutant, pVT/7.11.

7.2.3.2 The single-copy integrative expression constructs

As the hypothesis being investigated in this chapter was that Pyc contains an essential cysteine residue in the 2nd partial reaction site, verification of this hypothesis would require that substitution of this particular CYS residue with serine should produce a totally non-functional enzyme. Hence, it was expected expressing each of the CYS mutant Pyc1 enzymes in the double null mutant strain DM18 (*pyc1 / pyc2*) from the multi-copy pVT100-U vector should be sufficient to determine whether any the CYS residues in the pyruvate domain are catalytically essential. After all, the expression level should be irrelevant if an enzyme is totally non-functional. However, as a closer approximation to the native expression levels, it was decided to express each of the Pyc1 CYS mutants in DM18 from a single copy vector, in addition to the multi-copy approach.

The vector chosen for single-copy expression of the cysteine mutants was the integrative vector YCpLAC33 (Geitz and Sugino, 1988; see Fig. 7.2c). This vector was chosen because: a) it contained the lacZ gene simplifying the selection of recombinants, b) there were convenient restriction sites in the poly-linker to allow cloning of the *PYC1* gene as a single DNA fragment, and c) it contained a suitable marker gene for selection in strain DM18 (*URA3*).

An oversight was made in construction of the YCpLAC33 constructs. Instead of cloning the entire *PYC1* gene including the promoter region into the poly-linker of YCpLAC33 so that Pyc1 would be expressed from its own promoter, the constructs were prepared with the same 4.15 kb *Xho*I-*Xba*I *PYC1* fragments (lacking the *PYC1* promoter) that were used to prepare the pVT constructs. Hence, it was subsequently realised that the

Pyc1 enzymes would probably not be expressed from these constructs when transformed into yeast. Nevertheless these constructs were tried, as even if unsuccessfull the multi-copy approach should be sufficient to identify the "essential" cysteine, and genes are sometimes fortuitously expressed in yeast when cloned 3' of sequences derived from *E. coli.*

The wild-type, and each of the CYS mutant constructs in YCpLAC33 will be referred to as follows: wild-type Pyc1, YCp/13; cysteine "4" mutant, YCp/4.49; cysteine "5" mutant, YCp/5.3; cysteine "6" mutant, YCp/6.2; cysteine "7" mutant, YCp/7.11.

7.2.4 Expression of the wild-type and CYS mutant constructs

Each of the above mentioned multi-copy (pVT100-U) and single-copy (YCpLAC33) wild-type and CYS mutant constructs were transformed into the *pyc* double null mutant strain DM18 (*pyc1/pyc2*) by the Li^+ method. URA^+ colonies for each of the constructs were purified by several rounds of re-streaking, and representative clones containing each construct were selected. These clones were grown to mid-log phase in glucose minimal media under the necessary selection conditions, and in the presence of 8 mM aspartate. The quantity of Pyc1 being expressed in each clone was then analysed by SDS-PAGE and Western blotting, probing with avidin alkaline phosphatase.

7.2.4.1 Expression from the multi-copy pVT100-U vector

Figure 7.3a shows that approximately equal amounts of Pyc1 were present in all the clones containing the pVT100-U vector expressing either the wild-type Pyc1 enzyme (pVT/13), or any one of the CYS mutant enzymes (pVT/4.49, pVT/5.3, pVT/6.2, pVT/7.11) Also, the intensity of the 130 kDa biotinylated Pyc protein band in each of these clones was noticeably higher than that of the ~47 kDa unidentified biotinylated protein band, indicating that there was a high level of Pyc expression. With strains containing chromosomally encoded Pyc genes for example, the biotinylated ~47 kDa protein is always by far the most abundant biotinylated protein, ie the reverse of what was observed with these clones (compare Fig. 6.5 & Fig. 7.3).

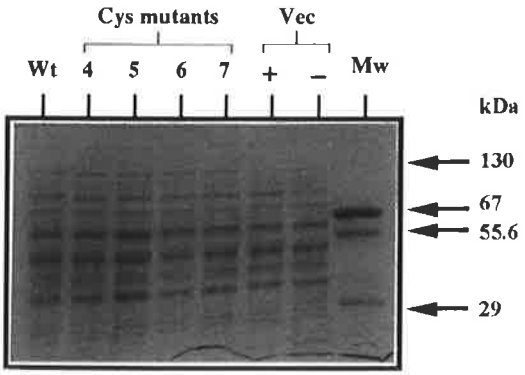
Western blotting also revealed that some of the Pyc1 protein in each of these clones was being degraded, as evident from the ladder of lower molecular weight biotinylated bands

Figure 7.3 Expression of the Pyc1 cysteine mutants.

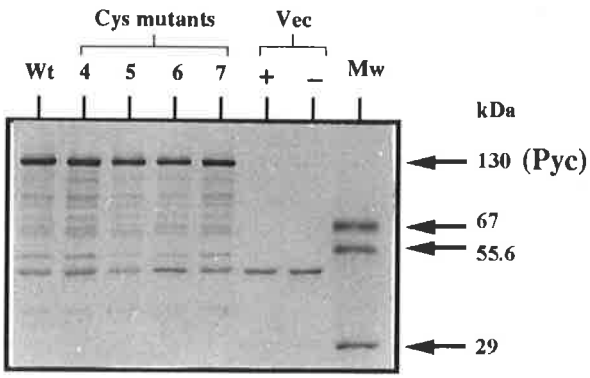
Duplicate protein samples (~10 μg) were prepared from crude cell lysates obtained from strains containing the vector alone (negative control), and from strains expressing either the wild-type Pyc1 enzyme (positive control) or one of the 4 cysteine mutant enymes. In each case the cells were grown to mid-log phase in glucose minimal media in the presence of 8 mM aspartate and the appropriate amino acids. Samples from the strains containing the same vector type (pVT100-U or YCpLAC33) were analysed together by running duplicate 10% Tris-glycine SDS-PAGE gels. One gel **(i)** was Coomassie stained for total proteins, and the other gel **(ii)** was used for the detection of biotinylated proteins by Western blotting with avidin alkaline phosphatase (as outlined in Chapter 2). **Part a)** shows the results for the strains containing the pVT100-U constructs, while **part b)** shows the results for the strains containing the YCpLA33 constructs. **Lanes**: **Mw**, biotinylated molecular weight markers; **Lane Wt**, strains expressing wild-type Pyc1, ie gel a) containing pVT/13, gel b) containing YCp/13; **Lanes 4 - 7**, strains containing pVT100-U (gel a) or YCpLAC33 (gel b) expressing the Pyc1 cysteine mutant constructs 4.49, 5.3, 6.2 or 7.11 respectively. **Lane Vec +**, strain containing the respective expression vector without the *PYC1* insert (negative control); **Vec -**, the untransformed host strain DM18 (negative control).

a) Multi-copy expression

(i) Coomassie Stain

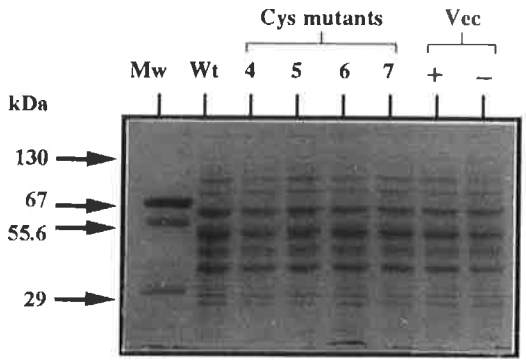


(ii) Western Blot

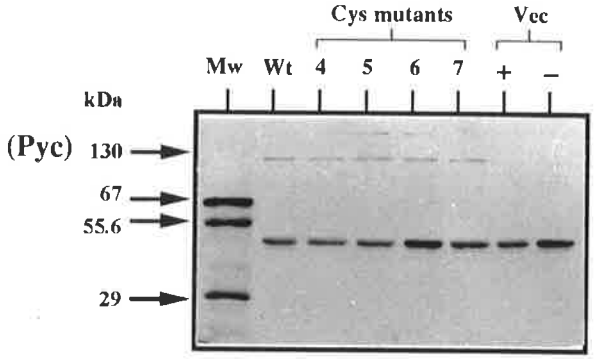


b) Single-copy expression

(i) Coomassie Stain



(ii) Western Blot



below the ~130 kDa Pyc1 band. Furthermore, this Pyc1 degradation was also observed in repeated experiments using protein samples prepared immediately after cell lysis (data not shown), suggesting that it was almost certainly occurring *in vivo*, and was probably a function of the very high levels of Pyc1 expression.

The Vec +/- Lanes revealed, as expected, that there was no Pyc protein present in the lysates from the host strain DM18 (Lane Vec -), nor in the host strain transformed with vector containing no *PYC1* DNA insert (Lane Vec+).

7.2.4.2 Expression from the single-copy YCpLAC33 vector

Figure 7.3b shows that approximately equal amounts of Pyc1 were present in all the clones containing the integrative single-copy YCpLAC33 vector expressing either the wild-type Pyc1 enzyme (YCp/13), or one of the CYS mutant Pyc enzymes (YCp/4.49, YCp/5.3, YCp/6.2, YCp/7.11), as was the case with the pVT100-U clones. However, a comparison of Fig. 7.3 a & b clearly shows that the level of Pyc1 expression in each of these clones was much less than that observed with the pVT100-U constructs, and indeed also less than the Pyc expression observed in laboratory strains containing the endogenous chromosomally encoded *PYC* gene(s) (compare Fig. 7.3 with Fig. 6.5).

Since all of these constructs contained only the *PYC1* coding and 3' non-coding sequences (due to the oversight explained in section 7.1.2.1) cloned into the coding region of the LacZ gene, there must have been some sequence element(s) in the vector 5' to the *PYC1* coding region which was fortuitously acting as a yeast promoter. This uncertainty regarding the sequence elements responsible for this expression was not considered to be a problem, as analysis of effects of the CYS mutations using these clones would simply require that the relative levels of Pyc1 protein in each of these clones be known. The actual promoter providing this expression, fortuitous or otherwise, was considered to be essentially irrelevant.

7.2.5 Growth phenotypes of the CYS mutants

7.2.5.1 Growth curves of the CYS mutants expressed from pVT100-U

Figure 7.4 shows the growth profiles for the clones expressing each of the CYS mutant enzymes from the multi-copy pVT100-U vector in the presence or absence of 8 mM aspartate (Fig. 7.4 a & b respectively). From figure 7.4a it can be seen that in the presence of 8 mM aspartate each of the clones displayed essentially the same growth characteristics, ie a lag phase of ~10 hours, and a log phase of ~ 15 hours with a doubling time of ~8 hours, reaching a maximum OD_{600} of 2 - 2.5 in the stationary phase. This similarity in growth profiles was to be expected, as in 8 mM aspartate the cells requirement for oxaloacetate would be entirely supplied by the aspartate in the growth media. Hence, the presence or absence of Pyc activity would make little difference to the growth rate.

In the absence of aspartate the total dependence on Pyc for growth can clearly be seen by the lack of growth in the negative control, ie yeast (DM18) transformed with the pVT100-U vector alone (Fig. 7.4b). In contrast, clones expressing the wild-type Pyc1 or any of the CYS mutants were all able to grow on glucose minimal media the absence of aspartate. This indicates that none of the CYS to SER substitutions in the pyruvate domain had totally destroyed the ability of the enzyme to carry out the 2^{nd} partial reaction (transcarboxylation).

The cultures containing the constructs with substitutions at CYS residues "5" and "7" appeared to grow equally well as those containing the wild-type construct, while those containing the constructs 4.49 and 6.2 reached a final OD_{600} of approximately half that of the other cultures. This may indicate that these two CYS residues are more important for the activity of the enzyme. However, previous experiments had revealed that the final OD_{600} is somewhat variable for cultures containing any one of the pVT/Pyc1 constructs (wild-type or mutant), depending on the particular isolate and experiment. Furthermore, on solid media (glucose minimal media plates), clones containing the pVT constructs 4.49 or 6.2 appeared to grow equally well as those containing the wild-type Pyc1, or the other CYS mutant constructs (5.3 & 7.11) (data not shown). Thus it would seem quite likely that the lower final OD_{600} of the clones containing the 4.49 or 6.2 constructs was probably due to clonal variation, possibly as a result of differences in the expression level.

7.2.5.2 Growth curves of the CYS mutants expressed from YCpLAC33

The growth profiles for the clones expressing each of the CYS mutant Pyc1 enzymes from the single-copy YCpLAC33 vector in the presence or absence of 8 mM aspartate are shown in Fig. 7.5a & Fig. 7.5b respectively. In the presence of 8 mM aspartate all the strains grew equally well, and, as the growth rate was independent of Pyc levels at this aspartate concentration, the growth curves were essentially identical those obtained with the multi-copy constructs (compare Fig. 7.4a and 7.5a).

The growth curve of the clone containing the YCpLAC33 vector alone (negative control) clearly shows once again that functional Pyc expression is absolutely required for DM18 transformants to grow in the absence of aspartate. The initial minor rise in OD_{600} of 0.1 -0.15 units which occurred with this clone was almost certainly the result of the small amount of aspartate present in the media containing the initial inoculum (also occurred with the pVT100-U vector). In contrast, the wild-type and CYS mutant constructs in YCpLAC33 exhibited at least minimal growth in the absence of aspartate, suggesting that even at the low levels of Pyc1 expression produced with this vector (section 7.2.3.2) each of the CYS mutant enzymes were still able to produce enough oxaloacetate to sustain a very slow rate of growth.

Unexpectedly, the clone expressing the 7.11 mutant exhibited a growth rate and final OD_{600} considerably higher than those expressing all the other Pyc1 enzymes, including the wild-type enzyme, which in fact sustained the least growth. However, no such differences were noticed with growth on solid media (glucose minimal media plates)

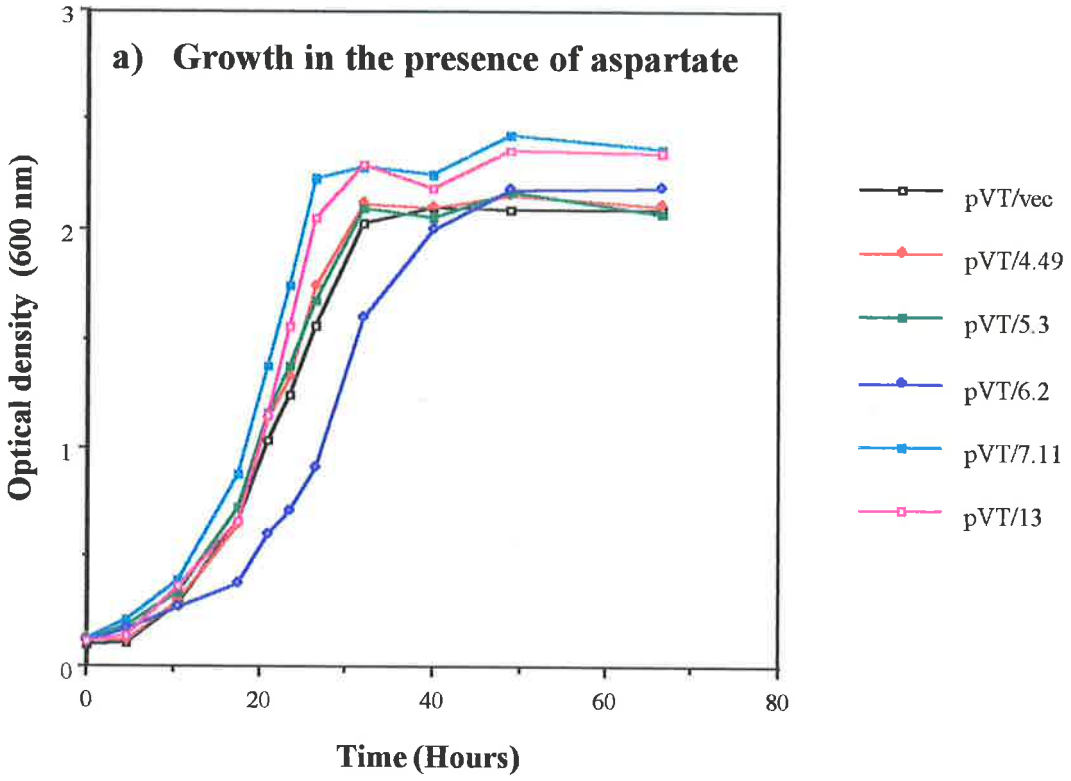
7.2.6 Enzyme activity of the CYS mutants

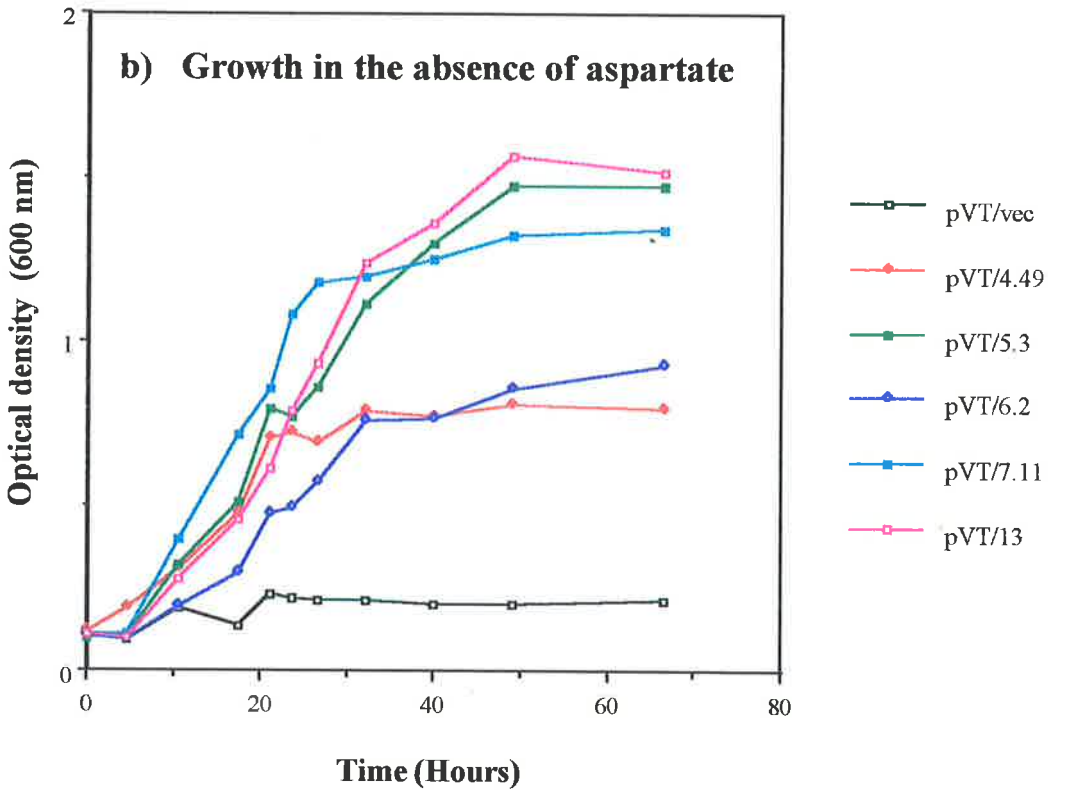
7.2.6.1 Activity of the CYS mutants expressed from pVT100-U

Table 7.1 shows the results of the Pyc enzyme assays performed on separate aliquots of the same mid-log phase lysates (+ aspartate) that were used to determine the Pyc expression levels for each of the CYS mutants expressed from the multi-copy vector pVT100-U.

Figure 7.4 Growth profiles of the clones expressing the cysteine mutants from the multi-copy vector

Shake flask glucose minimal media cultures (50 mL) of each of the pVT100-U yeast clones expressing either the wild-type Pyc1 or one of the four ("4" - "7") cysteine mutant Pyc1 enzymes were inoculated to a starting optical density at 600 nm (OD_{600}) of 0.1 in the presence **(a)** or absence **(b)** of 8 mM aspartate, and the OD_{600} of the cultures was monitored at regular time intervals. The abbreviations for each of the yeast clones refer to DM18 (*pyc1/pyc2*) transformants containing the following constructs: pVT/vec, the parental pVT100-U vector (Vernet *et al.*, 1987) (negative control) ; pVT/4.49, the pVT100-U vector expressing the Pyc1 C600S mutant enzyme (CYS "4" mutant); pVT/5.3, the pVT100-U vector expressing the Pyc1 C699S mutant enzyme (CYS "5" mutant); pVT/6.2, the pVT100-U vector expressing the Pyc1 C779S mutant enzyme (CYS "6" mutant);.pVT/7.11, the pVT100-U vector expressing the Pyc1 C839S mutant enzyme (CYS "7" mutant); pVT/13, the pVT100-U vector expressing the wild-type Pyc1 enzyme (positive control). The number used to signify each of the cysteine mutants (CYS "4" to "7') refer to the arbitrary numbers given to the corresponding cysteine residues in the pyruvate domain homology diagram (Fig. 4.9), and the number of the oligonucleotide used mutate these residues to serines.


a) Growth in the presence of aspartate
Optical density (600 nm)
Time (Hours)
0
1
2
3
0
20
40
60
80
pVT/vec
pVT/4.49
pVT/5.3
pVT/6.2
pVT/7.11
pVT/13


b) Growth in the absence of aspartate
Optical density (600 nm)
Time (Hours)
0
1
2
0
20
40
60
80
pVT/vec
pVT/4.49
pVT/5.3
pVT/6.2
pVT/7.11
pVT/13

Construct	Enzyme Activity (mU/mg)
pVT/vec	0.00
pVT/4.49	67.1 +/- 4.7
pVT/5.3	21.7 +/- 2.3
pVT/6.2	80.5 +/- 6.4
pVT/7.11	80.6 +/- 6.6
pVT/13	69.1 +/- 7.5

TABLE 7.1 **Enzyme activity of CYS mutant constructs expressed from the multi-copy vector pVT100-U.**

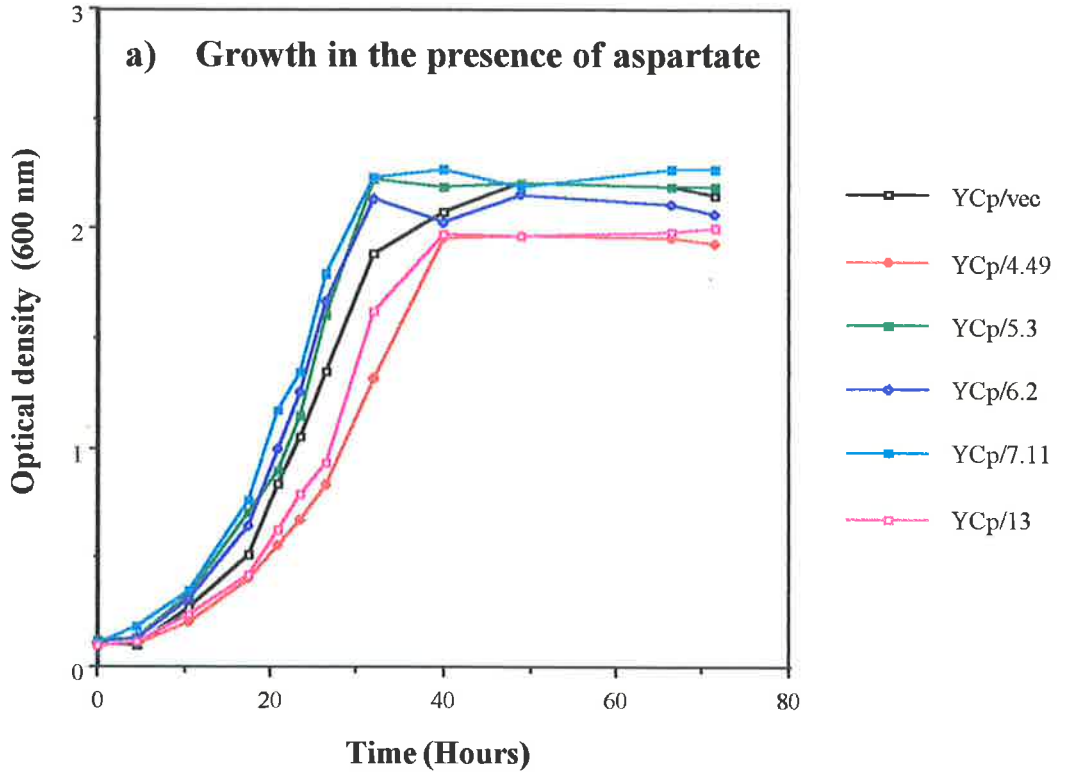
From Table 7.1 it can be seen that all of the CYS mutants contained appreciable levels of enzyme activity. In fact the clones expressing the mutant enzymes from pVT100-U had from 4 to 17x the activity present in DBY746 (4.65 mU / mg, see Table 6.3), the parental strain of DM18 which contains both the wild-type *PYC* genes. This high level of activity was in agreement with the high levels of Pyc detected on the Western blots by avidin alkaline phosphatase (see section 4.2.3.1). Furthermore, the CYS mutants 4.49, 6.2 and 7.11 had an enzyme activity equal to or greater than that of the wild-type positive control pVT/13. The pVT/5.3 clone had a level of Pyc activity approximately 3.3 fold less than pVT/13. This may indicate that the CYS "5" substitution had the greatest effect on the activity of the enzyme. However, the growth curves for the CYS 5.3 construct expressed from both the single and multi-copy vectors were both essentially the same as for the wild-type enzyme (see Fig 7.4 & 7.5). Hence it would appear that the most likely explanation for the lower activity of the pVT/5.3 clone was that the lysates for this clone were prepared during a different experiment than those for the other clones.

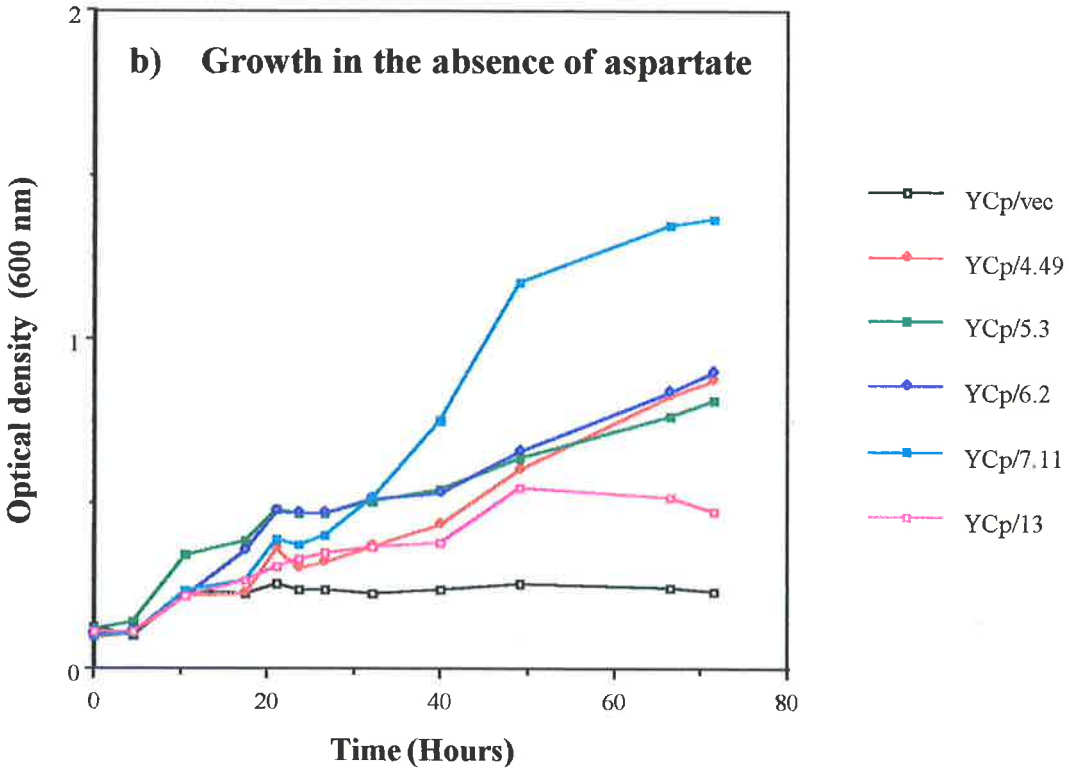
7.2.6.2 Activity of the CYS mutants expressed from YCpLAC33

Due to the technical difficulties associated with the low Pyc1 expression levels from YCpLAC33, it was not possible to determine accurate values for the Pyc enzyme activity in

Figure 7.5 Growth profiles of the clones expressing the cysteine mutants from the single-copy vector

Shake flask glucose minimal media cultures (50 mL) of each of the YCpLAC33 yeast clones expressing either the wild-type Pyc1 or one of the four ("4" - "7") cysteine mutant Pyc1 enzymes were grown in the presence (**a**) or absence (**b**) of 8 mM aspartate as described in Figure 7.4. The abbreviations for each of the YCpLAC33 yeast clones correspond to those for the equivalent pVT100-U wild-type and mutant clones as outlined in Figure 7.4.


a) Growth in the presence of aspartate
Optical density (600 nm)
Time (Hours)
0
1
2
3
0
20
40
60
80
YCp/vec
YCp/4.49
YCp/5.3
YCp/6.2
YCp/7.11
YCp/13


b) Growth in the absence of aspartate
Optical density (600 nm)
Time (Hours)
0
1
2
0
20
40
60
80
YCp/vec
YCp/4.49
YCp/5.3
YCp/6.2
YCp/7.11
YCp/13

the clones containing the various CYS mutant constructs. However, in the assays that were performed low levels of enzyme activity were detected in the positive control expressing wild-type Pyc1, and in each of the CYS mutants, while no activity could be detected in the negative control clone (YCpLAC33 vector alone).

7.3 DISCUSSION

7.3.1 The effect of the cysteine substitutions

As has been stated previously, the aim of the work presented in this chapter was to determine whether pyruvate carboxylase contains an essential cysteine residue involved in the transcarboxylation reaction (2nd partial reaction of Pyc). Clearly, if a particular cysteine residue was located in the active site of the enzyme, playing an essential role in the chemistry of the reaction, then substitution of this residue would be expected to destroy the activity of the enzyme.

The growth curves from the clones containing the mutant enzymes expressed from the multi-copy vector pVT100-U revealed that each of the 4 CYS mutant enzymes were able to complement the aspartate requirements of DM18 (Fig. 7.4) when expressed at higher than normal levels (compare Fig. 6.5 and Fig. 7.3a). In addition, growth curves from the clones with the single-copy vector showed that, when expressed in the absence of aspartate at levels less than that observed in the parental wild-type strains (compare Fig. 6.5 and Fig. 7.3b), the mutant enzymes were still able to support comparable levels of growth as the wild-type enzyme (Fig. 7.5). Enzyme assays also indicated that none of the mutant enzymes had activities that were markedly less than the wild-type enzyme.

On the basis of sequence comparisons CYS "6" (see Fig. 4.9) was considered to be the most likely candidate for a catalytically essential residue, as this was the only cysteine residue contained in all Pyc enzymes sequenced thus far. However, while some variations were observed in the ability of the various mutants to complement the aspartate auxotrophy of DM18 when they were expressed from the multi-copy vector, expression from the single copy vector failed to produce the same variation patterns. This suggested that the observed differences in the growth curves were most likely due to differences in expression levels

(possibly due to copy number variation) or other sources of clonal variation, rather than the mutations themselves. Furthermore, there were no consistent patterns maintained between the growth curve results and the enzyme assays to indicate that any one of the mutants (CYS "6" for example) was more essential for the activity of the enzyme than the other three residues.

In conclusion then, neither CYS "6" nor any of the other 3 residues located in the region of Pyc1 identified by sequence comparisons to be pyruvate binding / transcarboxylation domain (see Fig. 4.9) appeared to be essential for the activity of the enzyme.

7.3.2 Implications regarding the proposed reaction mechanisms

In Chapter 1 (section 1.4.3) the most likely proposed mechanisms for the 2nd partial reaction of pyruvate carboxylase were outlined. In this section it was explained that the 2nd partial reaction involves the transfer of the carboxyl group from the N-1 of biotin to pyruvate, with the concomitant transfer of a proton from pyruvate to the N-1 of biotin. Furthermore, Attwood *et al.* (1986b) showed by studying the ^{14}C and ^{2}H isotope effects that the carboxyl and proton transfers occur in different steps, and the reaction appears to involve the enol form of pyruvate. It was also pointed out in Chapter 1 that from the data available it is not clear whether the mechanism of the reaction involves the pathway proposed for Pyc by Attwood *et al.* (1986b) (Scheme IV), or an analogous pathway to that proposed for transcarboxylase by O'Keefe and Knowles (1986) (Scheme V).

The mechanism proposed by Attwood *et al.* (1986b) involves a carboxylation step(s) flanked by proton transfer steps, the first of which involves a base (S^- in an ion pair with a positive group) catalysed enolisation of pyruvate (Scheme IV). These workers chose this mechanism in preference to a mechanism analogous to that proposed for transcarboxylase by O'Keefe and Knowles (1986) on the basis that the empirical isotope effects did not appear to agree with the theoretical isotope effects calculated for this latter mechanism. They made this conclusion despite the fact that Attwood and Cleland (1986) had found no evidence for acid-base kinetics in the pH profiles for the decarboxylation of Pyc, a result which would appear

to be more consistent with the pH independence predicted by the mechanism of O'Keefe and Knowles (1986).

The inclusion of the cysteine-lysine (or other + amino acid) ion pair in the mechanism proposed by Attwood *et al.* (1986b) was made on the basis that the empirical isotope effects would only agree with the theoretical values for the mechanism they proposed as long as the enzymic base had a low deuterium fractionation factor. As cysteine was the only amino acid known to have a fractionation factor less than 1, these workers concluded that the enzymic base must be a cysteine residue. However the work presented in this chapter indicates that Pyc does not contain an essential cysteine residue in the pyruvate domain.

Similarly, the crystal structures of enolase (Lebioda and Stec, 1991) and adenosine deaminase (Wilson *et al.*, 1991) have revealed that neither of these enzymes contain a cysteine residue in their active sites, despite the fact that the low deuterium fractionation factors observed with these two enzymes had led to the belief that they both contained a cysteine residue acting as a catalytic base. Cleland (1992) has subsequently explained this apparent inconsistency, by pointing out that the groups responsible for the low fractionation factors were low-barrier hydrogen bonds formed between two groups with similar pKs.

Therefore, if the data and calculations of Attwood *et al.* (1986b) favouring their mechanism are correct, it would seem quite possible that an oxygen involved in a low-barrier hydrogen bond may be acting as the base in the pyruvate enolisation step instead of their proposed cysteine residue However, this would still not explain the apparent pH independence of oxaloacetate decarboxylation (Attwood and Cleland, 1992).

Clearly them, from the above discussions it can be seen that there are many features of the mechanism of transcarboxylation of pyruvate by Pyc that are still unknown, despite the fact that the mutagenesis experiments described in this chapter have been able to show that this reaction does not involve a cysteine residue functioning as a catalytic base. Another feature of the chemistry of the transcarboxylation which has yet to be understood, for example, is the observed increase in lability of the carboxy-biotin complex when in the locality of the 2nd partial reaction site (Easterbrook-Smith *et al.*, 1976).

7.3.2 Possible explanations to account for the data implying the presence of an essential cysteine residue

As was outlined in the introduction to this chapter, there have been a number of chemical modification studies which have shown that pyruvate carboxylase contains an essential cysteine residue. Briefly, the reagents DTNB (Bais, 1974; Goss, 1978; Bagley, 1982), NEM (Nielsen, 1970), and BP (Nielsen, 1970; Hudson *et al.*, 1975) have all clearly indicated that SLPyc contains 1 - 2 cysteine residues that are essential for the activity of the enzyme. However, these chemical modification studies have not been able to unequivocally locate this essential residue in either the 1st or the 2nd partial reaction site.

The inability of ATP and bicarbonate to protect against inactivation by BP (Hudson *et al.*, 1975) or DTNB (Goss, 1978; Bagley. 1982) has been interpreted by these workers to indicate that the essential cysteine was not located in the 1st partial reaction site. However, the partial protection provided by oxaloacetate or pyruvate, and the complete protection provided by oxaloacetate plus acetyl-CoA has been more difficult to interpret. Hudson *et al.* (1975) and Bagley (1982) have each interpreted their respective similar protection results to indicate that the modified cysteine residue is actually located in the 2nd partial reaction site being directly involved in catalysis. In contrast, Goss (1978) interpreted the incomplete protection in the absence of the allosteric activator acetyl-CoA to indicate that the cysteine may be close to, but is not actually within the active site.

The question that has to be resolved is, does the essential cysteine residue directly participate in catalysis, or is the residue simply required to maintain the enzyme's conformation, being located near enough to the active site to be protected from chemical modification by the conformational changes induced by the binding of substrates or acetyl-CoA ?. However, it has long been recognised that it is essentially impossible to distinguish between these two alternatives on the basis of chemical modification techniques alone (Spradlin and Thoma, 1970), especially in the case of allosteric enzymes such as Pyc (Froede *et al.*, 1968). Correctly distinguishing between these alternatives requires an investigation of the comparative physical properties of the native and chemically modified enzymes, so as to determine whether any conformational or other structural changes have occurred.

Hudson *et al.* (1975) and Bagley (1982) did not include an investigation of the comparative physical properties of the native and chemically modified enzymes. Hence, although they concluded that the modified cysteine residue was playing a role in catalysis, they could not actually rule out the possibility that the modified cysteine residue may simply be near the active site, playing an important role in maintaining the correct 3^o or 4^o structure of the enzyme. Therefore, the mutagenesis results presented in this chapter showing that none of the 4 cysteine residues in pyruvate domain of yeast Pyc1 are directly involved in catalysis (not essential for enzyme activity), are not at all inconsistent with these chemical modification studies. Instead, taken together these results may suggest that the essential cysteine residue is located outside of the pyruvate domain of Pyc, and has a structural rather than catalytic role.

Actually, there are other enzymes in the literature for which site-directed mutagenesis has shown that cysteines implicated as catalytic residues on the basis of chemical modification studies, were actually not involved in catalysis. This was the case for example with the pi, mu, and alpha-classes of glutathione S-transferases (reviewed in Wilce and Parker, 1994).

So, the question which needs to be addressed is, what is the actual role of the essential cysteine residue in Pyc ?. One possibility is that the residue may be critical for keeping Pyc in the tetrameric state, either directly by being involved in the subunit interaction surface, or indirectly by somehow maintaining the conformation necessary for the subunit interactions. Either way the residue would be essential for enzyme activity, as Pyc is only active in the tetrameric state (see Chapter 1, 1.5.4).

The chemical modification studies of Palacian and Neet (1970) with chicken liver enzyme would appear to be consistent with this proposal. Using ultra-centrifugation these workers found that inactivation of CLPyc by a range of different SH reagents was accompanied by dissociation of the 15S enzyme. This was evident by an increase in monomers (7S) with the reagents that attach a charged group to cysteine residues (cysteine, iodoacetate, DTNB, and *p*-hydroxymercuribenzoate ie PHMB), or the formation of high molecular weight aggregates with the reagents which attach uncharged groups (iodoacetamide, NEM, and 4,4'-dithiodipyridine). Palacian and Neet (1970) suggested that

the different association states caused by the two classes of reagents were consistent with the subunit contact surface being predominantly non-polar, a view which had earlier been proposed to explain the dissociation and inactivation of CLPyc caused by low temperatures (Scrutton and Utter, 1965; Irias *et al.*, 1969).

Similarly, the inactivation of rat Acc by PHMB was also found to be accompanied by dissociation into monomers, as indicated by sucrose density gradient centrifugation (Ahmad *et al.*, 1984). In view of the fact that vertebrate Accs were known to oscillate between the essentially inactive protomeric form, and the active polymeric form (reviewed in Wood and Barden, 1977), these workers concluded that the SH modification had "locked" the enzyme in the inactive protomeric state. This model also explained their observation that citrate, which activates rat Acc by inducing polymerisation, provided protection against inactivation but was unable to reverse the effects of SH group modification. In contrast, none of the substrates for the biotin carboxylation half reaction (ATP, ADP, Pi) or the transcarboxylation reaction (acetyl-CoA) were able to protect against PHMB inactivation, leading them to conclude that the inactivation was not caused by the modification of an active site reside.

However, the inability of the substrates of the biotin carboxylation reaction to protect either Acc (Ahmad *et al.*, 1984) or Pyc (Hudson *et al.*, 1975; Goss, 1978; and Bagley, 1982) against inactivation by various SH reagents does not entirely exclude the possibility that the modified cysteine residue has an essential catalytic role in the biotin carboxylation reaction. The decrease in stability of the quaternary structure of the enzymes may be a secondary rather than primary effect, and the cysteine residue may be in the active site without being sterically close enough to the binding sites for ATP or bicarbonate for these molecules provide protection. Indeed, this proposal would be consistent with a number of other studies implicating a cysteine residue as being involved in base catalysis of the enolisation of biotin, in the form of a cysteine-lysine ion pair (see section 1.4.1.3). Briefly, this has been suggested on the basis of: pH reaction profiles for sheep liver Pyc (SLPyc; Attwood and Cleland, 1986), pH profiles and isotope effects reported for the Bc subunit of *E. coli* Acc (Tipton and Cleland, 1988 a & b), cross-linking of an ATP analogue (oATP; 2',3'-dialdehyde derivative of ATP) to a lysine residue in SLPyc (Easterbrook-Smith *et al.*, 1976), and *o*-phalaldehyde cross-linking of a pair of adjacent (separated by < 3 angstroms)

cysteine and lysine residues in or near the 1st partial reaction site of CLPyc (Wernberg and Ash, 1993).

On the basis of sequence homology, there are two cysteine residues which would appear to be attractive candidates for the essential cysteine residue. These are residues 249 and 260 in the ATP domain of Pyc1 (see Fig. 4.8). Cysteine 249 in Pyc1 is contained within a region of high homology between ATP binding biotin carboxylases, and this residue is present in all the biotin carboxylases (see Fig. 4.8). Furthermore, inspection of model for the 2.4 angstrom crystal structure of biotin carboxylase (Bc, subunit of *E. coli* Acc; Waldrop *et al.*, 1994b) revealed that the corresponding cysteine residue to 249 in Pyc1 (230 in Bc, see Fig. 4.8) is in fact adjacent to the residues mentioned as making up the "putative" active site, including lysine 238.

Similarly, cysteine 360 (Pyc1) is contained in all Pyc, all Pcc, and 2/3 of Acc enzymes sequenced to date (see Fig. 4.8). In addition, the corresponding cysteine in Bc (237) is adjacent to R238, which is one of the residues believed to surround the "active site" phosphate ion in crystal structure of Bc (Waldrop, *et al.*, 1994b).

Therefore, Pyc1 resides 249 and 260 would appear to be reasonable candidates for a cysteine residue participating in the chemistry of the 1st partial reaction. Alternatively, modification of one of these residues may lead to conformational changes inactivating the enzyme, possibly due to dissociation. Clearly, before an accurate understanding of the role of these cysteine residues can be reached, the X-ray structure of biotin carboxylase will need to be solved at a higher resolution, and the biotinylated (rather than apo form) enzyme will need to be co-crystallised with the other substrates (or suitable analogues) of the biotin carboxylation reaction. Ultimately though, a satisfactory understanding of the observations made during the course of the various cysteine modification studies with pyruvate carboxylase will require that the three-dimensional structure of pyruvate carboxylase be solved.

CHAPTER 8

GENERAL DISCUSSION AND CONCLUSIONS

8.1 INTRODUCTION

One of the most fundamental objectives in the field of biochemistry is to be able to understand the functions of biological molecules. While much can be learnt about these functions using biochemical approaches, ultimately a comprehensive understanding of a given molecule requires a knowledge of its structure. The structural and functional attributes of the molecule can then be related in an effort to understand the mechanism of action and cellular role of the molecule.

In the case of complex allosteric enzymes such as pyruvate carboxylase, correctly establishing the structure / function relationships of the molecule can be particularly challenging. In addition to the complicated chemistry of the two partial reactions, understanding the workings of Pyc will require the characterisation of the precise enzyme-substrate interactions involved in catalysing each half reaction, the mechanism of carboxyl transfer between the two subsites, and the conformational changes which are responsible for the progress and regulation of the overall reaction.

Over the years much progress has been made towards understanding the mechanism of the Pyc reaction, some of the basic structural characteristics on the enzyme have been identified by protein chemistry techniques and electron microscopy, and chemical modification studies have implied that certain amino acids are involved in catalysis. However, there had been no structure / function studies investigating the role of specific amino acid residues using the more precise technique of site-directed mutagenesis. Therefore the work presented in this thesis was directed towards the long term goal of understanding the precise functional roles of specific key residues in the mechanism of action of pyruvate carboxylase.

More specifically, the primary research objectives of my project were to: a) clone and sequence the *PYC2* gene from *Saccharomyces cerevisiae*, b) disrupt the *PYC2* gene in a *pyc1* null so as to construct a strain with no Pyc activity suitable as a host for the expression of mutant Pyc molecules, and c) use site-directed mutagenesis to investigate the hypothesis that Pyc contains an essential cysteine residue which functions as a catalytic base in the 2nd partial reaction.

8.2 GENERAL DISCUSSION AND CONCLUSIONS

8.2.1 Chromosomal localisation of the yeast *PYC* genes

At the outset of this study the gene knockout experiments conducted by Dr. M. E. Walker in our laboratory had suggested that there was a second *PYC* isozyme in *Saccharomyces cerevisiae* (Walker *et al.*, 1991). The chromosmal localisation experiments described in Chapter 1 provided further confirmation of these results.

Using the technique of hybridisation (with a *PYC1* probe) of chromosome blots obtained from agarose gels containing yeast chromosomes separated out by pulse field electrophoresis, the *PYC1* gene was localised to chromosome VII, and *PYC2* to chromosome II (Walker *et al.*, 1991).

After isolating the *PYC2* gene, each of the *PYC* genes were further localised by hybridising gene specific probes to a set of prime λ-clone grid filters (kindly supplied by the M. V. Olson lab.) containing an ordered set of λ-clones of known location on the physical map of *S. cerevisiae* (Riles *et al.*, 1993). The *PYC1* gene was found to be on the left arm of chromosome VII, at or very close to the *RAD6* locus, while *PYC2* was found to be approximately 15 kb distal of the *DUR1,2* locus on the right arm of chromosome II. This localisation of *PYC2* agreed with the results of Stucka *et al.* (1991), and has since been confirmed by the complete sequencing of chromosome II (Feldman *et al.*, 1994).

8.2.2 The sequence of *PYC2*

Although it would have been possible to construct a *pyc2* / *pyc1* double null mutant host strain for the expression of mutant Pyc molecules without having sequenced *PYC2*, the isolated *PYC2* gene was sequenced, as the only *PYC* gene which had been sequenced at that stage was yeast *PYC1* (Lim *et al.*, 1988).

The DNA sequence of *PYC2* described in Chapter 4 revealed that within the open reading frame *PYC1* and *PYC2* are very similar, having an identity of 85.4%. However, there was no appreciable homology between the non-coding sequences, indicating that *PYC2* was indeed a separate non-allelic gene coding for a different isozyme of pyruvate carboxylase.

At the amino acid level Pyc2 has one additional amino acid at the N and C-termini giving a total of 1180 residues, and 92 internal sequence differences. The sequence differences are evenly dispersed throughout the majority of the protein with the exception of the greater variation in the C-terminal ~200 residues, the last 79 of which comprise the biotin domain which had a sequence identity to Pyc1 of only 67%. This low sequence identity is typical of biotin domains, although they all share the same pattern of hydrophobic key folding residues as those identified in the structures of the functionally similar lipoyl domains (Brocklehurst and Perham, 1993). This led to the prediction that lipoyl and biotin domains have very similar structures (Brocklehurst and Perham, 1993), a prediction which has since been confirmed by the NMR structure for the biotin domain from *E. coli* Acc (Brocklehurst *et al.*, 1995).

Regarding the specific amino acid differences between Pyc1 and Pyc2, it was not possible to make any reliable predictions concerning the effects of any of the 95 amino acid differences, as at this stage a detailed comparison of the kinetic and other characteristics of these isozymes have not been made. The preliminary kinetic comparisons of Stucka *et al.* (1991) suggested that there were no notable differences between the isozymes, while work with *pyc* null mutant strains in our laboratory has indicated that Pyc1 has approximately twice the specific activity of Pyc2 (M. Walker pers. comm.).

During the course of sequencing *PYC2*, the complete sequence of this gene was published by another group (Stucka *et al.*, 1991). Comparisons between the two *PYC2* sequences revealed that there were 36 differences within open reading frame (ORF), 35 transversions resulting in 12 predicted amino acid differences, and one single base insertion in the published sequence (relative to my data) close to the 3' end of the ORF. The insertion changed Q1178 into a P, and caused read-through of the TAA termination codon resulting in 5 additional amino acids at the C-terminus of the sequence reported by Stucka *et al.* (1991).

As the majority of sequence differences between the two sets of data were silent "changes", these variations may well be due to natural polymorphisms.

Confirmation of the predicted amino acid sequence for the biotin domain expressed from the *PYC2* gene I isolated was obtained by N-terminal sequencing and mass spectrometric analysis of an 86 amino acid peptide comprising the C-terminus of Pyc2 (described in Chapter 5). The expressed biotinylated peptide contained the correct N-terminal sequence (10 residues) and had a calculated Mw of 9234.9, which was within 0.1% of the theoretical value of 9235.7 and 620.6 atomic mass units less than predicted value for the peptide corresponding to the published sequence (Stucka *et al.*, 1991). Furthermore, having developed a convenient method to purify this peptide (see Chapter 5), it should be possible to obtain sufficient material for three-dimensional structure determination by NMR.

8.2.3 Effect of the sequence differences in the Pyc2 biotin domain on biotinylation *in vivo*

As the precise structural determinants which are responsible for the correct recognition and biotinylation of biotin domains by biotin ligase enzymes had not been determined (and are still yet to be determined), the effects of the sequence differences in the Pyc2 biotin domain on biotinylation were investigated. This was achieved by measuring the relative extent of *in vivo* biotinylation for control Pyc1 biotin domain peptides, and Pyc2 C-terminal peptides containing the alternative sequences with respect to the 2 differences in the Pyc2 biotin domain (between my data and the report of Stucka *et al.* 1991). More specifically, the peptides were expressed in the presence of excess biotin in the growth media, in a strain of *E. coli* which overexpresses the *E. coli* biotin ligase enzyme upon induction by arabinose. The relative extent of biotinylation was then determined by quantitating and comparing the amounts of each expressed biotin domain peptide with the amount of biotinylated material detected on Western blots by avidin alkaline phosphatase.

Using this technique the 5 amino acid extension to the C-terminus of Pyc2 present in the sequence of Stucka *et al.* (1991) was found to produce a noticeable increase in the extent of biotinylation. Similarly, the 104 amino acid C-terminal biotin domain peptide from Pyc1 was biotinylated ~6 times more efficiently than the N-terminally shorter 85 amino acid Pyc1 peptide. Therefore, these results suggest that the extensions to the N and C-termini of the core biotin domain in some way improve the ability of biotin ligase enzymes to correctly

recognise and modify these domains. This was unexpected as the extensions were to sequences N-terminal (in the case of the Pyc1 peptides) and C-terminal (in the case of the Pyc2 peptides) of the region predicted to contain the entire folded biotin domain. However, the similar effects of these extensions may be explained by the fact that the N and C-termini of lipoyl domains (Dardel *et al.*, 1993; Green *et al.*, 1995) and the BCCP biotin domain (Brocklehurst *et al.*, 1995) are known to be adjacent, located at the opposite end of the three-dimensional structures to the β-hairpin loop containing the recipient lysine. As the recipient lysine residue is at the other end of the domain, one possibility is that the increase in biotinylation produced by the N and C-terminal extensions may be due to a stabilisation of the protein folding at this end of the molecule, rather than any direct interactions with biotin ligase.

8.2.4 Construction of a *pyc1 / pyc2* double null strain as a host for the expression of mutant Pyc enzymes

In order to produce a yeast strain suitable for the expression of mutant Pyc enzymes in structure-function studies (DM18, *pyc1 / pyc2* double null mutant), it was necessary to use gene knock-out to disrupt the *PYC2* gene. Homologous recombination was used to construct *PYC1 / pyc2* single null mutant strains from the wild-type strains DBY746 and DBY747, and a *pyc1 / pyc2* double null mutants from strain MW21.3 with the *HIS3* and *TRP1* marker genes (described in Chapter 6). Confirmation of gene knockout was obtained by enzyme assays, growth profiles and Western blots.

8.2.5 Mutagenesis of the cysteine residues in the pyruvate domain

To investigate the hypothesis that the 2nd partial reaction of Pyc involves an essential cysteine residue acting as a base catalyst, each of the 4 cysteine residues in the pyruvate domain of yeast Pyc1 were changed to serine residues, and the mutant enzymes were expressed at high and low levels in strain DM18 (see Chapter 7). On the basis of enzyme assays and their ability to complement the aspartate auxotrophy of strain DM18, none of the cysteine residues were found to be essential for Pyc activity. Thus it was concluded that the 2nd partial reaction does not involve a cysteine residue acting as a catalytic base. Instead, the

transcarboxylation reaction may proceed by a mechanism analogous to pH independent mechanism proposed by O'Keefe and Knowles (1986; Chapter 1 Scheme V) for the enzyme transcarboxylase, or by the mechanism proposed for Pyc by Attwood (1995; Chapter 1 Scheme IV) with the enolisation of biotin being catalysed by a low-barrier hydrogen bond rather than a cysteine residue.

8.3 FUTURE WORK

8.3.1 The Pyc isozymes

From the results presented in this thesis it is clear that *Saccharomyces cerevisiae* contains two very similar isozymes of the enzyme pyruvate carboxylase. Furthermore, immunocytochemical localisation of Pyc in whole cells using polyclonal antibodies and protein A-gold revealed that Pyc2 (Walker *et al.*, 1991) and Pyc1 (Rohde *et al.*, 1991) are both found in the cytosol. This raises the question as to whether these two isozymes perform the same metabolic functions in *Saccharomyces cerevisiae*. Brewster *et al.*, (1994) found that both isozymes are expressed in wild-type yeast and are independently regulated. Pyc1 was found to be the predominant isozyme, expressed at relatively constant level throughout growth, while expression of Pyc2 was activated during early log phase and repressed during the remainder of the growth period. Also, expression of Pyc1 was found to be approximately 10x that of Pyc2 during growth on ethanol. Therefore, these preliminary regulation studies suggest that the two isozymes may well be performing different metabolic roles. However, more gene regulation studies, and kinetic studies on the purified enzymes will need to be performed before the differing metabolic roles of these two isozymes can be established.

8.3.2 Recognition and biotinylation of biotin domains

With the recent solving of the three-dimensional structure of the BCCP biotin domain (Brocklehurst *et al.*, 1995) it should be possible to investigate the recognition determinants of biotin domains by using molecular modelling software to perform "docking" between this

acceptor domain structure, and the structure of the *E. coli* biotin ligase enzyme (Wilson *et al.*, 1992). In conjunction with sequence homology data, this approach should enable more informed choices to be made when selecting potentially important residues responsible for the correct recognition of biotin domains. Site-directed mutagenesis could then be used to investigate the effect of changing these residues on the extent of biotinylation.

Alternatively, one of a number of random mutagenesis strategies could be used to mutate the BCCP biotin domain (and / or a Pyc biotin domain once its structure is determined). The mutant biotin domain sequences could then used to create a library of clones which could then be screened for the ability of the encoded domain to be biotinylated.

One way in which this screening could be achieved would be to fuse the coding sequences of the library of mutant clones to the 5' end of the gene encoding the filamentous phage coat III protein (reviewed in Cesareni, 1992). Phage displaying biotinylated fusion proteins could be selected by affinity chromatography with immobilised monomeric avidin, and mutant domains with different capacities to accept biotin could be selected by assaying the biotin dependence of a biotin auxotrophic host strain of *E. coli* (CY486) containing the phagemid which over-expresses that domain. Clones expressing a biotin domain which can be biotinylated efficiently would cause the biotin auxotroph to require higher concentrations of exogenous biotin for growth than those expressing domains which are biotinylated more poorly (Cronan, 1990).

8.3.3 Key residues in the structure and reaction mechanism of Pyc

The detailed sequence comparisons presented in Chapter 4 have highlighted a number of residues and motifs which may prove to be important for the function of this enzyme. Specifically, the sequence comparisons revealed that within a stretch of 11 residues in the ATP domain which are essentially unique to biotin carboxylases (248 - 258 of Pyc2) there is a cysteine residue (250 of Pyc1) which is present in all the biotin carboxylases. Similarly, all Pyc and Pcc enzymes, and 2/3 of the Acc enzymes contain a cysteine residue in the position corresponding to cysteine 361 in the ATP domain of Pyc2. Hence, it would be of interest to

determine whether either of these residues is the cysteine involved as the base catalyst of biotin enolisation, as proposed by mechanistic studies of Tipton and Cleland (1988 a & b).

The comparisons also revealed that all biotin domain carboxylases contain a QVEH motif in the ATP domain (315 - 318 of Pyc2), while in its place the Cps enzymes contain a SRSS motif. As Cps enzymes also carry out the formation of carboxy-phosphate (Wimmer *et al*, 1979) but do not carboxylate biotin, site-directed mutagenesis could be used to investigate the hypothesis that these residues may be involved in the carboxylation of biotin by carboxy-phosphate.

Site-directed mutagenesis could also be used to determine whether the HXHXH motif in the pyruvate domain forms part of the metal binding site, and whether tryptophan 619 is involved in pyruvate binding as inferred by the 2,4-dinitrophenylsulfenyl chloride (DNPS-Cl) modification studies with transcarboxylase Kumar *et al.* (1988b).

8.3.4 The three-dimensional structure

While much can be learnt about the structure-function relationships of pyruvate carboxylase by mutagenesis studies investigating predictions based on sequence comparisons, structure-function studies would be greatly facilitated by a knowledge of the three-dimensional structure. To date however efforts to crystallise Pyc have been essentially unsuccessful. Crystals of pyruvate carboxylase have been obtained for the chicken liver enzyme, but they were not of sufficient quality or quantity to permit structure determination by X-ray diffraction (Frey and Utter, 1977; Snoswell, 1982). In the case of yeast Pyc, Dr. F. Lim (1988) was unable to obtain any crystals. However it is possible that the inability of the purified yeast Pyc to crystallise may have been be due to the fact that the Pyc "enzyme" used in the crystallisation was actually a mixture of Pyc1 and Pyc2. Therefore, future efforts to crystallise this enzyme might best be performed by expressing and purifying the individual isozymes. If this approach is unsuccessful the pyruvate and ATP domains of the protein could be expressed using an approach similar to that used for the biotin domain (see Chapter 5), and crystallisation could be attempted with the individual domains, and various combinations of each of the domains.

With the aid of a three-dimensional structure it should then be possible to design experiments to investigate the many unanswered questions regarding the mechanism of action of pyruvate carboxylase. For example: the structural role of the Zn^{2+} (yeast Pyc) or Mn^{2+} (mammalian Pycs) ion, the role of magnesium ions in stabilising the biotin group in the 1st partial reaction site, the instability of the carboxy-biotin complex in the 2nd partial reaction site, the activation by acetyl-CoA, the signalling of translocation by pyruvate, and the interactions and conformational changes involved in the process of translocation.

PRESENTATIONS AND PUBLICATIONS

Brewster, N. K., Val, D. L., Walker, M. E., and Wallace, J. C. (1994) Regulation of pyruvate carboxylase isozyme (*PYC1, PYC2*) gene expression in *Saccharomyces* cerevisiae during fermentative and non fermentative growth. *Arch. Biochem. Biophys.* **310** 1-10

Val, D. L., Chapman-Smith, A., Walker, M. E., Cronan, J. E. Jr. (1995) Polymorphism of the yeast pyruvate carboxylase 2 gene and protein: effects on protein biotinylation. *Biochem. J.* (manuscript submitted)

Val, D. L., Chapman-Smith, A., Walker, M. E., Cronan, J. E. Jr, and Wallace, J. C. (1995) Yeast pyruvate carboxylases: sequences and biotinylation of variants of biotin carrier domains. *Proc. Lorne Conference on Protein Structure and Function.*

Val, D. L., Walker, M. E., Wallace, J. C. (1992) A second isozyme of yeast pyruvate carboxylase. *Proc. Lorne Conference on Protein Structure and Function.* pos C44.

Walker, M. E., Val, D. L., Wallace, J. C. (1991) Yeast pyruvate carboxylase: identification of two genes encoding isoenzymes. *Biochem. Biophys. Res. Comm.* **76** 1210-1217

Wallace, J. C., Val, D. L., Walker, M. E., Chapman-Smith, A., Cronan, J. E. Jr, Morris, T., Brocklehurst, S. M., Kalia, Y., Perham, R. N. (1994) Biotin-dependent carboxylases: structure and function. *Proc. Aus. Soc. Biochem. Mol. Biol.* SYM-12-2.

REFERENCES

Ahmad, F., Lygre, D.G., Jacobson, B.E. & Wood, H.G. (1972) Transcarboxylase: XII Identification of the metal-containing subunits of transcarboxylase and stability of the binding. *J. Biol. Chem.* **247**, 6299-6305

Ahmad, P.M., Gupta, S., Barden, R.E. & Ahmad, F. (1984) Inhibitory effects of sulfhydryl reagents in acetyl-CoA carboxylase from rat mammary gland. *Biochem. Biophys. Acta.* **789**, 152-158

Al-Feel, W., Chirala, S.S. & Wakil, S.J. (1992) Cloning of the yeast *FAS3* gene and primary structure of yeast acetyl-CoA carboxylase. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 4534-4538

Allen, S.H.G., Treble, D.H., Peterson, D.D. & Botti, C.F.J. (1984) *Fed. Proc., Fed. Am. Soc. Exp. Biol.* **43**, 1781

Ammerer, G. (1983) Expression of genes in yeast using the *ADH1* promoter. *Methods. Enzymol.* **101**, 192-201

Ashman, L.K., Keech, D.B., Wallace, J.C. & Nielsen, J. (1972) Sheep kidney pyruvate carboxylase. Studies on its activation by acetyl-CoA and characterisation of its acetyl coenzyme independent reaction. *J. Bio. Chem.* **247**, 5818-5824

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

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

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

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

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

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

Attwood, P.V. & Wallace, J.C. (1986) The carboxybiotin complex of chicken liver pyruvate carboxylase. *Biochem. J.* **235**, 359-364

Attwood, P.V. & Graneri, B.D.L.A. (1991) Pyruvate carboxylase catalysis of phosphate transfer between carbamoyl phosphate and ADP. *Biochem. J.* **273**, 443-448

Attwood, P.V. & Graneri, B.D.L.A. (1992) Bicarbonate-dependent ATP clevage catalysed by pyruvate carboxylase in the absence of pyruvate. *Biochem. J.* **287**, 1011-1017

Attwood, P.V., Johannssen, W., Chapman-Smith, A. & Wallace, J.C. (1993) The existence of multiple tetrameric conformers of chicken liver pyruvate carboxylase and their roles in dilution inactivation. *Biochem. J.* **290**, 583-590

Attwood, P.V. (1993) Locus of action of acetyl-CoA in the biotin-carboxylation reaction of pyruvate carboxylase. *Biochemistry* **32**, 12736-12742

Attwood, P.V. (1995) The structure and mechanism of action of pyruvate carboxylase. *Biochem. Int.* (in press)

Bai, D.-H., Moon, T.-W., Lopes-Casillas, F., Andrews, P.C. & Kim, K.-H. (1989) Analysis of the biotin-binding site on acetyl-CoA carboxylase from rat. *Eur. J. Biochem.* **182**, 239-245

Bais, R. & Keech, D.B. (1972) The magnesium ion (Mg^{2+}) activation of sheep kidney pyruvate carboxylase. *J. Biol. Chem.* **247**, 3255-3261

Bais, R. (1974) Pyruvate carboxylase. Physical and kinetic studies. PhD. Thesis, University of Adelaide, South Australia, Australia.

Bagley, C.J. (1982) Reactive cysteines of pyruvate carboxylase. B.Sc. (Hons), University of Adelaide, South Australia, Australia.

Barden, R.E., Fung, C.-H., Utter, M.F. & Scrutton, M.C. (1972) Pyruvate carboxylase from chicken liver: steady state kinetic studies indicate a "two-site" ping-pong mechanism. *J. Biol. Chem.* **247**, 1323-1333

Barden, R.E., Taylor, B. L., Isohashi, F., Frey, W. H., Zander, G. & Utter, M. F. (1975) Structural properties of pyruvate carboxylases from chicken liver and other sources. *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4308-4312

Barker, D.F. & Campbell, A.M. (1981) Genetic and biochemical characterization of the *bir*A gene and its product: evidence for a direct role of biotin holoenzyme synthetase in repression of the biotin operon in *Escherichia coli*. *J. Mol. Biol.* **146**, 469-492

Barritt, G.J. (1985) Regulation of enzymatic activity. In Pyruvate Carboxylase, ed. Keech, D.B. & Wallace, J.C. 179-212. CRC Press Inc., Boca Raton, Florida.

Becker, D.M. & Guarente, L. (1991) High-efficiency transformation of yeast by electroporation. *Meth. Enzymol.* **194**, 182-186

Bennetzen, J.L. & Hall, B.B. (1982) Codon selection in yeast. *J. Biol. Chem.* **257**, 3026-3031

Best, E.A. & Knauf, V.C. (1993) Organization and nucleotide sequences of the genes encoding the biotin carboxyl carrier protein and biotin carboxylase protein of *Pseudomonas aeroginosa* acetyl coenzyme A carboxylase. *J. Bacteriol.* **175**, 6881-6889

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

Bloom, S.J. & Johnson, M.J. (1962) The pyruvate carboxylase of *Aspergillus niger*. *J. Biol. Chem.* **237**, 2718-2720

Booker, G.W. (1990) Pyruvate carboxylase: relating the structure to its function. PhD. Thesis, The University of Adelaide, South Australia, Australia.

Bordo, D. & Argos, P. (1991) Suggestions for "safe" residue substitutions in site-directed mutagenesis. *J. Mol. Biol.* **217**, 721-729

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

Brewster, N.K., Val, D.L., Walker, M.E. & Wallace, J.C. (1994) Regulation of pyruvate carboxylase isozyme (*PYC1,PYC2*) gene expression in *Saccharomyces cerevisiae* during fermentative and nonfermentative growth. *Arch. Biochem. Biophys.* **310**, 1-10

Brocklehurst, S.M. & Perham, R.N. (1993) Prediction of the three-dimensional structures of the biotinylated domain from yeast pyruvate carboxylase and of the lipoylated H-protein from the pea leaf glycine cleavage system: a new automated method for the prediction of protein tertiary structure. *Protein Sci.* **2**, 626-639

Brocklehurst, S.M., Kalia, Y. N., Kulman, J. D., Chapman-Smith, A., Cronan, J. E. Jr., Wallace, J. C. & Perham, R. N. (1995) Three-dimensional structure of a biotinyl domain from *Escherichia coli* Acetyl-CoA carboxylase: relationship to lipoyl domains. (manuscript in preparation).

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

Buck, M.A., Olah, T.A., Weitzmann, C.J. & Cooperman, B.S. (1989) Protein estimation by the product of integrated peak area and flow rate. *Anal. Biochem.* **182**, 295-299

Carle, G.F. & Olson, M.V. (1984) Separation of chromosomal DNA molecules from yeast by orthogonal-field-alternation gel electrophoresis. *Nucleic Acids Res.* **12**, 5647-5664

Carle, G.F. & Olson, M.V. (1985) An electrophoretic karyotype for yeast. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3756-3760

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

Cassady, A.I. (1987) Pyruvate carboxylase : a molecular biological study. PhD thesis, University of Adelaide, Adelaide, South Australia, Australia.

Cazzulo, J.J. & Stoppani, A.O.M. (1967) Purification and properties of pyruvate carboxylase from baker's yeast. *Arch. Biochem. Biophys.* **121**, 596-608

Cazzulo, J.J., Claisse, L.M. & Stoppani, A.O.M. (1968) Carboxylase levels and carbon dioxide fixation in bakers yeast. *J. Biol. Chem.* **96**, 623-628

Cazzulo, J.J. & Stoppani, A.O.M. (1968) The regulation of yeast pyruvate carboxylase by acetyl-coenzyme A and L-aspartate. *Arch. Biochem. Biophys.* **127**, 563-567

Cesareni, G. (1992) Peptide display on filamentous phage cupsids: a new powerful tool to study protein-ligand interaction. *FEBS Lett.* **307**, 66-70

Chapman-Smith, A., Booker, G.W., Clements, P.R., Wallace, J.C. & Keech, D.B. (1991) Further studies on the localisation of the reactive lysyl residue of pyruvate carboxylase. *Biochem. J.* **276**, 759-764

Chapman-Smith, A., Turner, D.L., Cronan, J.E.Jr. & Wallace, J.C. (1994) Expression, biotinylation and purification of a biotin domain peptide from the biotin carboxyl carrier protein of *E. coli* acetyl-CoA carboxylase. *Biochem. J.* **302**, 831-837

Chen, E.J. & Seeburg, P.H. (1985) Supercoil sequencing: a fast simple method for sequencing plasmid DNA. *DNA* **4**, 165-170

Cleland, W.W. (1992) Low-barrier hydrogen bonds and low fractionation factor bases in enzymic reactions. *Biochemistry* **31**, 317-319

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

Cohen, N.D., Utter, M.F., Wrigley, N.G. & Barret, A.N. (1979b) Quaternary structure of yeast pyruvate carboxylase: biochemical and electron microscope studies. *Biochemistry* **18**, 2197-2203

Cohen, N.D., Beegan, H., Utter, M.F. & Wrigley, N.G. (1979c) A re-examination of the electron microscopic appearance of pyruvate carboxylase from chicken liver. *J. Biol. Chem.* **254**, 1740-1747

Cooper, T.G. & Benedict, C.R. (1966) The participation of acetyl-CoA in pyruvate carboxylase. *Biochem. Biophys. Res. Commun.* **22**, 285-290

Cronan, J.E.Jr. (1989) The *E. coli bio* operon: transcriptional repression by an essential protein modification enzyme. *Cell* **58**, 427-429

Cronan, J.E.Jr. (1990) Biotinylation of proteins *in vivo*. *J. Biol. Chem.* **265**, 10327-10333

Cryer, D.R., Ecclestall, R. & Marmur, J. (1975) Isolation of yeast DNA. *Methods Cell Biol.* **12**, 39-44

Dardel, F., Davis, A.L., Laue, E.D. & Perham, R.N. (1993) The three-dimensional structure of the lipoyl domain from *Bacillus stearothermophilus* pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **229**, 1037-1048

De Titta, G.T., Edmonds, J.W., Stallings, W. & Donohue, J. (1976) Molecular structure of biotin. Results of two independent crystal structure investigations. *J. Am. Chem. Soc.* **98**, 1920-1926

Dimroth, P. & Thomer, A. (1992) The sodium ion pumping oxaloacetate decarboxylase of *Klebsiella pneumoniae*. *FEBS. Letts.* **300**, 67-70

Dower, W.J., Miller, J.G. & Ragsdale, C. (1988) High efficiency transformation of *E. coli* by high voltage electroporation. *Nuc. Acids Res.* **16**, 6127-6145

Dujon, B., *et. al.* (1994) Sequence of yeast chromosome XI. *Nature* **369**, 371-378

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

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

Eibel, H. & Philippsen, P. (1983) Identification of the cloned *S. cerevisiae LYS2* gene by an integrative transformation approach. *Mol. Gen. Genet.* **191**, 66-73

Elgar, G. & Schofield, J.P. (1992) Carbamoyl phosphate synthetase (CPase) in the *PYR1-3* multigene of *Dictyostelium discoidium*. *DNA Seq.* **2**, 219-226

Fall, R.R. & Vagelos, P.R. (1973) Acetyl coenzyme A carboxylase: proteolytic modification of biotin carboxyl carrier protein. *J. Biol. Chem.* **248**, 2078-2088

Feinberg, A.P. & Vogelstein, B. (1983) A technique for radiolabelling DNA restriction fragments to a high specific activity. *Anal. Biochem.* **132**, 6-13

Feldman, H., *et al.* (1994) Complete DNA sequence of yeast chromosome II. *EMBO J.* **13**, 5795-5809

Freund, J.N. & Jarry, B.P. (1987) The rudimentary gene of *Drosophila melanogaster* encodes four enzymic functions. *J. Mol. Biol.* **193**, 1-13

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

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

Froede, H.C., Geraci, G. & Mansour, T.E. (1968) Studies on heart phosphofructokinase: thiol groups and their relationship to activity. *J. Biol. Chem.* **243**, 6021-6029

Fujiwara, K., Okamura-Ikeda, K. & Motokawa, Y. (1991) Lipoylation of H-protein of the glycine cleavage system. *FEBS Lett.* **293**, 115-118

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

Fung, C.H., Mildvan, A.S. & Leigh, J.S Jr. (1974) Electron and nuclear magnetic studies of the interaction of puruvate with transcarboxylase. *Biochemistry* **13**, 1160-1169

Geitz, R.D. & Sugino, A. (1988) New yeast-*Escherichia coli* shuttle vectors with *in vitro* mutagenized genes lacking six-base pair restriction sites. *Gene* **74**, 527-534

Genbauffe, F.S. & Cooper, T.G. (1991) The urea anidolyase (*DUR1,2*) gene of *Saccharomyces cerevisiae*. *DNA sequence* **2**, 19-32

Glusker, J.P. (1991) Structural aspects of metal liganding. *Adv. Pro. Chem.* **42**, 1-76

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

Gornicki, P., Scappino, L.A. & Haselkorn, R. (1993) Genes for two subunits of acetyl-CoA carboxylase of *Anabaena* sp. strain PCC 7120: biotin carboxylase and carboxyl carrier protein. *J. Bacteriol.* **175**, 5268-5272

Gornicki, P., Podkowinski, J., Scappino, L.A., DiMaio, J., Ward, E. & Haselkorn, R. (1994) Wheat acetyl-CoA carboxylase: cDNA and protein structure. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6860-6864

Goss, N.H. (1978) Pyruvate carboxylase: physico-chemical aspects. PhD. Thesis, University of Adelaide, Adelaide, South Australia, Australia.

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

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

Green, N.M., Valentine, R.C., Wrigley, N.G., Ahmad, F., Jacobson, B. & Wood, H.G. (1972) Transcarboxylase XI: electron microscopy and subunits structure. *J. Biol. Chem.* **247**, 6284-6298

Green, J.F.D., Laue, E.D., Perham, R.N., Ali, S.T. & Guest, J.R. (1995) *J. Mol. Biol.* (in press)

Grunstein, M. & Hogness, D. (1975) Colony hybridisation: a method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3961-3965

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

Guichhait, R.B., Polakis, S.E., Hollis, D., Fenselau, C. & Lane, M.D. (1974) Acetyl coenzyme A carboxylase systems of *Escherichia coli*. Site of carboxylation of biotin and enzymatic reactivity of 1'N-(ureido)carboxybiotin derivatives. *J. Biol. Chem.* **249**, 6646-6656

HA, J., Daniel, S., Kong, I-S., Park, C-K., Tae, H-J. & Kim, K-H. (1994) Cloning of the human acetyl-CoA carboxylase cDNA. *Eur. J. Biochem.* **219**, 297-306

Hale, G., Wallis, N.G. & Perham, R.N. (1992) Interaction of avidin with the lipoyl domains in the pyruvate dehydrogenase multienzyme complex: three-dimensional location and similarity to biotin domains in carboxylases. *Proc. R. Soc. Lond. B.* **248**, 247-253

Hanahan, D., Jessee, J. & Bloom, F.R. (1991) Plasmid transformations of *Escherichia coli* and other bacteria. *Methods Enzymol.* **204**, 63-113

Haraguchi, Y., Uchino, T., Takiguchi, M., Endo, F., Mori, M. & Matsuda, I. (1991) Cloning and sequence of a cDNA encoding human carbamoyl phosphate synthetase I: molecular analysis of hyper-ammonemia. *Gene* **107,** 335-340

Hegarty, A.F., Bruice, J.C. & Benkovic, S.J. (1969) Biotin and the nucleophilicity of 2-methoxy-2-imidizoline toward the sp2 carbonyl carbon. *Chem. Commun* 1173-1174

Henderson, J.R. & Lamonds, M.R. (1966) Effects of D_2O on citric acid cycle and pyruvate carboxylase formation by *Aspergillus niger*. *Arch. Biochem. Biophys.* **115**, 187-191

Henikoff, S. (1987) Unidirectional digestion with Exonuclease III in DNA sequence analysis. *Methods Enzymol.* **155**, 156-165

Hirel, P., Schmitter, D.P., Fayat, G. & Blanquet, S. (1989) Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 8247-8251

Hoekstra, M.F. (1988) Lid lysastes: an economical and rapid method for plasmid analysis. *BioTechniques* **6**, 929-932

Hoffman, N.E., Pichersky, E. & Cashmore, A.R. (1987) A tomato cDNA encoding a biotin binding protein. *Nuc. Acids Res.* **15**, 3928

Hudson, P.J. (1974) Bromopyruvate labelling of pyruvate carboxylase. B.Sc. (Hons) Thesis, Adelaide University, Adelaide, South Australia, Australia.

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

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

Ito, H., Fukada, Y., Murata, K. & Kimura, A. (1983) Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* **153**, 163-168

Jiang, W. & Bond, J.S. (1992) Families of metalloendopeptidases and their relationships. *FEBS. Lett.* **312**, 110-114

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

Kaziro, Y., Hass, L.F., Boyer, P.D. & Ochoa, S. (1962) Mechanism of the propionyl carboxylase reaction II. Isotopic exchange and tracer experiments. *J. Biol. Chem.* **237**, 1460-1468

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

Keech, D.B. & Barritt, G.J. (1967) Allosteric activation of sheep kidney pyruvate carboxylase by magnesium ion (Mg^{2+}) and the magnesium adenosine triphosphate ion ($MgATP^{2-}$). *J. Biol. Chem.* **242**, 1983-1987

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

Keech, D.B. & Attwood, P.V. (1985) The reaction mechanism. In Pyruvate carboxylase, Ed. Keech, D.B. & Wallace, J.C., pages 109-140. CRC Press Inc, Boca Raton, Florida .

Khew-Goodall, Y.S., Johannssen, W., Attwood, P.V., Wallace, J.C. & Keech, D.B. (1991) Studies on dilution inactivation of sheep liver pyruvate carboxylase. *Arch. Biochem. Biophys.* **284**, 98-105

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

Kondo, H., Uno, S., Komizo, Y. & Sunamoto, J. (1984) Importance of methionine residues in the enzymatic carboxylation of biotin-containing peptides representing the local biotinyl site of *E. coli* acetyl-CoA carboxylase. *Int. J. Pept. Protein Res.* **23**, 559-564

Kondo, H., Shiratsuchi, K., Yoshimoto, T., Matsuda, T., Kitazono, A., Tsuru, D., Anai, M., Sekiguchi, M. & Tenabe, T. (1991) Acetyl-CoA carboxylase from *Escherichia coli*: gene organization and nucleotide sequence of the biotin carboxylase subunit. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 9730-9733

Kozak, M. (1981) Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res.* **9**, 5233-5252

subunit of rat propionyl-CoA carboxylase. *Proc. Natl. Acad. Sci.* **83**, 8049-8053

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

Krebs, H.A. & Kornberg, H.L. (1957) A survey of energy transformations in living matter. *Ergeb. Physiol. Biol. Chem. Exp. Pharmakol.* **49**, 212

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

Kumar, G.K., Beegan, H. & Wood, H.G. (1988b) Involvement of tryptophans at the catalytic and subunit-binding domains of transcarboxlase. *Biochemistry* **27**, 5972-5978

Kunkel, T.A., Roberts, J.D. & Zakour. (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154**, 367-382

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

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

Lamhonwah, A.-M., Mahuran, D. & Gravel, R.A. (1989) Human mitochondrial propionyl-CoA carboxylase: localization of the N-terminus of the pro- and mature α chains. *Nucleic Acids Res.* **17**, 4396

Lane, M.D., Rominger, K.L., Young, D.L. & Lynen, F. (1964) The enzymatic synthesis of holotranscarboxylase from apotranscarboxylase and (+)-biotin: II investigation of the reaction mechanism. *J. Biol. Chem.* **239**, 2865-2871

Le Gouill, C., Desmardis, D. & Dery, C.V. (1993) *Saccharopolyspora hirsuta* 367 encodes clustered genes similar to ketoacyl synthetase, ketoacyl reductase, acyl carrier protein and biotin carboxyl carrier protein. *Mol. Gen. Genet.* **240**, 146-150

Lebioda, L. & Stec, B. (1991) Mechanism of enolase: the crystal structure of enolase-Mg^{2+}-phosphoglycerate/phosphoenolpyruvate at 2.2 Å resolution. *Biochemistry* **30**, 2817-2822

Li, S.-J. & Cronan, J.E.Jr. (1992) The gene encoding the biotin carboxylase subunit of *Escherichia coli* acetyl-CoA carboxylase. *J. Biol. Chem.* **267**, 855-863

Libor, S.M., Sundaram, T.K. & Scrutton, M.C. (1978) 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. *Biochem. J.* **169**, 543-558

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

Lim, F., Rohde, M., Morris, C.P. & Wallace, J.C. (1987) Pyruvate carboxylase in the yeast *pyc* mutant. *Arch. Biochem. Biophys.* **258**, 259-264

Lim, F. (1988) Pyruvate carboxylase: the protein and gene in yeast. PhD. Thesis, University of Adelaide, Adelaide, South Australia, Australia.

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

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

Lorber, V., Lifson, N., Wood, H.G., Sakami, W. & Shreeve, W.W. (1950a) Conversion of lactate to liver glycogen in the intact rat, studied with isotopic lactate. *J. Biol. Chem.* **183**, 517-530

Lorber, V., Lifson, N., Sakami, W. & Wood, H.G. (1950b) Conversion of propionate to liver glycogen in the intact rat, studied with isotopic propionate. *J. Biol. Chem.* **183**, 531-537

Lusty, C.J., Windgren, E.E., Broglie, K.E. & Nyunoya, H. (1983) Yeast carbamoyl phosphate synthetase: structure of the yeast gene and homology to *Escherichia coli* carbamoyl phosphate synthetase. *J. Biol. Chem.* **258**, 14466-14472

Lynen, F. (1975) Structure of biotin enzymes. In Lypmann symposium: Energy regulation & biosynthesis in Molecular biology, Ed. Richter, D. pages 671-698.

MacKay, N., Rigat, B., Douglas, C., Shiene Chen, H. & Robinson, B.H. (1994) cDNA cloning of human kidney pyruvate carboxylase. *Biochem. Biophys. Res. Comm.* **202**, 1009-1014

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

Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982) Molecular Cloning: A laboratory Manual. Cold Spring Harbour Laboratory, Cold Spring Harbor, New York, USA

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

McAllister, H.C. & Coon, M.J. (1966) Further studies on the properties of liver propionyl coenzyme A holoenzyme synthetase and the specificity of holocarboxylase formation. *J. Biol. Chem.* **241**, 2855-2861

McClure, W.R., Lardy, H.A. & Cleland, W.W. (1971) Rat liver pyruvate carboxylase: isotopic exchange studies of the first partial reaction. *J. Biol. Chem.* **246**, 3584-3590

McCullough, W. & Roberts, C.F. (1980) Genetic regulation of isocitrate lyase in *Aspergillus nidulans*. *J. Gen. Microbiol.* **120**, 67-84

McGurk, K.S. & Spivey, H.O. (1979) Ligand induced conformational transitions and secondary structure composition of chicken liver pyruvate carboxylase. *Biochem. J.* **177**, 697-705

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

Mildvan, A.S. & Scrutton, M.C. (1967) Pyruvate carboxylase. X. The demonstration of direct co-ordination of pyruvate and α-ketobutarate by the bound magnesium and formation of enzyme-metal-substrate bridge complexes. *Biochemistry* **6**, 2978-2994

Miles, J.S., Guest, J.R., Radford, S.E. & Perham, R.N. (1988) Investigation of the mechanism of active site coupling in pyruvate dehydrogenase multienzyme complex of *Escherichia coli* by protein engineering. *J. Mol. Biol.* **202**, 97-106

Morris, C.M., Lim, F. & Wallace, J.C. (1987) Yeast pyruvate carboxylase: gene isolation. *Biochem. Biophys. Res. Comm.* **145**, 390-396

Mortimer, R.K., Schild, D., Contopoulou, C.R. & Kans, J.A. (1989) Genetic map of *Saccharomyces cerevisiae*, Edition 10. *Yeast* 321-403

Moss, J. & Lane, M.D. (1972) The biotin-dependent enzymes. *Adv. Enzymol.* **35**, 321-442

Mossé, M.-O., Linder, P., Lasowska, J. & Slonimski, P.P. (1993) A comprehensive compilation of 1001 nucleotide sequences coding for proteins from the yeast *Saccharomyces cerevisiae* (=ListA2). *Curr. Genet.* **23**, 66-91

Mozier, N.M., Walsh, M.P. & Pearson, J.D. (1991) Characterization of a novel zinc binding site of protein kinase C inhibitor-1. *FEBS. Lett.* **279**, 14-18

Munkres, K.D. & Richards, F.M. (1965) *Arch. Biochem. Biophys.* **109**, 466-478

Murtif, V.L., Bahler, C.R. & Samols, D. (1985) Cloning and expression of the 1.3S biotin-containing subunit of transcarboxylase. *Proc. Natl. Sci. U.S.A.* **82**, 5617-5621

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

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

Nielsen, J. (1970) Aspects of the mechanism of sheep kidney pyruvate carboxylase. PhD. Thesis,University of Adelaide, Adelaide, South Australia, Australia.

Norman, E., De Smet, K.A., Stoker, N.G., Ratledge, C., Wheeler, P.R., Dale, J.W. (1994) Lipid synthesis in mycobacterium: characterisation of the biotin carboxyl carrier protein genes from *Mycobacterium leprae* and *M. tuberculosis*. *J. Bacteriol.* **176**, 2525-2531

Northrop, D.B. & Wood, H.G. (1969) Transcarboxylase: exchange reactions and kinetics of oxalate inhibition. *J. Biol. Chem.* **244**, 5801-5807

Nyunoya, H. & Lusty, C.J. (1983) The *car*B gene of *Escherichia coli*: a duplicated gene coding for the large subunit of carbamoyl-phosphate synthetase. *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4629-4633

Nyunoya, H., Broglie, K.E., Windgren, E.E. & Lusty, C.J. (1985) Characterization and derivation of the gene coding for mitochondrial carbamoyl phosphate synthetase I of rat. *J. Biol. Chem.* **260**, 9346-9356

O'Keefe, S.J. & Knowles, J.R. (1986) Biotin-dependant carboxylation catalysed by transcarboxylase is a stepwise process. *Biochemistry* **25**, 6077-6084

Obermayer, M. & Lynen, F. (1976) Structure of biotin enzymes. *Trends Biochem. Sci.* **1**, 169-171

Ogata, K. (1970) Microbiol synthesis of dethiobiotin and biotin. *Methods Enzymol.* **18A**, 390-394

Oliver, S., *et al.* (1992) The complete DNA sequence of yeast chromosome III. *Nature* **357**, 38-46

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

Osmani, S.A. & Scrutton, M.C. (1983) The sub-cellular localisation of pyruvate carboxylase and of some other enzymes in *Aspergillus nidulans*. *Eur. J. Biochem.* **133**, 551-560

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

Osmani, S.A. & Scrutton, M.C. (1985) The subcellular localisation and regulatory properties of pyruvate carboxylase from *Rhizopus arrhizus*. *Eur. J. Biochem.* **147**, 119-128

Oura, E., Haarasilta, S. & Londesborough, J. (1980) Carbon dioxide fixation by bakers yeast in a variety of growth conditions.*Gen. Microbiol.* **118**, 51-58

Palacian, E. & Neet, K.E. (1970) Pyruvate carboxylase: effect of chemical modification of the sulfhydryl groups on the activity and quaternary structure. *Biochem. Biophys. Acta* **212**, 158-169

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

Pares, S., Cohen-Addad, C., Sieker, L., Neuberger, M. & Douce, R. (1994) X-ray structure determination at 2.6-Å resolution of a lipoate-containing protein: The H-protein of the glycine decarboxylase complex from pea leaves. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4850-4853

Perham, R.N. (1991) Domains, motifs and linkers in 2-oxo acid dehydrogenase multienzyme complexes. A paradigm in the design of a multifunctional protein. *Biochemistry* **20**, 8501-8512

Polakis, S.E., Guchhait, R.B. & Lane, M.D. (1972) On the possible involvement of a carbonyl phosphate intermediate in the triphosphate-dependent carboxylation of biotin. *J. Biol. Chem.* **247**, 1335-1337

Polakis, S.E., Guchhait, R.B., Zwergel, E.E. & Lane, M.D. (1974) Acetyl coenzyme A carboxylase system of *Escherichia coli*: studies on the mechanism of the biotin carboxylase- and carboxytransferase-catalysed reactions. *J. Biol. Chem.* **249**, 6657-6667

Post, L.E., Post, D.J. & Raushel, F.M. (1990) Dissection of the functional domains of *Escherichia coli* carbamoyl phosphate synthetase by site-directed mutagenesis. *J. Biol. Chem.* **265**, 7742-7747

Quinn, C.L., Stephenson, B.T. & Switzer, R.L. (1991) Functional organization and nucleotide sequence of the *Bacillus subtilis* pyrimidine biosynthesis operon. *J. Biol. Chem.* **266**, 9113-9127

Radford, S.E., Laue, E.D., Perham, R.N., Martin, S.R. & Apella, E. (1989) Conformational flexibility and folding of synthetic peptides representing an inter-domain segment of polypeptide chain in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J. Biol. Chem.* **264**, 767-775

Reed, L.J. (1974) Multienzyme complexes. *Acc. Chem. Res.* **7**, 40-46

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

Reed, L.J. & Hackert, M.L. (1990) Structure-function relationships in dihydrolipoamide acyltransferases. *J. Biol. Chem.* **265**, 8971-8974

Reed, K.E. & Cronan, J.E.Jr. (1991) *Escherischia coli* exports previously folded and biotinated protein domains. *J. Biol. Chem.* **266**, 11425-11428

Riles, L., Dutchik, J. E., Baktha, A., McCauley, B. K., Thayer, E. C., Leckie, M. P., Braden, V. V., Depke, J. E. & Olson, M. V., (1993) Physical maps of the six smallest chromosomes of *Saccharomyces cerevisiae* at a resolution of 2.6 kilobase pairs. *Genetics*. **134**, 81-150

Rock, C.O. & Cronan, J.E.Jr. (1980) Improved purification of acyl carrier protein. *Anal. Biochem.* **102**, 362-364

Roesler, K.R., Shorrosh, B.S. & Ohlrogge, J.B. (1994) Structure and expression of an *Arabidoposis* acetyl-CoA carboxylase gene. *Plant Physiol.* **105**, 611-617

Roessler, P.G. & Ohlrogge, J.B. (1993) Cloning and characterization of the gene that encodes acetyl-CoA carboxylase in the alga *Cyclotella cryptica*. *J. Biol. Chem.* **268**, 19254-19259

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

Rohde, M., Lim, F. & Wallace, J.C. (1991) Electron microscopic localisation of pyruvate carboxylase in rat liver and Saccharomyces cerevisiae by immunogold procedures. *Arch. Biochem. Biophys.* **290,** 197-201

Rose, I.A., O'Connell, E.L. & Solomon, F. (1976) Intermolecular tritium transfer in the transcarboxylation reaction. *J. Biol. Chem.* **251**, 902-904

Rose, M., Grisafi, P. & Boltstein, D. (1984) Structure and function of the yeast *URA3* gene: expression in *Escherichia coli*. *Gene* **29**, 113-124

Rothstein, R.J. (1983) One step gene disruption in yeast. *Methods Enzymol.* **101**, 202-211

Ruiz-Amil, M., de Torrontegui, G., Palacian, P., Catalina, L. & Losada, M. (1965) Properties and function of yeast pyruvate carboxylase. *J. Biol. Chem.* **240**, 3485-3492

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

Sakai, A., Chibazakura, T., Shimiza, Y. & Hishinuma, F. (1992) Molecular analysis of *POP2* gene, a gene required for glucose-derepression of gene expression in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **20**, 6227-6223

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) Molecular cloning: a laboratory manual. 2nd ed., NY: Cold Spring Harbor Laboratory, Cold Spring Harbor.

Samols, D., Thornton, C.G., Murtif, V.L., Kumar, G.K., Haase, F.C. & Wood, H.G. (1988) Evolutionary Conservation among Biotin Enzymes. *J. Biol. Chem.* **263**, 6461-6464

Saraste, M., Sibbald, P.R. & Wittinghofer, A. (1990) The P-loop a common motif in ATP and GTP-binding proteins. *Trends Biochem. Sci.* **15**, 430-434

Sauers, C.K., Jencks, W.P. & Groh, S. (1975) The alcohol-bicarbonate-water system. Structure-reactivity studies on the equilibria for formation of alkyl monocarbonates and on rates of their decomposition in aqueous alkali. *J. Am. Chem. Soc.* **97**, 5546-5553

Schagger, H. & von Jagow, G. (1987) Tricine-sodium dodecyl sulphate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 - 100 kDa. *Anal. Biochem.* **166**, 368-379

Schwartz, D.C. & Cantor, C.R. (1984) Separation of yeast chromosome-sized DNAs by pulse field gradient gel electrophoresis. *Cell* **37**, 67-75

Schwarz, E., Oesterheld, D., Reinke, H., Beyreuther, K. & Dimroth, P. (1988) The sodium ion translocating oxaloacetate decarboxylase of *Klebsiella pneumoniae*. *J. Biol. Chem.* **263**, 9640-9645

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

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

Scrutton, M.C. & Utter, M.F. (1965) Pyruvate carboxylase. III Some physical and chemical properties of highly purified enzymes. *J. Biol. Chem.* **240**, 1-9

Scrutton, M.C., Utter, M.F. & Mildvan, A.S. (1966) Pyruvate carboxylase. IV The presence of tightly bound manganese. *J. Biol. Chem.* **241**, 3480-3487

Scrutton, M.C., Young, M.R. & Utter, M.F. (1970) Pyruvate carboxylase from bakers yeast: the presence of bound zinc. *J. Biol. Chem.* **245**, 6220-6227

Scrutton, M.C. & Young, M.R. (1972) Pyruvate carboxylase. In The Enzymes, Ed. Boyer, P. D., 3rd ed., Vol. 6, page 13, Acedemic Press, New York

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

Scrutton, M.C., Griminger, P. & Wallace, J.C. (1973a) Bound metal content of the vertebrate liver enzyme as a function of diet and species. *J. Biol. Chem.* **247**, 3305-3313

Scrutton, M.C., Reed, G.H. & Mildvan, A.S. (1973b) Application of physical methods to the study of enzymes containing bound manganese: problems and prospects. *Adv. Exp. Med. Biol.* **40**, 79-102

Scrutton, M.C., Pearce, P.H. & Fatabene, F. (1977) 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. *Eur. J. Biochem.* **76**, 219-231

Seubert, W. & Remberger, U. (1961) Reinigung und Wirkungsweise der Pyruvatcarboxylase aus *Pseudomonas citronellolis*. *Biochem. Z.* **334**, 401-414

Sharp, P.M., Tuohy, T.M.F. & Mosursky, K.R. (1986) Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Res.* **14**, 5125-5143

Shenoy, B.C., Xie, Y., Park, V.L., Kumar, G.K., Beegen, H., Wood, H.G. & Samols, D. (1992) The importance of methionine residues for the catalysis of the biotin enzyme, transcarboxylase: analysis by site-directed mutagenesis. *J. Biol. Chem.* **267**, 18407-18412

Shorrosh, B.S., Dixon, R.A. & Ohlrogge, J.B. (1994) Molecular cloning, characterization, and elicitation of acetyl-CoA carboxylase from Alfalfa. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 4323-4327

Simon, E.J. & Shemin, D. (1953) The preparation of S-succinyl coenzyme A. *J. Amer. Chem. Soc.* **75**, 2520

Skinner, V.M. & Armitt, S. (1972) Mutants of *Aspergillus nidulans* lacking pyruvate carboxylase. *FEBS Lett.* **20**, 16-18

Snoswell, M.A. (1982) Chicken liver pyruvate carboxylase: mechanism of the first partial reaction. B. Sc. (Hons) Thesis, University of Adelaide, South Australia, Australia.

Southern, E.M. (1975) Detection of specific sequences and DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503-517

Spradlin, J. & Thoma, J.A. (1970) β-Amylase thiol groups: possible regulator sites. *J. Biol. Chem.* **245**, 117-127

Stallings, W.C., Paltridge, K.A., Strong, R.K. & Ludwig, M.L. (1985) The structure of manganese superoxide dismutase from *Thermus thermophillus* HB8 at 2.4 Å resolution. *J. Biol. Chem.* **260**, 16424-16432

Sternberg, M.J.E. & Taylor, W.R. (1984) Modelling the ATP-binding site of oncogene products, the epidermal growth factor receptor and related proteins. *FEBS Lett.* **175**, 387-391

Stinchcomb, D.T., Struhl, K. & Davis, R.W. (1979) Isolation and characterization of a yeast chromosomal replicator. *Nature* **282**, 39-43

Struhl, K., Stinchcomb, D.T., Scherer, S. & Davis, R.W. (1979) High frequency transformation of yeast: autonomous replication of hybrid DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1035-1039

Stucka, R., Dequin, S., Salmon, J. & Gancedo, C. (1991) DNA sequences in chromosomes II and VII code for pyruvate carboxylase isoenzymes in *Saccharomyces cerevisiae*: analysis of pyruvate carboxylase deficient strains. *Mol. Gen. Genet.* **229**, 307-315

Studier, F.W. & Moffat, B.A. (1986) Use of bacterophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113-130

Sumegi, B., McCammon, M.T., Sherry, A.D., Keys, D.A., McAlister-Henn, L., & Srere, P.A. (1992) Metabolism of [3-^{13}C]pyruvate in TCA cycle mutants of yeast. *Biochem.* **31**, 8720-8725

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

Takeshita, S., Sato, M., Tabo, M., Masahashi, W. & Hashimoto-Gotoh, T. (1987) High-copy-number and low-copy-number plasmid vectors for *LacZ* α-complimentation and chloramphenicol - or kanamycin resistance selection. *Gene* **61**, 63-74

Taylor, B.L., Barden, R.E. & Utter, M.E. (1972) Identification of the reacting form of pyruvate carboxylase. *J. Biol. Chem.* **247**, 7383-7390

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

Texter, F.L., Radford, S.E., Laue, E.D., Perham, R.N., Miles, J.S. & Guest, J.R.(1988) Site-directed mutagenesis and ^{1}H-NMR spectroscopy of an interdomain segment in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *Biochemistry* **27**, 289-296

Thornton, C.G., Kumar, G.K., Shenoy, B.H., Haase, F.C., Phillips, N.F.B., Park, V.M., Magner, W.J., Hejlik, D.P., Wood, H.G. & Samols, D.(1993) Primary structure of the 5S subunit of transcarboxylase as deduced from the genomic DNA sequence. *FEBS Lett.* **330**, 191-196

Tipton, P.A. & Cleland, W.W. (1988a) Catalytic mechanism of biotin carboxylase: steady state kinetic investigations. *Biochemistry* **27**, 4317-4325

Tipton, P.A. & Cleland, W.W. (1988b) Carbon-13 and deuterium isotope effects on the catalytic reactions of biotin carboxylase. *Biochemistry* **27**, 4325-4331

Toh, H., Kondo, H. & Tanabe, T. (1993) Molecular evolution of biotin-dependent carboxylases. *Eur. J. Biochem.* **215**, 687-696

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

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

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

Utter, M.F. & Keech, D.B. (1960) Formation of oxaloacetate from pyruvate and CO_2. *J. Biol. Chem.* **235**, 17-18

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

Utter, M.F. (1969) Metabolic roles of oxaloacetate. In Citric acid acid cycle. Control and compartmentalization, ed. Lowenstein, J.M. pages 249-296. Marcel Dekker, New York .

Vallee, B.L. & Auld, D.S. (1990) Zinc coordination, function, and structure of zinc enzymes and other proteins. *Biochemistry* **29**, 5647-5659

Vernet, T., Dignard, D. & Thomas, D.Y. (1987) A family of yeast expression vectors containing the phage f1 intergenic region. *Gene* **52**, 225-233

Visser, C.M. & Kellogg, R.M. (1978) Biotin. Its place in evolution. *J. Mol. Evol.* **11**, 171-187

Vollrath, D., Davis, R.W., Connelly, C. & Heiter, P. (1988) Physical mapping of large DNA by chromosome fragmentation. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6027-6031

Waldrop, G., Holden, H.M. & Rayment, I. (1994a) Preliminary X-ray crystallographic analysis of biotin carboxylase isolated from *Escherichia coli*. *J. Mol. Biol.* **235**, 367-369

Waldrop, G., Rayment, I. & Holden, H.M. (1994b) Three-dimensional structure of the biotin carboxylase subunit of acetyl-CoA carboxylase. *Biochemistry* **33**, 10249-10256

Walker, J.E., Saraste, M., Runswick, M.J. & Gay, N.J. (1982) Distantly related sequences in the α- and β-subunits of the ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide fold. *EMBO J.* **1**, 945-951

Walker, M.E., Val, D.L., Rohde, M., Devenish, R.J. & Wallace, J.C. (1991) Yeast pyruvate carboxylase: Identification of two genes encoding isoenzymes. *Biochem. Biophys. Res. Comm.* **176**, 1210-1217

Walker, M.E. & Walace, J.C. (1991) Isolation of a yeast mutant deficient in pyruvate carboxylase activity. *Biochem. Int.* **23**, 697-705

Walker, M.E., Jitrapakdee, S., Val, D. L. & Walace, J.C. (1995) Human pyruvate carboxylase: heterogeneity at the 5' end of the mRNA. *Biochem. Biophys. Acta.* (manuscript in preparation)

Wallace, J.C. (1985) Distribution and biological functions of pyruvate carboxylase in nature. In Pyruvate Carboxylase, ed. Keech, D.B. & Wallace, J.C., pages 6-63, CRC Press Inc., Boca Raton, Florida .

Wallace, J.C. & Easterbrook-Smith, S.B. (1985) The structure of pyruvate carboxylase. In Pyruvate carboxylase, Ed. Keech, D.B. & Wallace, J.C., pages 66-108, CRC Press Inc., Boca Raton, Florida .

Wallis, N.G. & Perham, R.N. (1994) Structural dependence of post-translational modification and reductive acylation of the lipoyl domain of the pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* **236**, 209-216

Warren, G.B. & Tipon, K.F. (1974) Pig liver pyruvate carboxylase: the reaction pathway for the carboxylation of pyruvate. *Biochem. J.* **139**, 311-320

Welch, J.W., Fogel, S., Buchman, C. & Karin, M. (1989) The *CUP2* gene product regulates the expression of *CUP1* gene coding for yeast metalothionine. *EMBO J.* **8**, 255-260

Wang, D., Waye, M.M., Taricani, M., Buckingham, K. & Sandham, H.J. (1993) Biotin-containing protein as a cause of false positive clones in gene probing with streptavidin / biotin. *BioTechniques* **209**, 209-212

Werneberg, B.G. & Ash, D.E. (1993) Chemical modifications of chicken liver pyruvate carboxylase: Evidence for essential cysteine-lysine pairs and a reactive sulfhydryl group. *Arch. Biochem. Biophys.* **303**, 214-221

Whal, G., Ong, E., Meinkoth, J., Franko, R. & Barinaga, M. (1981) Methods for transfer of DNA, RNA and protein to nitrocellulose and diazotized paper solid supports. Schleicher and Schuell, Inc.

Whittle, A.M. (1986) Propionyl-CoA carboxylase from chicken liver. B. Sc. (Hons), University of Adelaide, Adelaide, South Australia, Australia.

Wilce, M.C.J. & Parker, M.W. (1994) Structure and function of glutathione S-transferases. *Biochem. Biophys. Acta.* **1205**, 1-18

Wilson, D.K., Rudolph, F.B. & Quiocho, F.A. (1991) Atomic structure of adenosine deaminase complexed with a transition-state analog: understanding catalysis and immunodeficiency mutations. *Science* **252**, 1278-1284

Wilson, K.P., Shewchuk, L.M., Brennan, R.G., Otsuka, A.J. & Matthews, B.W. (1992) *Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure deliniates the biotin- and DNA-binding domains. *Proc. Natl. Acad. Sci. U.S.A* **89**, 9257-9261

Wimmer, M.J., Rose, I.A., Powers, S.G. & Meister, A. (1979) Evidence that carboxyphosphate is a kinetically competant intermediate in the carbamyl phosphate synthetase reaction. *J. Biol. Chem.* **254**, 1854-1859

Woehlke, G., Wifling, K. & Dimroth, P. (1992) Sequence of the sodium ion pump oxaloacetate decarboxylase from *Salmonella typhimurium*. *J. Biol. Chem.* **267**, 22798-22703

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

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

Wood, H.G. (1979) The anatomy of transcarboxylase and the role of its subunits. *CRC Crit. Rev. Biochem.* **7**, 143-160

Wood, H.G., Harmon, F.R., Wühr, B., Hübner, K. & Lynen, F. (1980) Comparison of the biotination of apotranscarboxylase and its aposubunits. *J. Biol. Chem.* **255**, 7397-7409

Wurtele, E.S. & Nikolau, B.J. (1990) Plants contain multiple biotin enzymes: discovery of 3-methylcrotonyl-CoA carboxylase, propionyl-CoA carboxylase and pyruvate carboxylase in the plant kingdom. *Arch. Biochem. Biophys.* **278**, 179-186

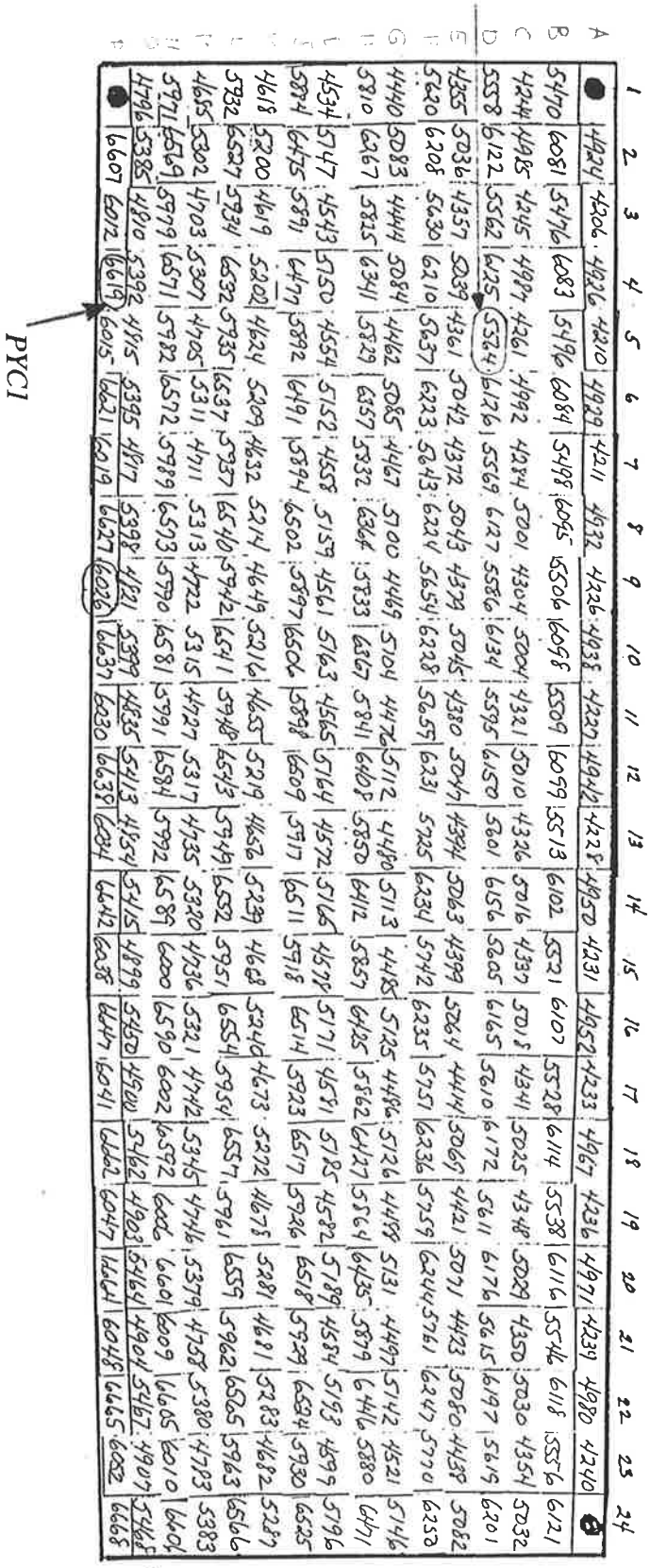
Young, M.R., Tolbert, B. & Utter, M.F. (1969) Pyruvate carboxylase from *Saccharomyces cerevisiae*. *Meth. Enzymol.* **13**, 250-258

Zhang, J., Xia, W., Brew, K. & Ahmad, F. (1993) Adipose pyruvate carboxylase: amino acid sequence and domain structrure from cDNA sequencing. *Biochemistry* 1766-1770

Zhang, S., Sanyal, I., Bulboaca, G.H., Rich, A. & Flint, D.H. (1994) The gene for biotin synthase from *Saccharomyces cerevisiae*: cloning, sequencing, and complementation of *Escherischia coli* strains lacking biotin synthase. *Arch. Biochem. Biophys.* **309**, 29-35

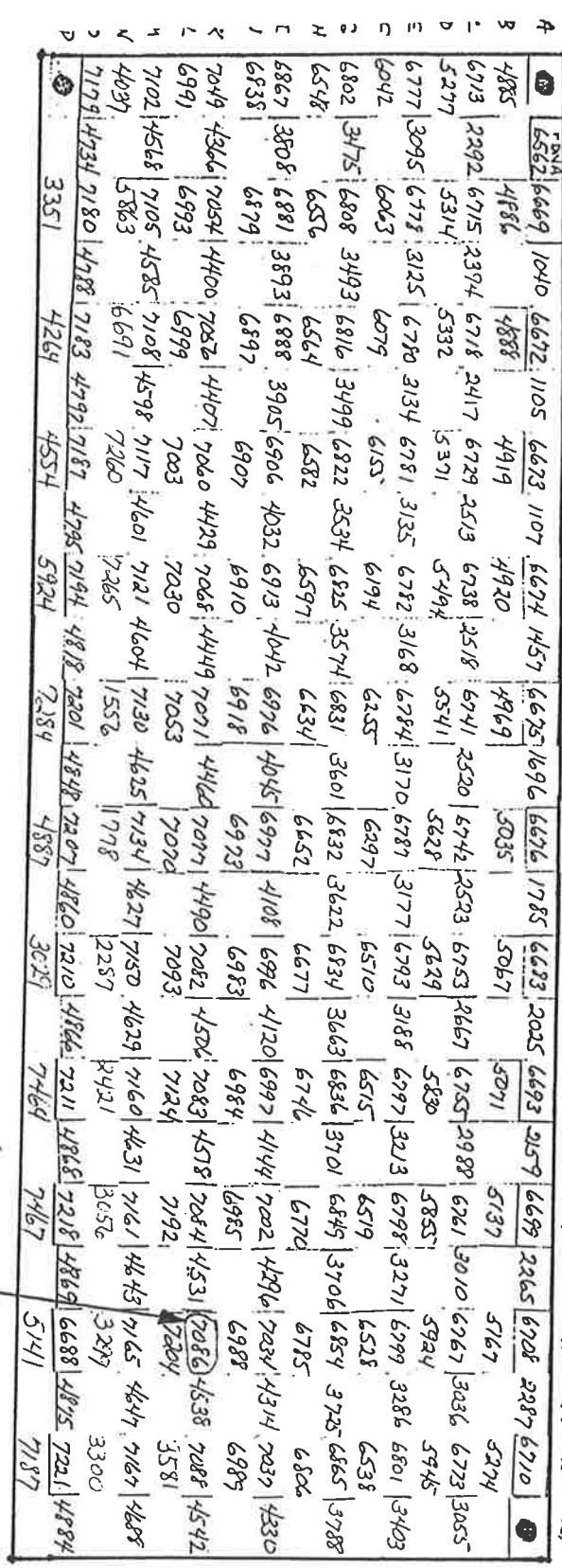
Appendix A

Key to the clone numbers for the grid positions on the "Olson" yeast chromosome mapping filter No. 2.



Appendix B

Key to the clone numbers for the grid positions on the "Olson" yeast chromosome mapping filter No. 3.



PYC1