Studies of the glyoxylate bypass of *Streptomyces coelicolor* A3(2)

by

JO DOWMAN

Thesis submitted to the Faculty of Science, University of Glasgow,
for the degree of Master of Science.

Department of Genetics.

December 1997

The research reported in this thesis is my own original work except where otherwise stated and has not been submitted for any other degree.

Jo Dowman

Abstract

In this study, molecular cloning of the structural gene encoding malate synthase (*aceB*) from Streptomyces coelicolor was attempted using two different approaches. Despite identification of potential sequences using genetic fingerprinting and southern hybridisation techniques, all my attempts to clone the full coding sequence of the gene were unsuccessful. Analysis of DNA sequences upstream of the structural gene encoding isocitrate lyase (*aceA*), led to the identification of potential open reading frame (pORF) which shows a very close similarity to pORF from Mycobacterium tuberculosis. The function of such an ORF in both organisms remains to be discovered.

CONTENTS

List of contents

CHAPTER 1 General Introduction

CHAPTER 2 Materials and Methods

List of Figures

CHAPTER 1

CHAPTER 3

Abbreviations

AMP	adenosine monophosphate
ATP	adenosine triphosphate
kDa	kilodaltons
DMSO	dimethylsulphoxide
DNase	deoxyribonuclease
DTT	dithiothreitol
EDTA	ethylene diamine tetra acetate
EtBr	ethidium bromide
ICL	Isocitrate lyase (EC 4.1.3.1)
ICDH	Isocitrate dehydrogenase (EC 1.1.1.42)
IPTG	isopropyl-b-D-thiogalactoside
kb	kilo base pairs
knt	kilonucleotides
MS	Malate synthase (EC 4.1.3.2)
MES	2-(N-Morpholino) ethanesulphonic acid
MOPS	Morpholino propane sulphonic acid.
MW	molecular weight
NAD+	nicotinamide adenine dinucleotide
NADP+	nicotinamide adenine dinucleotide phosphate
NADPH	reduced nicotinamide adenine dinucleotide phosphate
OAA	oxaloacetate.
PCR	polymerase chain reaction
PEG	polyethylene glycol
PEP	phosphoenol pyruvate
3-PG	3-phosphoglycerate
pfu	plaque forming units
Rf	Mobility of a protein compared to the dye front
RNase	ribonuclease
SDS	sodium dodecyl sulphate
SSC	standard saline citrate
TAE	Tris-acetate/EDTA buffer
TBE	Tris-borate/EDTA buffer
TEMED	N,N,N'N'-tetramethylethylene diamine
TES	N-tris (Hydroxymethyl) methyl-2-aminoethane sulphonic acid

Tris	Tris (hydroxymethyl) aminomethane
U	unitsof enzyme activity
UDP	uridine diphosphate
UV	ultra violet
X-gal	5-bromo-4-chloro-3-indoyl-b-galactoside

Acknowledgements

Many thanks to everyone who has helped me in both the practical work and the writing of this thesis.

The practical work was supported by a grant from the MRC and writing-up was funded by various means!!

Thanks go to Iain Hunter, my supervisor, and everyone else in the lab who aided me with my practical work. Particular thanks to Paul for his support and help.

Thanks to John and Mike for doing some of the most tedious typing.

Finally, thanks to my mum, without whose endless encouragement, I would never have completed the write-up.

Chapter 1

General Introduction

1.1 General Introduction

This chapter aims to provide a background to the work presented in this thesis. First there is a general introduction to *Streptomyces*. This is followed by a more detailed review of fatty acid and acetate catabolism, particularly the glyoxylate bypass, the enzymes involved and their regulation. The final section of this chapter describes the aims of this project.

1.2 Introduction to *Streptomyces*

Streptomyces are Gram-positive, aerobic, filamentous, eubacteria commonly found in soil and leaf litter. They obtain nutrients by the degradation of organic material such as cellulose, lignin and xylan by the excretion of degradative enzymes. Streptomycetes possess a complex life cycle of morphological differentiation.

The soil community usually exists in a state of substrate limitation and streptomycetes have adapted to growth in hostile environments. During times of unfavourable conditions, aerial hyphae appear which subsequently differentiate to form spores. This is concurrent with the lysis of the substrate mycelia and secondary metabolite production (Chater, 1984). Production of spores allows the dispersal and continued survival of the species.

Streptomycetes have a characteristic DNA composition with a high G+C bias of 69-73%. It is conceivable that such a high G+C content might have substantially influenced the regulatory mechanisms in streptomycetes. The genome size has been estimated using pulse-field gel electrophoresis and found to be 6-9Mb of DNA, approximately 75% larger than that of *Escherichia coli* (Keiser *et al.*, 1992). Lin *et*

al., (1993) also reported that a number of streptomycetes including *Streptomyces coelicolor* have linear chromosomes.

Streptomyces coelicolor is the best studied streptomycete. It is genetically the best characterised, allowing for a variety of classical and recombinant DNA techniques (Hopwood *et al*, 1985) to be used in the study of this organism.

Streptomycetes produce a wide range of biologically-active molecules known as secondary metabolites. Secondary metabolites are not considered to be essential for growth of the organism but may give a competitive advantage. They are usually produced after vegetative growth and in conjuction with differentiation. They include immunosuppresants, anti-helminthics, anti-cancer agents and most importantly, antibiotics. Over 60% of all known antibiotics and over 75% of all commercially produced antibiotics are produced by *Streptomyces* (Omura, 1986). Much of the work concerned with the species has centred around production of antibiotics whilst they have also been studied as a model of differentiation.

1.2.1 Life Cycle

An environmental stimulus leads to spore germination and growth occurs by cell wall extension at the hyphal tips to form a branched vegetative mycelium throughout the substrate. Vegetative septa are infrequent and hyphal compartments contain many copies of the genome, especially the tip cells. In response to nutrient limitation, aerial hyphae are formed. This is associated with antibiotic production in some species and the breakdown of the substrate mycelium. It is thought that aerial hyphal growth is driven by osmotic pressure derived from the solubilisation of storage macromolecules such as glycogen. Regularly spaced cross walls divide the aerial

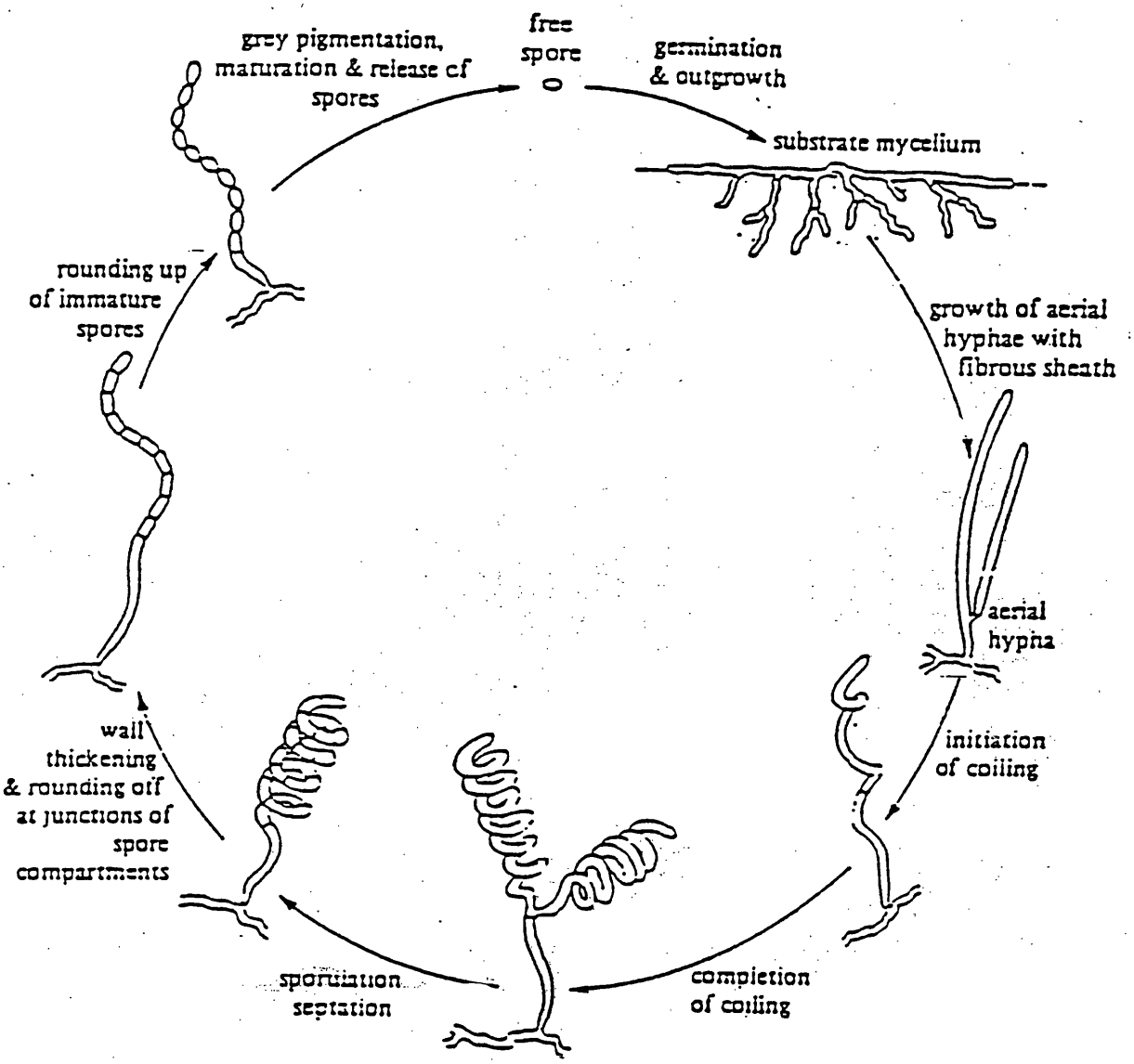


Figure 1.1 Life Cycle of Streptomycetes

This idealised diagram summarises the life cycle of *S.coelicolor* from germination to sporulation. Modified from Chater and Merrick, 1979.

hyphae and these become coiled chains of hydrophobic spores. Maturation results in thickening, rounding and usually pigmentation of the spore walls. A diagram of the streptomycete life cycle is shown in Figure 1.1.

1.2.2 Growth Studies

As previously mentioned, *Streptomyces* produce a great number of commercially-important antibiotics. Much of the work on *Streptomyces* has centred on antibiotic production and there has been very little investigation into the pathways of central metabolism. It is important to study these pathways in order to investigate the possibility of controlling carbon flux, enhancing the flux of metabolic precursors to secondary metabolism and therefore reducing unnecessary/wasteful metabolism during industrial fermentations.

Investigation of gene regulation in streptomycetes has been mainly confined to genes involved in antibiotic resistance and biosynthesis, and morphological differentiation. However, some research has also been done on carbon catabolism. The systems for utilisation of glycerol (Seno and Chater, 1983; Seno *et al.*, 1984; Smith and Chater, 1988a,b) and galactose (Fornwald *et al.*, 1987; Adams *et al.*, 1988) in *Streptomyces* have been studied. Both are subject to glucose repression, however, glucose repression in *Streptomyces* species does not appear to be mediated by cyclic AMP, unlike in *E.coli*. Thus studies of the operons involved in glycerol (*gyl*) and galactose (*gal*) utilisation of streptomycetes are important to elucidate the general features of operon organisation and the novel regulatory mechanisms used by this bacterium.

1.2.2.1 Glycerol Utilisation in *S.coelicolor*

Two enzymes are essential for glycerol catabolism, glycerol kinase and L-glycerol-3-phosphate dehydrogenase (G3PD). These enzymes are encoded by two co-ordinately expressed genes *gylA* and *gylB* respectively, which are contained in the glycerol-inducible, glucose-repressible glycerol (*gylABX*) operon (Seno and Chater, 1983; Smith and Chater, 1988a). *GylR* is found immediately upstream from the *gylABX* promoter region and encodes a regulatory protein essential for growth on glycerol (Smith and Chater, 1988b). *GylX* of unknown function is not essential for glycerol catabolism but it is thought it may be a facilitator for glycerol uptake which would not be essential when cultures are grown in high levels of glycerol as the carbon source would enter the cell by passive diffusion.

Northern analysis revealed several different *gyl* mRNA transcripts originating from a single transcription initiation site. The 4.3 kb mRNA species is thought to encode the two inducible polypeptides of 101 kDa and 51 kDa, glycerol kinase and G3PD respectively. A 5.4 kb transcript includes the region of *gylX* (Seno *et al.*, 1994).

A 27.6 kDa protein is encoded by *gylR*. The predicted sequence of this protein contains a helix-turn-helix motif characteristic of DNA-binding proteins. The protein is similar to *AsnC*, a transcriptional activator of *E.coli*. In prokaryotic systems an activator gene that is specific to particular pathway is typically contained in a monocistronic transcription unit immediately upstream from the transcription unit that it controls. Therefore evidence suggests that *gylR* encodes a *gyl* activator protein. There is also a region of dyad symmetry in the *gylABX* promoter region which resembles several prokaryotic *cis*-acting elements that bind regulatory proteins.

The pathways for glycerol catabolism are similar in *S.coelicolor* and *E.coli* but they do not appear to share a common mode of regulation. Unlike in *S.coelicolor*, the *glp* genes of *E.coli* are subject to negative control (Yang *et al.*, 1997).

A representation of the *gyl* operon can be seen in Figure 1.2a.

1.2.2.2 Galactose Utilisation in *S.lividans*

Three enzymes are required for galactose utilisation in *E.coli* and *Streptomyces lividans* and both systems comprise a metabolic operon which is induced in the presence of galactose and repressed by glucose. UDP galactose-4-epimerase is encoded by *galE*, galactose-1-phosphate (G-1-P) uridyltransferase by *galT* and galactokinase by *galK* (Shapiro and Adhya, 1969; Lemarie and Miller, 1986; Fornwald *et al.*, 1987).

In *E.coli* the genes are transcribed in the order *galE*, *galT*, *galK*. Only three nucleotides separate *galT* and *galK* helping to ensure coordinate expression (Schumperli *et al.*, 1982). This organisation is essential to ensure coordinate expression of the two genes (Schumperli *et al.*, 1982). The *gal* operon has been cloned from *S.lividans* and the gene order was found to be *galT*, *galE*, *galK* which differs from that of *E.coli*. The genes are separated by quite long intercistronic regions. Thus it seems likely that translational coupling is not a regulatory feature of the *Streptomyces gal* operon.

Fornwald *et al.* (1987) found that in *S.lividans* two promoters, *galP1* and *galP2*, direct transcription of two distinct polycistronic transcripts. *galP1*, which is located

immediately upstream of the operon, is induced by galactose and results in the transcription of all three *gal* genes. It appears that this promoter may be controlled in part by *cis*-acting elements (Mattern *et al.*, 1993). The second promoter, *galP2*, is located within the operon, just upstream of *galE* and is responsible for the constitutive expression of the *galE* and *galK* genes. The apparent total lack of homology between the *Streptomyces gal* promoters suggests that different factors may be required for transcription from each of these promoters.

In *E.coli* there is also a constitutive promoter upstream of the *galE* gene. UDP galactose-4-epimerase, the gene product of *galE*, is required for cell wall biosynthesis regardless of galactose being the sole carbon source or not, hence its constitutive expression. Basal levels of galactokinase are also high due to constitutive transcription from *galP2*. It is not clear what role this may play but unpublished evidence suggests that galactokinase may be involved in induction of the *gal* genes.

The *gal* operons of *E.coli* and *S lividans* show some similarities. Both are polycistronic operons that are induced by galactose and transcription is regulated by two promoters. However, organisation of the promoters is quite different. The promoters that regulate the *E.coli* operon overlap and are located at the 5' end of the operon whilst *S.lividans* has two totally separate promoters separated by *galT* and regulation is essentially separate.

A representation of the *gal* operon is shown in Figure 1.2b.

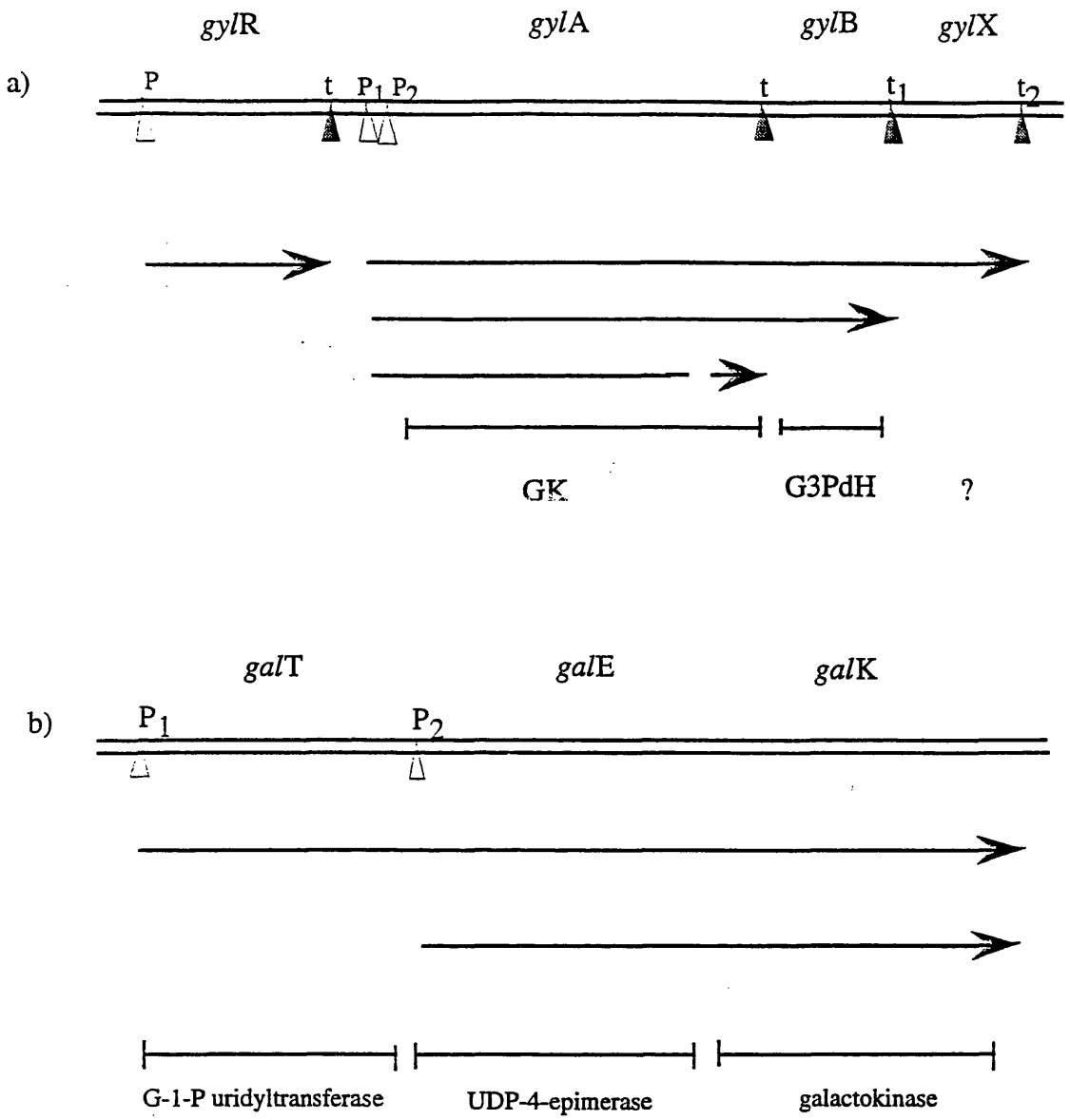


Figure 1.2 Organisation of the Streptomyces glycerol and galactose operons

a) Diagramatic representation of the operon for glycerol utilisation from *S.coelicolor*. The two structural genes, *glyA* and *gylB*, are shown, as well as the repressor, *glyR*. The transcripts obtained from the promoters (P,P_1,P_2) are also shown. The terminators are represented by t,t_1 and t_2.

b) Diagramatic representation of the operon for galactose utilisation from *S.lividans*. The three structural genes, *galT*, *galE* and *galK*, are shown. The transcripts from the constituitive promoter P_2 and the galactose inducible promoter, P_1 are shown.

1.3 Central Metabolism

Cells have a number of central metabolic pathways concerned with production of energy (in the form of ATP), reducing power and biosynthetic precursors from a carbon source. In aerobic organisms, energy and reducing power are obtained from the complete combustion of nutrients to CO_2 and H_2O.

The Tricarboxylic acid (TCA) cycle serves as a central pathway in most aerobic organisms. Acetyl-CoA enters the cycle by condensing with oxaloacetate (OAA) to form citrate. This then undergoes a series of reactions whereby energy and reducing power are generated with the loss of two carbon atoms as carbon dioxide. OAA is regenerated so the cycle can continue, but there is no net accumulation of carbon.

Some intermediates of the TCA cycle are required for biosynthesis, and these are channelled away from the cycle. If these are not replaced the OAA will not be regenerated resulting in the eventual breakdown of the cycle. Thus there are anaplerotic pathways which operate to maintain the levels of intermediates of the TCA cycle (Kornberg, 1966a).

During growth on glucose, TCA cycle intermediates are replaced by the action of the enzyme phosphoenolpyruvate (PEP) carboxylase which catalyses the carboxylation of PEP to OAA. This OAA can then be used to replenish TCA cycle intermediates. However, compounds such as fatty acids, acetate and ethanol are metabolised directly to acetyl-CoA, bypassing the production of PEP. Therefore the OAA must be regenerated by an alternative pathway when these compounds are provided as the sole carbon source. Many streptomycete fermentations are carried

out using media based on oils since this is often more cost effective and it is therefore important to understand the process by which growth on oils is carried out.

1.4 Fatty Acid Metabolism

When fatty acids enter the cell they are incorporated into more complex lipids or catabolised to acetyl-CoA. The pathway by which *E.coli* degrades fatty acids is very similar to the β-oxidative pathway of mammals and other eukaryotic organisms. Many bacteria can grow on fatty acids as a sole carbon source and most of the knowledge of the genetics and biochemistry of fatty acid degradation has been derived from studies with *E.coli*.

There are two systems by which *E.coli* can utilise fatty acids. The fatty acid - oxidative (FAO) system is concerned with transportation and catabolism of medium chain(C_7-C_{11}) fatty acids (MCFA) and long chain (C>11) fatty acids (LCFA), whilst the utilisation of short chain (C_4-C_6) fatty acids (SCFA) involves SCFA-degradative (ATO) enzymes as well as the FAO system.

The degradation of fatty acids involves the oxidation of a series of homologous substrates through a series of homologous intermediates. The synthesis of at least five FAO enzymes is coordinately induced when LCFAs are present in the growth media.

1.4.1 The FAO System

The genes encoding the FAO enzymes are located at several sites on the *E.coli* chromosome and are part of a regulon known as the *fad* regulon which is primarily responsible for the transport, acylation and β-oxidation of MCFAs and LCFAs.

Fatty acids must first enter the cell through the cell membrane. LCFAs enter via a carrier mechanism which involves two proteins encoded by *fadD* and *fadL*. *fadD* encodes the membrane-bound protein acyl-CoA synthetase and *fadL* encodes a 43 kDa outer membrane bound protein, FLP. It has been suggested that the LCFAs first bind to FLP and are transported across the outer membrane then they are somehow transferred across the cytoplasmic membrane to the peripheral membrane-bound acyl-CoA synthetase. Here they are activated and released into the cytoplasm. MCFAs can be actively transported by the same mechanism but if FLP is non-functional they can also diffuse across the cell membrane to the acyl-CoA synthetase.

As already described, the first step in the degradation of the free fatty acids is their activation to a fatty-acyl-CoA as part of the transport process, catalysed by acyl-CoA synthetase. The fatty-acyl-CoA then undergoes two oxidation steps, yielding one molecule of $FADH_2$ and one of NADH, followed by thiolytic cleavage to give acetyl-CoA and a fatty-acyl-CoA which is shorter than the original by two carbon atoms. The β-oxidation steps are catalysed by acyl-CoA dehydrogenase which is a multi-enzyme complex. The shortened fatty-acyl-CoA then re-enters the degradation cycle without further activation, thus yielding several molecules of acetyl-CoA for entry into the TCA cycle for each fatty acid molecule (Overath *et al.*, 1969) (Fig 1.3).

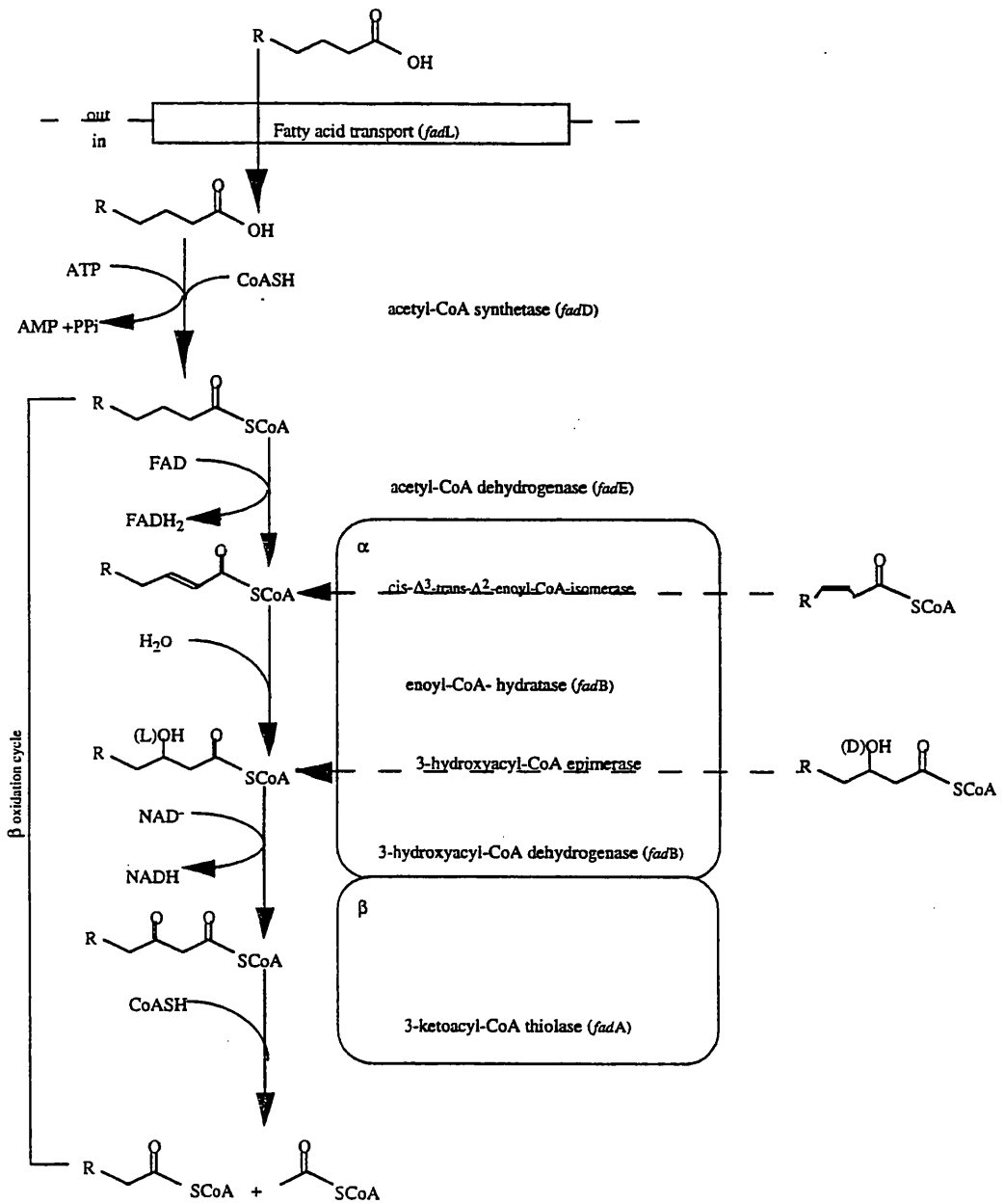


Figure 1.3 The fatty acid oxidation system of *E.coli*

Cyclic pathway of fatty acid degradation. principle enzymes of the pathway are shown on the right. Acetyl-CoA is further metabolised in the TCA cycle (Redrawn from Nunn, 1986)

fadA and *fadB* are transcribed together in the direction of *fadA* to *fadB* (Spratt *et al*, 1984). The gene products form a multi-enzyme complex which has five of the FAO enzyme activities associated with it. The complex has been purified and found to have a $\alpha_2\beta_2$ subunit structure where α=72 kDa *fadB* gene product and β=48 kDa *fadA* gene product (Binstock *et al.*, 1977; Pawar and Schultz, 1981; Pramanik *et al.*, 1979). The thiolase is associated with the 48 kDa subunit and the remaining four enzyme activities are associated with the larger, 78 kDa, subunit. Two of these activities, isomerase and epimerase, are only required during growth on unsaturated fatty acids.

Synthesis of the FAO enzymes is induced by the presence of LCFAs. Therefore only fatty acids longer than 11 carbons may be used as sole carbon source by wild-type *E.coli* and only after a distinct lag period. MCFAs can serve as substrates for the FAO enzymes but cannot induce the synthesis of these enzymes.

1.4.2 SCFA Metabolism

It appears that SCFAs are transported into the cell by a mechanism different from that of the FAO system. It is thought that acetoacetyl-CoA (AA-CoA) transferase may be involved in transport of SCFAs (Freeman, 1973) since this enzyme is associated with membranes and SCFA uptake is inhibited by butyryl-CoA and acetate, the products of the reaction catalysed by AA-CoA transferase.

The *ato* genes map together at 47 min on the *E.coli* chromosomal map. *atoD/atoA* encode acetyl-CoA:AA-CoA transferase, *atoB* encodes thiolase II and *atoC* encodes a regulatory protein. It appears that the *atoD* and *atoA* gene products, 26.5 kDa and 24 kDa proteins respectively, combine to form a tetrameric protein with an $\alpha_2\beta_2$

structure. Thiolase II is a tetrameric protein composed of four identical subunits, the 42 kDa gene product of *atoB*. Jenkins and Nunn (1987a and b) suggested that the *atoDAB* locus is transcribed as an operon and the *atoC* gene product acts as a transcriptional activator of this operon in the presence of acetoacetate.

The degradation of acetoacetate to acetyl-CoA is a two step reaction. Firstly acetoacetate is activated to acetoacetate -CoA catalysed by Acetyl-CoA:AA-CoA transferase, followed by cleavage of AA-CoA to yield acetyl-CoA which is catalysed by thiolase II.

Metabolism of SCFAs also requires the presence of at least two FAO enzymes and therefore wild-type *E.coli* cannot grow on SCFAs as the sole carbon source.

1.5 Acetate Metabolism

The bacterial plasma membrane is permeable for acetate by passive diffusion but it is not thought that acetate transport by diffusion is sufficient to explain the high uptake rate needed for growth on acetate, especially at low substrate concentration. Recent studies have suggested that an active transport mechanism may be involved. In *Azotobacter vinelandii* acetate is transported as the undissociated acid by an energy-linked potassium pump. Work on *Corynebacterium glutamicum* (Ebbighausen *et al*, 1991) suggested that a specific carrier is involved and uptake of acetate proceeds by a secondary proton-coupled mechanism and is competitively inhibited by propionate.

Once acetate enters the cell it can be activated by one of two mechanisms. One involves the acetylation of CoA and the cleavage of ATP to AMP, catalysed by acetyl-CoA synthetase (Berg, 1956). The second mechanism involves the

conversion of acetate to acetyl-phosphate with cleavage of ATP to ADP , catalysed by acetate kinase. Phosphotransacetylase then catalyses the transfer of the acetyl group from acetyl-phosphate to CoA, liberating inorganic phosphate.

1.6 The Glyoxylate Bypass

The major anaplerotic pathway operating during growth of a variety of organisms on acetate or fatty acids is the glyoxlate bypass. This pathway bypasses the decarboxylative steps of the TCA cycle (Fig 1.4).

The glyoxylate bypass involves two unique enzymes, isocitrate lyase (ICL) and malate synthase A (MS), which are induced when *E.coli* is grown on fatty acids or acetate as the sole carbon source (Kornberg, 1966b).

The glyoxylate bypass has been found in a wide range of organisms including bacteria, fungi, protozoa, nematodes, algae, gymnosperms and angiosperms. However for a long time it was not thought to exist in higher animals. Recently the cycle has been found in the liver of fetal guinea pigs and the urinary bladder of toads. The role of the cycle is not clear but it must have a role in some eukaryotic cells. In fungi and higher plants the enzymes are located in a microbody known as the glyoxysome and it is here that the glyoxylate bypass reactions take place.

In plants the glyoxylate bypass functions exclusively in the utilisation of storage lipids. Fat is stored in considerable amounts in seeds of a number of plants. When these seeds germinate the glyoxylate bypass is in operation. It is a process which allows the net synthesis of sugar compounds from fat stores at a fast rate and in a stage of the plant's life cycle where the photosynthetic apparatus is not yet available.

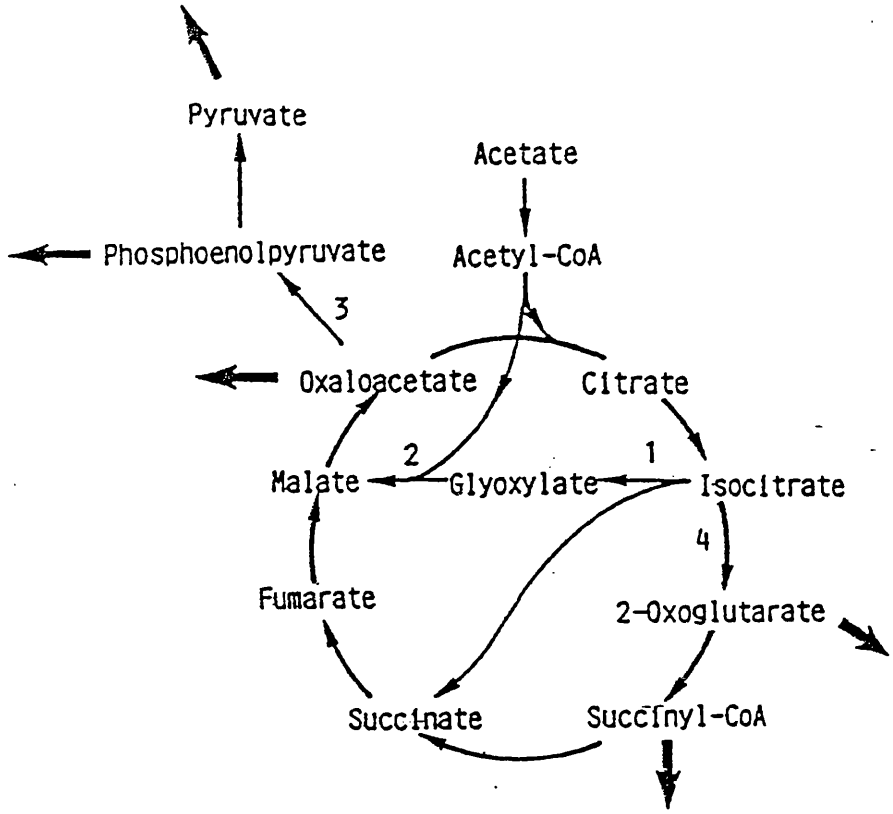


Figure 1.4 The Glyoxylate bypass and the TCA cycle

The reactions numbered 1-4 are catalysed by the enzymes isocitrate lyase, malate synthase-A, phosphoenolpyruvate carboxykinase and isocitrate dehydrogenase repectively. The heavy arrows indicate fluxes to biosynthesis.

At later stages, when the plant is fully developed, the enzymes of the glyoxylate cycle disappear.

Isocitrate lyase [Ds-isocitrate glyoxylate-lyase E.C.4.1.3.1.] was first discovered in sonic extracts of *Pseudomonas aeruginosa*, and was found to catalyse the reversible aldol cleavage of isocitrate to glyoxylate and succinate on addition of citrate or *cis*-aconitate (Campbell *et al.*, 1953) (fig 1.5a). It was later established that Ds-(+) isocitrate is the true substrate (Olson, 1954; Saz, 1954).

Malate synthase A [L-malate glyoxylate-lyase (CoA-acetylating) E.C.4.1.3.2.], the second enzyme in the pathway, was first observed in extracts of *E.coli* by Wong and Ajl (1956) and was shown to catalyse the condensation of glyoxylate and acetyl-CoA to form malate (fig 1.5b).

Kornberg (1966b) was the first to suggest that the two enzymes co-operate to form the glyoxylate bypass.

The overall reaction is as follows:-

acetyl-CoA + isocitrate + H_2O => succinate + malate + HS-CoA

The glyoxylate bypass results in the formation of two C_4 dicarboxylic acids, one is used to regenerate OAA and therefore ensure the continuation of the TCA cycle, and the other is used to supply the precursors required for biosynthesis. OAA can be converted to PEP, catalysed by PEP carboxykinase. The PEP can then be used to provide the cell with phosphorylated biosynthetic precursors. Other precursors are

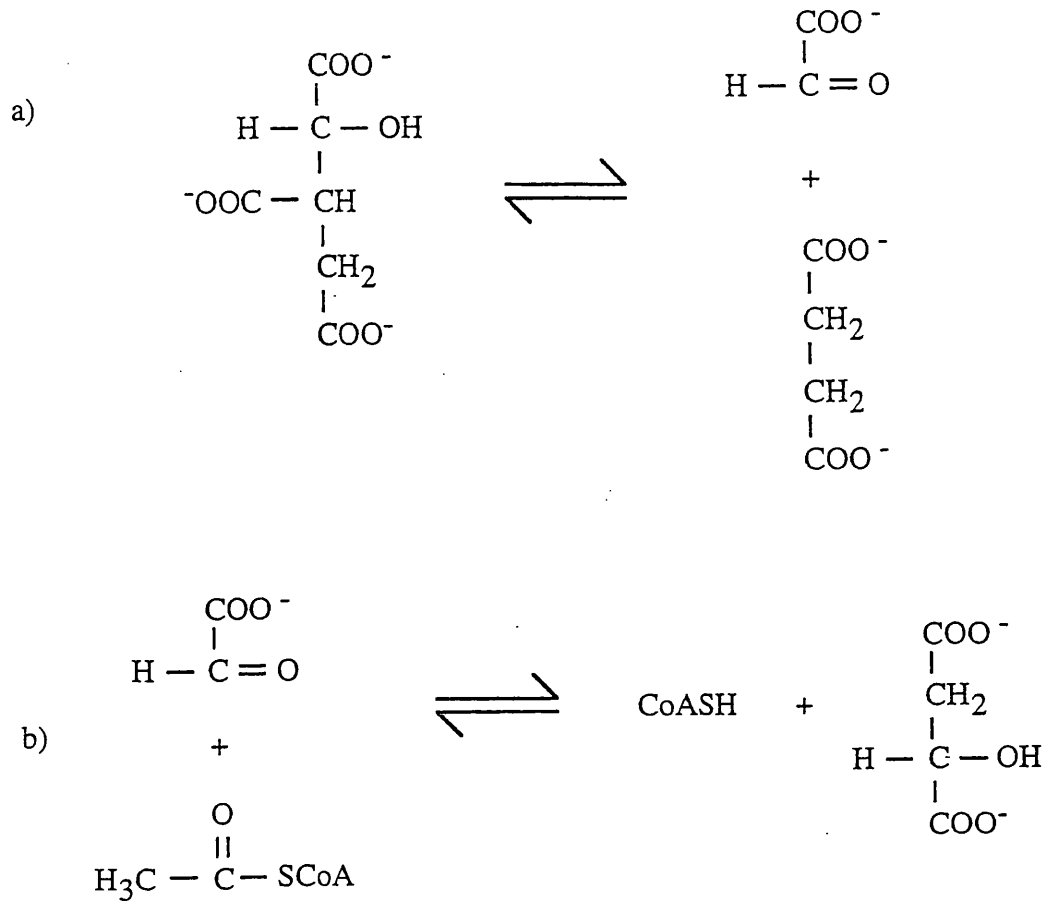


Figure 1.5 Reactions catalysed by Isocitrate lyase and Malate synthase

a) Isocitrate lyase catalyses the irreversible aldol cleavage of isocitrate to give succinate and glyoxylate.

b) Malate synthase-A catalyses the condensation of glyoxylate with acetyl-CoA to give malate and CoASH.

generated from 2-oxoglutarate and succinyl-CoA which are supplied by the TCA cycle.

1.6.1 Isocitrate Lyase

Isocitrate lyase is known to exist in a wide variety of organisms including bacteria, algae, ferns, gymnosperms, angiosperms, fungi, protozoa and nematodes (Cioni *et al.*, 1981) and there are a number of reports that ICL exists in vertebrates. The enzyme has been purified from a number of bacteria, fungi, plants and yeast. *E.coli* ICL has a subunit molecular weight of 47.2 kDa (Robertson and Reeves, 1987; Matsuoka and McFadden, 1988), which is similar to the values obtained for ICL from other prokaryotes (Chell *et al.*, 1978; McFadden *et al.*, 1968). The plant enzymes however, generally have a higher subunit value of between 62 and 70 kDa (Uchida *et al.*, 1986; Ruchti and Widmer, 1986). The native molecular weight of the *E.coli* ICL is approximately 180 kDa, suggesting that like most other ICLs it is a tetramer. Now that the amino acid sequence has been determined it is thought that the subunits are identical (Mackintosh and Nimmo, 1988).

ICL requires Mg^{2+} for catalysis. Hoyt *et al.* (1988) measured the activity of the enzyme when different ions were present and found that no activity was observed without any metal ion. Highest activity was found with Mg^{2+} followed by Mn^{2+}(54%) and Co^{2+}(17%). The enzyme also requires the presence of thiol reducing agents or DTT. It is thought that the catalytic mechanism probably involves acid-base catalysis. Nimmo *et al.* (1989) proposed that a cysteine residue of ICL donates a proton to the C-3 atom of isocitrate during cleavage, which results in the formation of glyoxylate and a succinate anion which can then be protonated to give succinate. The cysteine residue is situated in a highly conserved part of the protein

and is therefore thought to be part of the active site of the enzyme. In *E.coli* this highly conserved region occurs between residues 177 and 200 and includes two lysine residues, two histidine residues and the cysteine residue at position 195. Diehl and McFadden (1993) substituted lysine 193 with various amino acids and found that if the cationic nature of this region is not maintained the activity of the enzyme is lost confirming the idea that this region is extremely important.

Robertson *et al.* (1987a, b) suggested that ICL may be controlled by phosphorylation and this was later confirmed. Hoyt and Reeves (1988) found that ICL is phosphorylated *in vitro* in an ATP-dependent reaction. They showed that [^{32}P] was incorporated *in vivo* at a histidine residue and that this phosphorylation is essential for activity. This is unusual since often phosphorylation results in the inactivation of enzymes (e.g. isocitrate dehydrogenase). ICL in *Acinetobacter calcoaceticus* was also shown to be phosphorylated *in vivo* (Hoyt and Reeves, 1992a) and it was found that in both organisms the enzymes remain phosphorylated throughout the growth curve during growth on acetate (Hoyt *et al.*, 1994). It was originally thought that isocitrate dehydrogenase was solely responsible for control of partitioning of isocitrate, but this discovery of contrasting patterns of phosphorylation and enzymatic activity suggests that there may be a more complex control mechanism.

1.6.2 Malate Synthase

Malate synthase has been purified from a number of organisms including bacteria, plants, fungi and yeast. The bacterial and fungal enzymes have a molecular weight of 60-70 kDa and show a high level of similarity between amino acid sequences.

In many organisms two or more malate synthase isoenzymes have been observed. In *E.coli* two isoenzymes have been identified, MSA and MSG. MSA consists of 533 amino acids and has a molecular weight of 61 kDa. MSG has 723 amino acids and a molecular weight of 82 kDa. The two isoenzymes both catalyse the condensation of glyoxylate with acetyl-CoA to form malate and CoASH. They display very similar kinetic parameters for the substrate and are distinguished on the basis of chromatographic separation, thermostability and sensitivity to inhibitors such as glycollate and oxaloate.

MSA is induced during growth on fatty acids, acetate and other C_2 compounds whose metabolism does not proceed through PEP. The enzyme catalyses the condensation of glyoxylate, formed by the cleavage of isocitrate catalysed by ICL, with acetyl-CoA to form malate, as part of the glyoxylate bypass. In *E.coli* MSA is a soluble enzyme found in the cytosol. However, in plants, it is situated in the microbodies and therefore has a short C-terminal extension for import into the microbody.

The gene encoding MSA (*aceB*) is located upstream of the gene encoding ICL (*aceA*) at 90 min on the functional map of *E.coli* K12. They form part of a metabolic operon.

The MSA enzymes from *Saccharomyces cerivisiae* and *Pseudomonas ovalis* have an optimum pH of 8.5 and an absolute requirement for Mg^{2+}. Work by Durchschlag *et al.* (1981) on MSA from *S.cerevisiae* indicates that it is the binding of glyoxylate that is dependent on Mg^{2+} whilst that of acetyl-CoA is independent of Mg^{2+}. They also found that in *S.cerevisiae*, synthesis of the enzyme is induced by ethanol and repressed by glucose.

A number of amino acids in MSA are conserved between different organisms. In *E.coli* and yeasts, cysteine residues are conserved at positions 146, 496 and 617. In plants a serine residue is conserved at position 617. It is thought that the conserved cysteine 146 may be important in the active site.

The gene encoding malate synthase A in *Streptomyces arenae* has been cloned and sequenced and the enzyme purified (Huttner *et al.*, 1997). The 1632 bp open reading frame codes for a 61.36 kDa protein which shows similarity between amino acid sequences of bacteria, plants and fungi. The holoenzyme appears to be dimeric. All other known malate synthase from eubacteria are monomeric, while those from plants are oligmeric. The K_m value for glyoxylate is significantly higher than that of other known malate synthase enzymes.

MSG is induced during growth on glycollate and compounds which are metabolised to glyoxylate as the sole carbon source (Vanderwinkel and DeVlieger, 1968). This enzyme can substitute for MSA (Vanderwinkel *et al.*, 1963). When glycollate is metabolized it is first oxidised to glyoxylate, catalysed by glycollate oxidase. The glyoxylate can then undergo one of two reactions resulting in the formation of tartronic semialdehyde, catalysed by glyoxylate carboligase, or malate, catalysed by MSG.

In *E.coli*, MSG is monomeric but in *S.cerevisiae* the enzyme is homotrimeric and in germinating soybean cotyledons it is dimeric. MSG is encoded by *glcB* which is found at 64.5 min on the *E.coli* map next to the gene encoding glycollate oxidase and is transcribed as a monocistronic transcript. Three out of the nine consensus boxes defined for most MSAs are also conserved in MSG. The cysteine residue at position

496 is also conserved supporting the idea that it is important for the catalytic mechanism.

1.7 Isocitrate dehydrogenase

Isocitrate dehydrogenase(ICDH) has been purified from a number of bacteria and they can be classed into two distinct types based on quaternary structure. Most belong to type I which is typified by ICDH from *E.coli* and ICDH-I from *Vibrio* spp.ABE-1. They are homodimers with a subunit molecular weight of around 45 kDa and have similar protein sequences suggesting they are related evolutionarily. None of the enzymes are controlled allosterically by possible effectors such as AMP, ADP, ATP, $NAD(P)^+$ or intermediates of central metabolism.

Type II ICDHs are monomers with a molecular weight of around 80 kDa and have been found in *Vibrio* spp, *A.vinelandii* and *Rhodmicrobium vannielli*. *Vibrio* spp has also been found to contain one other ICDH isoenzyme but it is not known to which type it belongs. *A.calcoaceticus* also contains more than one ICDH isoenzyme, one is a type I, the other appears to belong to an entirely separate group and has a molecular weight of 300 kDa.

The regulation of type I ICDH in *E.coli* has been well studied, as described earlier (1.7.1), and has been shown to play a key role in controlling flux between the TCA cycle and the glyoxylate bypass. Both subunits are involved in catalysis and both are subject to control by phosphorylation. Other type I ICDHs have conserved residues involved in binding of isocitrate and Mn^{2+} and it is therefore thought that they may have similar control mechanisms.

Little is known about the control of type II ICDH. Work has been carried out on *Vibrio* spp which has both type I and type II ICDH's but the role of each enzyme in cell metabolism is not known. During growth on succinate, the activity of ICDH-II is 5.6 times higher than that of ICDH-I (Ishii *et al.*, 1987). However, during growth on acetate the activity of ICDH-I is 3 times higher than that on glucose, whilst the activity of ICDH-II remains constant. The ratio of ICDH-I to ICDH-II is 1:1.5 and this is due to increased synthesis of ICDH-I during growth on acetate. It is assumed that these changes in ICDH levels play a role in the control of metabolism during growth on different carbon sources. Many eukaryotic organisms have several forms of ICDH, some of which are specific to certain organelles and have different control mechanisms.

ICDH from *Streptomyces coelicolor* has been purified and the gene encoding it has been cloned and sequenced (Taylor, 1992). The protein is monomeric with a molecular weight of 80 kDa and the N-terminal sequence is 69% identical to that of ICDH-II from *Vibrio* spp and 63% identical to that of *R.vannielli*. There was no cross reaction with an anti-*E.coli* ICDH antiserum and it is therefore thought that the enzyme can be classed as a type-II ICDH (Taylor, 1992).

1.8 The Branchpoint between ICL and ICDH

The glyoxylate bypass generates some reducing power, but not enough to satisfy the cell's needs. A lot more is obtained by the complete oxidation of acetyl-CoA to CO_2. Also, the production of 2-oxoglutarate and succinyl-CoA, intermediates of the TCA cycle, is essential. Therefore, during growth on acetate or fatty acids, ICL of the glyoxlate bypass competes with ICDH, the TCA cycle enzyme, for the common

substrate, isocitrate. In *E.coli*, the control of flux of carbon at this branchpoint is tightly regulated in order to ensure efficient use of the available acetate.

Britten (1954) observed that acetate accumulates in the medium of *E.coli* cultures growing on glucose. When the glucose is exhausted, oxidation of the accumulated acetate occurs. This is concurrent with increased activity of the glyoxylate bypass enzymes and decreased activity of ICDH to 25% of its original value (Holms and Bennet, 1971). After acetate was exhausted the activity of ICDH increased. Bennet and Holms (1975) observed that the specific activity of ICDH from acetate-grown *E.coli* was 30% of that found on glucose- or pyruvate-grown cultures. When 1 mM pyruvate was added to an acetate-grown culture the specific activity of ICDH increased until the pyruvate was exhausted. Addition of chloramphenicol before adding pyruvate did not stop the increase in ICDH activity and they therefore concluded that *de novo* synthesis of ICDH was not responsible for increased activity since chloramphenicol inhibits protein synthesis. It was proposed that a reversible covalent modification of the enzyme or reversible association with another molecule may be responsible for the differing levels of activity. It was also suggested that the role of inactivation was to restrict flux through the TCA cycle and so permit isocitrate to be metabolised through the glyoxylate bypass.

Phosphorylation of ICDH was first implicated in the control of ICDH by the work of Garnak and Reeves (1979a,b) when they added [^{32}P] to an *E.coli* culture grown on limiting glycerol in a low phosphate medium. On addition of acetate to the medium, there was a partial reduction in ICDH activity. A radioactive protein co-purified with ICDH, it was found to be immunologically identical to ICDH and it was suggested that the protein was [^{32}P] ICDH. Partial hydrolysis of the purified protein revealed that [^{32}P]phosphate was incorporated into a serine residue in ICDH. It had been

generally assumed previously, that protein phosphorylation catalysed by protein kinases did not occur in bacteria.

Borthwick *et al.* (1984a) purified ICDH from glycerol-grown cells and the inactive and partially inactive forms of ICDH from acetate-grown cells. Using non-denaturing gel electrophoresis they showed that the proteins had differing mobilities which was due to a difference in charge, rather than molecular weight. They showed that active ICDH is dephosphorylated, whilst the inactive ICDH is phosphorylated and the two differ in one phosphate group per subunit. Partially-inactive ICDH is thought to be a mixture of the two forms.

Borthwick *et al.* (1984b) showed that phosphorylation of ICDH results in complete inactivation of the enzyme, which is unusual since many other phosphorylated enzymes show only partial loss of activity.

During growth on acetate approximately 70% of ICDH is maintained in the inactive, phosphorylated form (LaPorte and Koshland, 1983; Borthwick *et al.*, 1984a).

The K_m of ICL is higher than that of ICDH for isocitrate (Nimmo *et al.*, 1987). Therefore, when intracellular levels of isocitrate are low, ICL cannot compete for available substrate, so it is metabolised by ICDH and channelled through the TCA cycle (Nimmo *et al.*, 1984). When *E.coli* is grown on glucose the intracellular level of isocitrate is too low to detect (El-Mansi *et al.*, 1985) and the glyoxylate bypass does not operate. However, during growth on acetate, ICDH is partially inactivated by phosphorylation which allows the concentration of isocitrate to rise and ICL can then compete with ICDH for the available substrate (Holms, 1987).

Laporte and Koshland (1982) attempted to purify ICDH kinase, the enzyme responsible for phosphorylation of ICDH. To their surprise they found that the enzyme responsible for dephosphorylation of ICDH co-purified and the ratio of kinase to phosphatase activity remained constant throughout the purification. Fractions from the affinity column which contained both kinase and phosphatase activities gave a single band of 66 kDa on SDS PAGE gels and they concluded that the two activities were associated with the same polypeptide. This was confirmed by LaPorte and Chung (1985) when they isolated a clone which restored growth on acetate to an *aceK* mutant. The clone expressed a 66 kDa protein which had both kinase and phosphatase activities. Nimmo *et al* (1984) reported the native MW to be 135 kDa suggesting that ICDHK/P exists as a dimer. Laporte and Chung (1985) suggested that there may be two active sites on the protein to account for the opposing activities but it was also suggested that there may be only one active site which underwent a conformational change in order to catalyse both reactions.

Sequencing of the *ace*K gene encoding ICDHK/P revealed a gene of 1731 nucleotides encoding a polypeptide of 577 amino acid residues with a molecular weight of 66.5 kDa (Chung *et al.*, 1988; Klumpp *et al.*, 1988; Cortay *et al.*, 1988). The deduced amino acid sequence revealed none of the consensus motifs characteristic of other protein kinases (Celenza and Carlson, 1986) with the exception of a consensus ATP binding site. Steuland *et al.* (1989) showed that both activities use the same ATP binding site. They changed an "invariant" lysine residue, found in every other protein kinase to date and involved in ATP binding, to a methionine. This resulted in the inhibition of both kinase and phosphatase activities.

ICDHP has an absolute requirement for ATP or ADP (LaPorte and Koshland, 1982; Nimmo *et al.*, 1984). Since [^{32}P] released from [^{32}P]ICDH appears as inorganic phosphate rather than [^{32}P]ATP, the dephosphorylation cannot simply be the back reaction of the kinase. ICDHK catalyses the transfer of the γ-phosphate group of ATP to a serine residue of ICDH with a stoichiometry of one phosphate per subunit. ICDHP catalyses the release of inorganic phosphate from phospho-ICDH causing full reactivation of ICDH in the presence of inhibitors of the kinase and either ATP or ADP. Since ICDHK and ICDHP can be active simultaneously *in vivo* (Borthwick *et al.*, 1984) the phosphorylation state of ICDH is representative of the steady state balance of ICDHK and ICDHP activities.

Wild-type ICDHK/P is maintained in massive excess over the level required for steady state phosphorylation of ICDH (Steuland *et al.*, 1989). This allows rapid responses to changes in the available carbon source and is therefore not excessive during metabolic transitions. Under these conditions, this enzyme represents the primary rate limiting step (Laporte, 1993).

1.8.1 Control of Phosphorylation of ICDH

It is believed that the role of phosphorylation and inactivation of ICDH is to diminish the flux control coefficient of ICDH which in turn allows the intracellular concentration of isocitrate to increase to a level high enough to maintain flux through ICL (El-Mansi *et al.*, 1984).

ICDHK/P is controlled by a wide variety of metabolites. Isocitrate, pyruvate, OAA, AMP, ADP, 3-phosphoglycerate (3-PG), 2-oxoglutarate and PEP activate ICDHP and inhibit ICDHK, while NADP, citrate, fructose 6-phosphate and glyoxylate

inhibit ICDHK but have little effect on ICDHP (Nimmo and Nimmo, 1984). During growth on acetate, isocitrate, PEP, 3-PG, NADPH, AMP and ADP are the most important *in vivo* in control of the ICDH phosphorylation cycle (Nimmo and Nimmo, 1984). It is thought that they act as general indicators of the levels of metabolic intermediates and thus the need for isocitrate to be directed to the glyoxylate bypass. AMP is also important in the activation of ICDH as AMP inhibits ICDHK and activates ICDHP. The energy requirements of the cell can be monitored through AMP. Therefore when energy levels are low, there is a high level of AMP, ICDHP is activated and the activity of ICDH increases forcing isocitrate through the TCA cycle.

Borthwick *et al.* (1984b) confirmed that ICDH is inactivated by ICDHK/P *in vivo*, that the phosphorylation state and activity of ICDH are inversely related and that the reversible activation and deactivation of the enzyme could be accounted for by its dephosphorylation and phosphorylation.

It was suggested that phosphorylated ICDH could not bind $NADP^+$ (Garland and Nimmo, 1984; Borthwick *et al.*, 1984b,c) and this may have some effect on its inactivation. They proposed that phosphorylation of ICDH occurs at or near the $NADP^+$ binding site of the enzyme so that the introduction of a negatively-charged phosphate group could prevent binding of $NADP^+$ by charge repulsion and also cause a conformational change similar to that caused by binding of $NADP^+$. However, Dean *et al.* (1989) showed that the phosphorylated enzyme was unable to bind isocitrate but retains the ability to bind $NADP^+$ and NADPH.

Phosphorylation of ICDH from *E.coli* occurs at a single serine residue (Borthwick *et al.*, 1984b,c; LaPorte and Koshland, 1983; Nimmo *et al.*, 1984) and inactive ICDH

contains one phosphoserine group per subunit. It was found that this residue corresponds to Ser-113 (Borthwick *et al.*, 1984c; Malloy *et al.*, 1984). Studies were carried out to investigate what the effect of substitution of this serine residue would on the activity of the enzyme.

Site directed mutagenesis of serine 113 was carried out (Thorness and Koshland, 1987; Dean *et al.*,1989). It was found that if the serine residue was replaced by negatively-charged aspartate the enzyme became completely inactive but if it was substituted with lysine, threonine, cysteine, tyrosine or alanine, the enzyme remained active. It was concluded that the negative charge of aspartate or phospho-serine inactivates the enzyme and it is the negative charge rather that stearic hindrance which prevents binding of isocitrate.

Hurley *et al.* (1990a) used X-ray crystallographic studies to show that replacing serine with aspartate or glutamate (both negatively charged) resulted in very little conformational change in the protein and yet resulted in rendering the protein inactive. They also showed that isocitrate binds within an inter-domain pocket and Ser-113 lies at the edge of this pocket. Ser-113 forms a hydrogen bond to the g-carboxy group on isocitrate and this is the reason for its importance.

It can be seen that the system responds efficiently to a number of different effectors and because some of the metabolites affect both opposing activities, this amplifies the sensitivity of the system. The overall outcome is that flux of isocitrate through the glyoxylate bypass and the TCA cycle can be tightly regulated during growth on acetate or fatty acids as the sole carbon source.

Other examples of opposing enzymic activities encoded on the same polypeptide chain have been reported and they all play major roles in the regulation of intermediary metabolism.

1.9 The *ace* operon

The genes encoding malate synthase (*aceB*) and isocitrate lyase (*aceA*) were found close to each other at 90 min on the functional map of *E.coli* K12 and it was suggested then that they may form an operon (Brice and Kornberg, 1968) since they are co-ordinately expressed and show close genetic linkage. The gene *metA* was shown to be alongside (Vanderwinkel and De Vlieghere, 1968) and the gene order was found to be *metA-aceB-aceA*. The *ace* genes appeared to be under the negative control of an adjacent gene, *iclR*, since mutations in this gene lead to constitutive expression of isocitrate lyase (ICL) (Brice and Kornberg, 1968).

The first evidence that the genes encoding the glyoxylate bypass enzymes may form an operon came from Maloy and Nunn (1982). They created transposon Tn10 insertions which can be polar and prevent expression of genes downstream of the point of insertion in an operon. They found that *aceA*::Tn10 insertions eliminated ICL activity, but not malate synthase (MS) activity. When *aceB*::Tn10 insertions were constructed in a *glc*- background, ICL and MS activities were eliminated. They concluded that *aceA* and *aceB* form an operon with transcription in the direction *aceB* to *aceA*. They also confirmed the gene order to be *metA-aceB-aceA-iclR*. *AceB* mutants were examined in a *glc*- background since the gene product of *glc* is MSG which can be substituted for MSA.

LaPorte and Chung (1985) showed that *aceK*, the gene encoding ICDHK/P was also closely linked and therefore may be part of the *ace* operon. They confirmed this by creating *ace*::Mu insertions and placed *aceK* downstream of *aceA*. Inclusion of *aceK* in the *ace* operon provides a mechanism for co-ordinating its expression, and therefore the activity of ICDH, with that of the glyoxylate bypass enzymes.

The *ace* operon has been cloned from *E.coli* by several groups (El Mansi *et al.*, 1987; Chung *et al.*, 1988) and the sequences of *aceA* (Rieul *et al.*, 1988), *aceB* (Byrne *et al.*, 1988; Matsuoka and McFadden, 1985), *aceK* (Cortay *et al.*, 1988) and *iclR* (Sonnarborg *et al.*, 1990; Negre *et al.*, 1991) have all been determined. Chung *et al.* (1988) identified the promoter by deletion mapping, S1 mapping and sequencing upstream of the start site and showed that the operon is expressed from a single promoter located upstream of *aceB*, during growth on acetate (fig 1.6).

AceB and *aceA* are separated by 32 base pairs which contains no structural motifs. However, there is an intergenic region of 184 nucleotides between *aceA* and *aceK* which contains two consecutive, long, dyad symmetries. Each gene is preceded by a ribosome binding site (RBS) (Shine and Dalgarno, 1974): AGAGG for *aceB*, GGAG for *aceA* and GAGG for *aceK*.

In *Salmonella typhimurium* the *ace* operon has been identified (Wilson and Maloy, 1987). The genes encoding the glyoxylate bypass enzymes are arranged in the same order as in *E.coli* and are transcribed as an operon from a single promoter.

During growth on acetate the cellular level of ICL is up to 1000 fold greater than that of ICDHK/P (Laporte and Chung, 1985). This is consistent with a striking

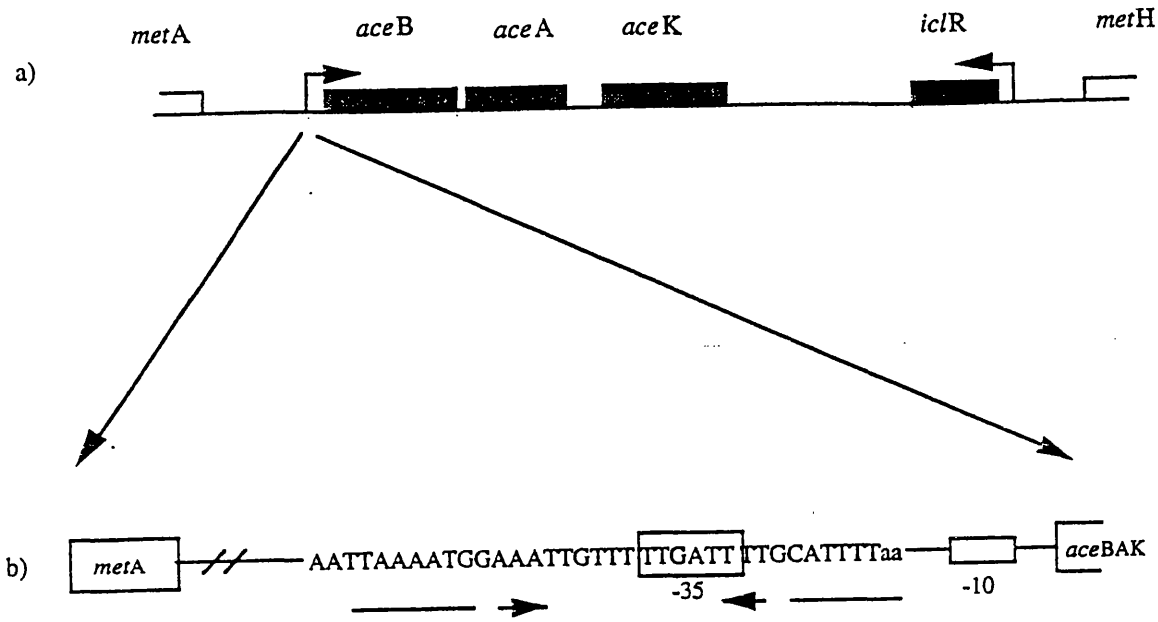


Figure 1.6 The *ace* operon of *E.coli*

a) A diagrammatic representation of the ace operon of *E.coli*. The *ace*A, *ace*B and *ace*K genes are transcribed from the same promoter and the *icl*R gene is transcribed in the opposite direction from its own promoter.

b) A detailed representation of the operator/promoter region of the ace operon. The two arrows represent the imperfect palindromes to which the ICL repressor protein binds.

downshift in expression between *aceA* and *aceK* even though they are part of the same operon (Cortay *et al.*, 1988). Chung *et al.* (1993) created both gene and operon fusions with the *lacZ* gene to characterise differential expression of the glyoxylate bypass operon genes by measuring β-galactosidase activity. Relative cellular levels of the products of *aceA, aceB and aceK* were found to be 0.3 : 1 : 0.003 respectively. The upshift in expression between *aceB* and *aceA* was shown to be due to differences in translational efficiency.

The *aceA-aceK* intercistronic region contains several repetitive extragenic palindromic (REP) elements. A number of stem-loop structures are possible within this region which also contains the Shine-Dalgarno RBS. It was first thought that these stem-loop structures may contribute to the downshift in expression between *aceA* and *aceK*. However, deletion of these sequences had no effect on expression of *aceK* but instead, halved the expression of *aceA* (Klumpp *et al.*, 1988). It was then suggested that the stem-loop structures did not inhibit translation but instead, stabilised the message upstream by preventing 3' exonucleolytic digestion of the message (Cortay *et al.*, 1989). However, it has now been shown that inefficient translation and premature transcriptional termination are responsible for the downshift in expression between *aceA* and *aceK*. The sequences responsible for inefficient expression of *aceK* lie within its RBS. The *aceK* RBS (GAGG) was replaced by that of *aceA* (GGAG) and this resulted in a dramatic increase in expression of β-galactosidase in an *aceK* gene fusion. The downshift in expression between *aceA* and *aceK* undoubtedly reflects the fact that MS and ICL are metabolic enzymes while ICDHK/P is a catalytic regulatory protein and therefore not required in large amounts by the cell.

1.10 Control of the *ace* Operon

1.10.1 Metabolic Regulation

The mechanism of control of the *ace* operon is complex and the metabolites which regulate its expression have yet to be identified. The regulation may be similar to that for ICDHK/P in that it responds to a number of effectors. Evidence suggests that the glyoxylate bypass is not subject to catabolite repression (Kornberg, 1966a; Wilson and Maloy, 1987) and must therefore depend on the presence of inducers. The intracellular signal for induction of the glyoxylate bypass remains to be elucidated. Several observations suggest that the expression of the *ace* operon does not simply respond to the presence of acetate or acetyl-CoA. When bacteria are grown simultaneously on acetate and a preferred carbon source such as glucose, operation of the glyoxylate bypass does not take place. During growth on fatty acids as the sole carbon source the glyoxylate bypass is induced even though the utilisation of fatty acids does not proceed through acetate.

1.10.2 IclR and FadR

The operation of the glyoxylate bypass is not only controlled by the phosphorylation state of ICDH, but also at the level of gene expression. Brice and Kornberg (1968) were the first to suggest that the *ace* genes were under the negative control of an adjacent gene, *iclR*. In addition, the *fadR* gene, which maps at 25 min on the *E.coli* linkage map and is involved in the regulation of fatty acid metabolism (Simons *et al*, 1980), is also involved in control of the *ace* operon (Maloy *et al.*, 1980).

Simons *et al.* (1980) found that FadR is a diffusable repressor protein of subunit mass 29 kDa (DiRusso and Nunn, 1985) with two functional domains which act in *trans* to control both the *fad* and the *fab* regulons (Maloy and Nunn, 1981; Nunn, 1986). The protein is involved in the negative regulation of fatty acid regulation (*fad* regulon) and the positive regulation of fatty acid biosynthesis (*fab* regulon).

Mutations in the *iclR* or *fadR* genes result in elevated expression of the enzymes of the glyoxylate bypass under normal inducing conditions. Merodiploid studies also showed that *iclR* and *fadR* act in a *trans*-dominant manner suggesting that the gene products are repressor proteins. Maloy and Nunn (1982) found that the two proteins, IclR and FadR, can act independently to cause partial repression of the operon, but when acting together they cause its full repression.

This dual repressor system results in different levels of expression of the *ace* operon depending on the carbon source and therefore confers a metabolic advantage to the cell. During growth on fatty acids a lot more energy is produced and therefore more carbon can be directed through the glyoxylate bypass so the *ace* operon is highly expressed. However, during growth on acetate, less energy is produced so there is greater flux through the TCA cycle and the *ace* operon is expressed to a lesser extent.

Sunnarborg *et al.* (1990) cloned *iclR* from *E.coli* by taking advantage of its tight linkage with the glyoxylate operon. They cloned a 0.9 kb fragment which complemented an *iclR* mutant. The 855 bp gene is located downstream of *aceK* and is transcribed in the opposite direction to the *ace* operon. The gene encodes a 274 amino acid protein of molecular weight 29741. The gene encoding IclR has also been cloned from *Salmonella typhimurium* (Galinier *et al.*, 1991) and found to have 89% identity with that of *E.coli* (Negre *et al.*, 1991). A putative helix-turn-helix motif

has been recognised at the N-terminal of IclR which is characteristic of prokaryotic transcriptional regulatory proteins.

Cortay *et al.* (1991) purified IclR to homogeneity from *E.coli* in a one-step procedure by cationic exchange chromatography after ammonium sulphate fractionation. They investigated the specific interaction of the protein with the operator/promoter region of the *ace* operon using gel retardation and DNase I footprinting. They found that the repressor protein recognises a 35 bp palindromic sequence which overlaps the -35 recognition site of RNA polymerase.

It is thought that the repressor protein and the transcriptional enzyme compete for the same binding site on the DNA, reinforcing the concept that repression is a matter of blocking access to the promoter. However, no protection of the -10 region has been detected.

Several observations suggest that the expression of the *ace* operon does not simply respond to the presence of acetate or acetyl-CoA. Therefore Cortay *et al.* (1991) investigated the effects of various metabolites on the binding of IclR to the promoter of the *ace* operon. They found that acetate, acetyl-CoA, pyruvate and OAA had no effect and only PEP impaired the formation of the complex between IclR and the operator/promoter region and this metabolite is therefore thought to be a good candidate as an inducer of transcription of the operon.

Negre *et al.* (1992) carried out further binding studies on IclR and found that IclR is a dimer which contacts the operator on both strands of the DNA helix and makes specific interactions with two nucleotide domains arranged in a palindrome-like

structure. It was suggested that each monomer of the protein binds to one half of the palindrome-like structure.

1.11 Growth of Streptomycetes on acetate

ICL was purified from *S.coelicolor* by Chapman (1994) and was found to be produced during growth on Tween. The gene encoding ICL (*icl*) was also cloned and found to be located immediately upstream of the gene encoding malate synthase. This is the opposite way round compared to *E.coli*. The presence of ICL and MS indicate the glyoxylate bypass may be present in *S.coelicolor*.

Malate synthase has also been purified and the gene cloned from *S.arenae* (Huttner *et al.*, 1997), indicating that the glyoxylate bypass is used by this organism for growth on C_2 compounds.

However, this is not the case for all Streptomycetes. A novel anaplerotic pathway has been identified in *S.collinus* (Han and Reynolds, 1997). Two genes have been identified, *ccr* and *meaA* which exhibit strong similarity to *adhA* and *meaA*, respectively, of *Methylobacterium extoquens*. *AdhA* and *meaA* are involved in the assimilation of C_1 and C_2 compounds in an unknown pathway in the ICL-negative *Methylobacterium*. Growth studies suggested that *ccr* and *meaA* are involved in a novel pathway for growth of *S.collinus* when acetate is its sole carbon source.

1.12 Aims of the Project

This chapter has described how the glyoxylate bypass is involved in growth of *E.coli* on fatty acids or acetate as the sole carbon source. Complex regulatory mechanisms

control the flux of acetyl-CoA between the glyoxylate bypass to supply precursors for biosynthesis, and the TCA cycle to produce energy.

Polyketides are an important group of antibiotics that are derived from acetyl-CoA. They are made by Gram-positive bacteria, *Streptomyces* spp. If the mechanisms involved in controlling flow of acetyl-CoA in *Streptomyces* can be deduced, it may be possible to control this flux and produce antibiotics more efficiently.

Work has already been carried out on the glyoxylate bypass in *S.coelicolor*. ICDH has been purified and found to be a type II ICDH (Taylor, 1992). This class of enzyme has control mechanisms different to the type I ICDH found in *E.coli*. The gene encoding ICDH in *S.coelicolor* has been cloned and sequenced (Taylor, 1992). The gene encoding ICL has also been cloned from *S.coelicolor* and the amino terminal portion of a gene thought to encode malate synthase was found immediately downstream (Chapman, 1994).

The aim of this project was to study the glyoxylate bypass in *S.coelicolor* further by cloning the entire gene, evaluating whether it does encode malate synthase and sequencing the regions upstream and downstream in an attempt to locate other genes involved in the glyoxylate bypass.

CHAPTER 2

Materials and Methods

2.1 Introduction

This chapter contains the general procedures used in the experiments which were the basis of this thesis. The chapter is divided into three main sections for convenience; bacterial strains, vectors and chemicals (2.2), microbiological techniques and standard media(2.3) and general DNA techniques (2.4).

2.2 Bacterial strains, vectors and chemicals

2.2.1 Bacterial strains

The bacterial stains used are listed below (Table 2.1):

STRAIN	GENOTYPE	REFERENCE/SOURCE
Escherichia coli strains		
TG1	*sup*E, *hsd*AEΔ, *thi*, AE(*lac-pro*AB), F' [*tra*D36, *pro*AB+, *lac*Iq, *lac*ZAEM15].	Gibson (1984)
LE392	*sup*E, *sup*F, *hsd*R.	Murray *et al* (1977)
***Streptomyces coelicolor* strains**		
GLW1147	SCP1+, SCP2+	G.Hobbs, UMIST

Table 2.1 Bacterial strains

2.2.2 Plasmid and bacteriophage vectors

The plasmid pUC18 was obtained from Pharmacia Biotech Ltd. (St. Albans, UK). Plasmid pBluescript^{TM}II (KS+) was obtained from Stratagene (La Jolla, Ca., USA) pM13mp18 and M13mp19 are bacteriophage cloning vectors from which single-stranded DNA may be isolated for DNA sequencing. They were obtained from Pharmacia Biotech Ltd. The *Streptomyces coelicolor* genomic library (Taylor, 1992) was constructed using the λGEM-11 replacement vector supplied as "*BamHI* arms" by Promega Corporation. Plasmid pICL92 is a pBluescript-derived plasmid containing the wildtype isocitrate lyase gene and part of the gene encoding malate synthase from *Streptomyces coelicolor* (Chapman, 1994). pICL92B and pICL92C are derived from pICL92.

2.2.3 Chemicals and biochemicals

Ampicillin and ethidium bromide were obtained from Sigma Chemical Co. (Poole, UK). Bactotryptone, yeast extract and Bactotryptone (agar) were obtained from Difco (Detroit, USA).

ATP, DTT, TEMED (N,N,N',N'- tetramethylethylene diamine), and Tris buffer were obtained from Boehringer Mannheim (Lewes, UK).

DMSO, phenylalanine, polyethylene glycol 8000, and tryptophan were obtained from BDH Chemicals (Poole, UK).

Agarose, IPTG (isopropyl-β-thiogalactoside), phenol (ultrapure), and X-gal (5-bromo-4-chloro-3-indolyl-β-galactoside) were obtained from BRL (Gibco Ltd., Paisley, UK) Acrylamide, bisacrylamide and SDS were obtained from FSA Laboratory Supplies (Loughborough, UK).

All other chemicals were of analytical reagent grade and were obtained from one of the following suppliers: BDH Ltd.; Formachem Ltd. (Strathaven, UK); FSA Laboratory Supplies; Koch-Light Ltd. (Haverhill, UK); Sigma Chemical Co.

2.2.4 Enzymes, and kits

All restriction enzymes and T4 DNA ligase were obtained from BRL and Promega Corporation.

The sequencing kit used was the Sequenase^{TM} sequencing kit (USB Biochemicals, La Jolla, Ca., USA).

Wizard ^{TM} Lambda preps DNA purification system was supplied by Promega.

Dynabeads^{TM} were supplied by Dynal Ltd.

2.3 Standard media and microbiological techniques

2.3.1 Media used for growth of *S.coelicolor*

All growth media were sterilised by heating to 120^{o}C for 15 mins in an autoclave. Supplements and buffer solutions were heated to 108^{o}C and $CaCl_2$ to 114^{o}C for 10 mins. Heat-labile solutions, such as antibiotics, were sterilized by filtration through Nalgene 0.22 μm pore membranes (Nalge Co., New York, USA).

2.3.1 Complex media

a) Soya Mannitol Agar (SM)

This was used as a general plating medium for *Streptomyces*, particularly for production of spores. It consists of 2% (w/v) mannitol, 2% (w/v) soya bean flour, and 1.6% (w/v) agar, made up in tap water.

b) Yeast extract-Malt extract (YEME)

This was the only liquid medium used for growth of *Streptomyces*. It consists of 0.3% (w/v) Difco yeast extract, 0.5% (w/v) Difco bacto peptone, 0.3% (w/v) Oxoid malt extract, 1% (w/v) glucose, 34% (w/v) sucrose, made up in distilled water. To prevent pelletting, and produce a well-dispersed growth, $MgCl_2$ was added to a final concentration of 5mM, and glycine to a final concentration of 0.5% (w/v).

2.3.2 Media for propagation of *E.coli*

Chemicals of good quality were used in the preparation of the growth media and solutions (AnalaR grade when available). The sources of many of the chemicals varied during the course of this work. The most common suppliers were BDH Chemicals Ltd.; Difco Laboratories (Detriot, Michigan, USA) and Sigma Chemical Co. Ltd.

a) L-broth

1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl, 0.1% (w/v) glucose, 0.002% (w/v) thymine, made up in distilled water and adjusted to pH 7.0 with NaOH.

b) L-agar

As L-broth, but without glucose and with the addition of 1.5% (w/v) bacto-agar.

c) 2xYT medium

1.6% (w/v) bacto-tryptone, 1% (w/v) bacto-yeast extract, 0.5% (w/v) NaCl, made up in distilled water and adjusted to pH 7.0 with NaOH.

d) Top agar

As L-agar, but with only 1% (w/v) bacto-agar.

e) NZCYM Top agar

1% (w/v) NZ amine, 0.5% (w/v) NaCl, 0.5% bacto-yeast extract, 0.1% (w/v) casamino acids, 0.2% $MgSO_4.7H_2O$, 1% (w/v) bacto-agar, made up in distilled water and adjusted to pH 7.0 with NaOH.

f) Minimal agar

Agar was made to 1.75% with water. To 75 ml of this, 25 ml of D+M salts were added and glucose and thiamine were added to final concentrations of 2 mg/ml and 20 μg/ml, respectively. Other supplements were added if necessary.

g) Davis and Mingoli (D+M) Salts (x4)

2.8% (w/v) K_2HPO_4, 0.8% (w/v) KH_2PO_4, 0.4% (w/v) $(NH_4)_2SO_4$, 0.1g trisodium citrate, $MgSO_4.7H_2O$, made up in distilled water.

2.3.3 Growth of *Streptomyces* mycelia in liquid media

Cultures were typically grown in 500 ml conical flasks containing 100 ml of medium at 30°C on an orbital shaker at 200 rpm.

Growth experiments employed YEME complex medium which produced reproducible and rapid growth. Spores from frozen suspensions or suspensions that were freshly-prepared from a frozen slope were used to inoculate the medium. Cells grown on YEME could be harvested after 48-72 hours and stored indefinitely as a cell pellet at -20°C.

2.3.4 Harvesting of mycelia

After growth, mycelia were diluted in an equal volume of distilled H_2O (dH_2O), recovered from the media by centrifugation (10mins, 10,000g) and resuspended in dH_20.

2.3.5 Growth of *E.coli*

Liquid cultures of *E. coli* strains from which plasmids were to be isolated were grown in L broth with the appropriate antibiotic selection (typically ampicillin at 50μg/ml). The volume of broth inoculated depended on the quantity of plasmid required. Routinely, 1.5 ml and 100 ml cultures were used for small and large scale plasmid preparations, respectively (sections **2.4.2.3** and **2.4.2.2** respectively). For the preparation of competent cells, liquid cultures of *E. coli* TG1 were grown in 2xYT (section **2.3.2**). To maximise aeration of the culture, the volume of the Ehrlenmeyer flask used was at least five times that of the broth. All cultures were incubated at $37^{o}C$ in an orbital shaker at 250 rpm.

2.3.6 Production of *S.coelicolor* spores

2.3.6.1 Preparation and storage of suspensions

Concentrated spore suspensions were required for inoculating liquid cultures of S.*coelicolor* or *S.lividans*. The protocol described by Hopwood *et al.* (1985) was followed with minor modifications:

A boiling tube containing a slant of SM agar (produced by pouring 20 ml of molten agar into the tube and allowing it to solidify with the tube held at 5 degrees from the

horizontal) was inoculated with 150 µl of a spore or mycelial fragment suspension and incubated at 30°C. After 5-10 days the surface of the culture was covered in a dark grey mass of spores. The slant was then sealed using parafilm and frozen at -20°C. Spores could be harvested immediately or stored indefinitely at -20°C. The spores were harvested by adding 5 ml of dH_2O to the frozen slant and rubbing the surface of the slant with a 10 ml glass syringe. Contaminating agar or mycelial fragments were removed by a single passage through a cotton wool filter, as described in Hopwood *et al.* (1985). The filtered spore suspension was then either used fresh to inoculate YEME media or frozen at -20°C.

2.3.6.2 Spore counts

Colony forming units were determined by plating suitably diluted spore samples on Soya plates. Counts of the number of colonies were made after incubation at 30°C for 5 days.

2.3.7 Preservation of E.coli strains

E.coli strains were stored in glycerol. An 800 µl aliquot of an overnight culture was mixed with an equal volume of 40% (v/v) glycerol, 2% peptone (w/v) and frozen at -70°C. The strains were revived by scraping the surface of the frozen suspension with a toothpick and either inoculating liquid broth or streaking onto agar to isolate a single colony.

2.3.8 Introduction of plasmid DNA into *E.coli*

2.3.8.1 Preparation of competent cells

a) $CaCl_2$ method

An overnight culture of the recipient strain was diluted 1 in 100 into 30 ml L-broth and incubated for 90-120 mins to a density of approximately 10^8/ml cells (OD_{600} 0.4-0.6). The cells were harvested using a centrifuge (12,000g, 5 min, 4^oC) and resuspended in 10 ml of ice-cold 50 mM $CaCl_2$. The cells were pelleted again, resuspended in 1ml of ice-cold 5 mM $CaCl_2$ and either kept on ice for at least 15 min before use, or stored in 100 µl aliquots following the addition of 200 µl 100% glycerol.

b) Hanahan method

For cells with high transformation efficiencies, the following steps were carried out (Hanahan, 1983). Cells were prepared and harvested as above but were then resuspended in 2.5 ml of ice cold TFB (10 mM MES/KOH pH 6.3, 100 mM RbCl, 45 mM $MnCl_2$, 10 mM $CoCl_2$, 3 mM hexaminecobaltic chloride) and incubated on ice for 15 mins. 100 µl of DMSO was then added and the cells incubated on ice for 5mins, followed by the addition of 100 µl of 2.25 M DTT, 40 mM potassium acetate (pH 6.0) and the cells incubated on ice for a further 10 minutes. Finally, 100 µl of DMSO was added, and the cells were stored on ice and used on the day of preparation.

2.3.8.2 Transformation procedure

Transformations were carried out in sterile 1.5 ml microfuge tubes. An aliquot (maximum 10 µl) of ligation mix or plasmid DNA was added to 100 µl aliquots of competent cells and the mixture was incubated on ice for at least 30 minutes. The DNA/cell mix was heat shocked at 42°C for 2 minutes before being placed back on ice. The cells were then plated onto L broth plates containing the appropriate antibiotic/chromogenic substances and incubated overnight at 37°C.

2.3.8.3 Transfection with bacteriophage M13

This procedure is as above, except that, after heat-shock, the transfected cells and 100 µl exponentially-growing (plating) cells were added to 3 ml 0.6% soft agar containing 10 µl IPTG (stock 24 mg/ml) and 50 µl X-gal (stock 20 mg/ml). The mixture was plated onto L-agar plates, which were incubated at 37°C overnight.

2.3.8.4 Selection of pUC-derived recombinant clones

a) Ampicillin

Stock solutions (20 mg/ml made up in water) were added to molten agar (cooled to 55°C) to a final concentration of 50 µg/ml.

b) X-gal (5-bromo-4-chloro-3-indolyl-β-galactosidase)

This was used in conjunction with IPTG to identify E.coli strains containing pUC-based or M13mp18/19 vectors with inserts in their multiple cloning sites. Recombinants containing inserts are generally white, while those lacking inserts are blue. X-gal was stored at a concentration of 20 mg/ml in dimethylformamide (DMF) at -20°C while IPTG was stored at aconcentration of 24 mg/ml in dH_2O at -20°C. X-

gal and IPTG were added to L-agar plates to a final concentration of 20 μg/ml and 50 μg/ml, respectively.

2.4 General DNA methods

2.4.1 Commonly used buffers

a) TE buffer (10x)

100 mM Tris-HCl pH 8.0, 10 mM EDTA. Sterilised using an autoclave and stored at room temperature. It was used as a 1x solution for most applications.

b) Phage buffer (1x)

20 mM Tris-HCl pH 7.4, 100 mM NaCl, 10 mM $MgSO_4$.

c) TAE Buffer (10x)

4.84% (w/v) Tris, 1.64% (w/v) Sodium acetate, 0.36% (w/v) $Na_2EDTA.2H_2O$, made up in distilled water, pH adjusted to 8.2 with glacial acetic acid.

d) TBE buffer (10x) pH8.3

10.9% (w/v) Tris, 5.5% (w/v) boric acid, 0.93% (w/v) $Na_2EDTA.2H_2O$ made up in distilled water.

e) Agarose gel loading buffer (10x) pH 7.4

0.5% (w/v) bromophenol blue, 0.05% (w/v) xylene cyanol, 50% (w/v) ficoll, 1% (w/v) SDS, 100 mM EDTA.

f) λ-Hind III DNA markers

λclind 1 ts857 sam7 DNA was obtained from BRL. This DNA was cleaved with the restriction enzyme *Hind*III and resulting DNA fragments diluted to a final concentration of 27 ng/μl in TE (final concentration 1x) with loading buffer added to 1x concentration. Typically, 10 μl was used on agarose gels as markers for

comparing the size and concentration of bands in samples.Sizes of fragments produced: 23,130 bp; 9,416 bp; 6,557 bp; 4,361 bp; 2,322 bp; 2,027 bp and 564 bp.

2.4.2 Preparation of plasmid DNA

2.4.2.1 Reagents for isolation of plasmid DNA

Protocols based on the alkaline lysis method (Birnboim and Doly, 1979) were used for the isolation of plasmid from small (5ml) or large (50-200ml) cultures of *E. coli.*

a) Birnboim Doly I (BDI)

50 mM glucose, 25 mM Tris-HCl pH 8.0, 10 mM EDTA.

b) Birnboim Doly II (BDII)

0.2 M NaOH, 1%(w/v) SDS which was stored in a plastic container.

c) Birnboim Doly III (BDIII)

5 M Potassium acetate pH 4.8; prepared by mixing equal volumes of 3 M KCH_3COO and 2 M CH_3COOH.

d) DNAse-free RNAse

Pancreatic RNAse (RNAse A) was dissolved at a concentration of 10 mg/ml in dH_2O, heated to 100^oC for 15 mins and allowed to cool slowly to room temperature. The RNAse was then aliquoted and stored at -20^oC.

e) TES buffer

25 mM Tris-Cl pH 8.0, 25 mM EDTA, 10% sucrose

f) Acid phenol/chloroform

50 g phenol was dissolved in 50 ml chloroform and 10 ml dH_2O. Hydroxyquinoline was added to prevent oxidation of the phenol and its colour helps in separation of the phases.

g) Alkaline SDS

0.3 M NaOH, 2%SDS.

2.4.2.2 Large-scale plasmid preparation from *E.coli*

a) Caesium chloride gradient

100 ml cultures of stationary phase cells were harvested using a centrifuge (12,000 g, 5 min at 4°C). The pellet was resuspended in 8 ml of Birnboim-Doly I solution and incubated at room temperature for 5 mins. 16 ml of Birnboim-Doly II solution was added and the solution left on ice for 5-10 mins, before 12 ml of cold Birnboim-Doly III solution was added. The suspension was mixed gently and left on ice for 15-30 mins. The cell debris and most of the chromosomal DNA were removed by centrifugation (32,000g, 5 mins at 4°C). The remaining nucleic acid was precipitated by the addition of 0.6 volumes of isopropanol and then harvested by centrifugation (39,200 g, 15 mins).

The nucleic acid pellet was washed with 70% (v/v) ethanol. The plasmid DNA was further purified by equilibrium density centrifugation on a caesium chloride/ethidium bromide (CsCl/EtBr) gradient. The nucleic acid pellet was redissolved in 1 ml of dH_2O and 4.5 g of CsCl dissolved in 3.5 ml of dH_2O. The DNA and CsCl solutions were combined with 250 ml of EtBr (10 mg/ml), creating a solution with a density of 1.58 g/ml. The nucleic acid-CsCl solution was spun in a Beckman Ti70 angled rotor at 289,000 g for 16 hours at 20°C. Two bands were visible in the gradients after centrifugation, a lower supercoiled plasmid band and an upper chromosomal and relaxed plasmid DNA band.

The lower band was removed using a 1ml syringe and the EtBr removed by repeated extractions with water-saturated butanol. After dilution with 3 volumes of dH_2O, 9 volumes of absolute ethanol were added. The precipitate was pelleted by centrifugation (27,000 g, 4°C for 30 mins). The resulting plasmid pellet was washed twice with 70% (v/v) ethanol and dried *in vacuo* before being redissolved in 1 ml dH_2O. This procedure yielded very large amounts of pure plasmid DNA (up to 1 mg from *E. coli* cultures) suitable for all *in vitro* manipulations.

b) PEG precipitation

The protocol was the same as that for a), up to the resuspension of the DNA, in dH_2O and CsCl. Instead the DNA was resuspended in 3 ml of dH_2O. 3 ml of ice-cold 5 M LiCl was added and mixed. The precipitated DNA was removed by centrifugation (12,000 g, 4°C for 20 mins). The pellet was resuspended in 500 µl 1x TE buffer (containing 10mg/ml RNase), and incubated at room temperature for 30 mins. 500 µl of 1.6 M NaCl containing 13% PEG-8000 was added, the solution mixed, and centrifuged (12,000 g, 4°C for 5 mins). The pellet was washed with 70% EtOH, before being dried under vacuum and redissolved in dH_2O.

2.4.2.3 Small-scale plasmid preparation from *E.coli*

Routinely, plasmids were isolated from 1.5 ml of *E. coli* cultures. The cells were pelleted by centrifugation in a 1.5 ml microfuge tube (12,000 g for 30 secs) and resuspended in 100 µl of BDI, containing lysozyme at a concentration of 1 mg/ml, using a vortex mixer. This was followed by the addition of 200 µl of BDII and repeated inversion of the microfuge tube to thoroughly mix the suspension. Immediately afterwards, 150 µl of pre-chilled BDIII was added to the viscous bacterial lysate, mixed gently on the vortex mixer and placed on ice for 5-10 mins.

The cell debris and most of the chromosomal material were harvested by centrifugation (12,000 g, 4°C for 10 min) in a microfuge. The supernatant was transferred to a fresh tube and extracted with half volumes of phenol/chloroform and chloroform. The nucleic acid was then precipitated by the addition of 2 volumes of ethanol and allowed to stand at room temperature for 5 mins. The precipitate was harvested by centrifugation in a microfuge (12,000 g, 4°C for at least 15 mins). The resulting pellet was rinsed twice with 70% (v/v) ethanol before it was allowed to dry by leaving the tube open on the bench. The nucleic acid was then resuspended in 50 µl dH_2O containing DNAse-free RNAse (20 µg/ml).

If further purification of the DNA was required (for example for double-stranded sequencing reactions), the pellet was resuspended in 16 µl of 1x TE, to which 4 µl of 4 M NaCl and 20 µl of 13% PEG were added, followed by incubation on ice for 20 mins. Centrifugation for 15 mins at 4°C produced a pellet, which was washed in 70% ethanol, and resuspended in 1x TE.

2.4.3 Preparation of "total" DNA from *S.coelicolor*

50 mg fresh *Streptomyces* culture was resuspended in 500 µl lysis solution (containing 2 mg/ml lysozyme) and incubated at 37°C for 30 mins until the cells became translucent, followed by the addition of 250 µl 2% SDS and mixing by vortex for at least 1 min until the viscosity of the mix decreased noticably. Cell debris was removed by extraction with phenol/chloroform and centrifugation (12,000 g, 4°C for 2 mins). Such extraction was repeated until no interface remained (a minimum of three extractions). DNA was then precipitated by the addition of 0.1 vol of 3 M Sodium acetate and 1 vol isopropanol, with 5 mins incubation at room temperature followed by centrifugation (12,000 g, 4°C for 2 mins). The DNA was

resuspended in 1x TE buffer with 25 µl 100 mM spermidine-HCl and incubated at room temperature for 5 mins followed by centrifugation (12,000 g, 4°C for 2 mins). The pellet was resuspended in 300 µl 0.3 M sodium acetate,100 mM $MgCl_2$ and 700 µl ethanol. This was incubated at room temperature for 1 hr followed by centrifugation (12,000 g, 4°C for 2 mins). The DNA was resuspended in 500 µl 1xTE buffer and stored at 4°C.

2.4.4 Organic solvent extraction

Protein was removed from DNA solutions by phenol/chloroform extraction. An equal volume of TE-saturated phenol was added to samples which were then mixed by vortexing and centrifuged in a microfuge for 1-5 mins. The upper aqueous phase was removed to a fresh microfuge tube and the process repeated this time using phenol/chloroform (1:1 v/v). Finally, traces of phenol were removed by extraction with an equal volume of chloroform in an identical manner. Precipitation with ethanol or isopropanol removed any remaining solvent.

2.4.5 Precipitation of DNA using ethanol or isopropanol

DNA solutions were precipitated by the addition of 1/50 volume of 5M NaCl and 2 volumes of cold ethanol or an equal volume of isopropanol. After mixing, the DNA was pelleted by centrifugation (27,000 g, 4°C, 30 mins for volumes of 7.5-20 ml, or 12,000 g, 4°C, 15 mins for small volumes in microfuge tubes). The pellet was washed in 70% (v/v) ethanol and dried briefly in a vacuum desiccator, or in an open tube on the bench.

2.4.6 Spectrophotometric measurement of nucleic acid

Nucleic acid concentrations were determined spectrophotometrically at 260 nm. In a 1 cm path length an absorbance value of 1.0 corresponds to 50 μg/ml for double stranded DNA and 33 μg/ml for single stranded DNA.

2.4.7 Digestion of DNA with restriction enzymes

Restriction digests were carried out using the BRL restriction enzymes and REact buffers which were provided with each batch of enzyme. There are ten different REact buffers with a range of salt concentrations, each one suitable for a range of enzymes. Alternatively, digests were performed using Promega restriction enzymes and the accompanying restriction buffers. Analytical digests were carried out in a volume of 20 ml at 37^{o}C. Preparative digests were carried out in larger volumes. When DNA was digested with two restriction enzymes, the endonuclease requiring the lower salt buffer was used first. After the recommended duration of digestion, the salt concentration was adjusted and the second enzyme added.

2.4.8 Ligation of DNA fragments

The ligation of DNA fragments was carried out usually at a DNA concentration of 6 mg/ml. The molar ratio of insert fragment to vector was 3:1, when the vector could not ligate to itself (for example when using a vector that has been dephosphorylated or had been cut with two non-complementary enzymes). A molar ratio of 10:1 was used when the ends of the vector could ligate to each other. Ligations were performed usually in 10 μl of 1x ligation buffer provided by BRL, containing 1 U of

T4 ligase per μg of DNA. The reactions were incubated for 4 hours at room temperature or overnight at 16°C.

2.4.9 Removal of the 5' phosphate from linearised DNA

10x CIP Buffer: 200 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 10 mM $ZnCl_2$ and 0.5 mg/ml bovine serum albumin.

Procedure: Calf Intestinal Alkaline phosphatase (CIP) was used to remove the 5' phosphate from DNA. Around 5 pmoles of 5'-terminal phosphorylated DNA with 5' protruding ends (approximately 7 μg of a 5 kb molecule) were incubated in 1x CIP buffer, containing 0.1 U of CIP at 37°C for 30 mins. The reaction was terminated by heating to 65°C in 1x gel loading buffer for 10 mins. The 5'-terminal dephosphorylated DNA was recovered from an agarose gel after electrophoresis.

2.4.10 Agarose gel electrophoresis

DNA was visualized on horizontal neutral agarose gels. Although 0.8% (w/v) gels were most commonly used, 1-2% (w/v) gels were occasionally used to separate fragments of <1.5 kb. Gels were routinely prepared and run in TAE buffer. λ-Hind III markers were used on all gels as size markers and for quantification of the amount of DNA by comparing the intensity of bands to those of the samples (section **2.4.1**f).

a) Mini gels

BRL model H6 gel kits were used for the rapid analysis of DNA after digestion with restriction enzymes or precipitation steps. 0.16 g agarose was added to 20 ml of 1x TAE, boiled, and then cooled to 60°C. EtBr was added to 200 ng/ml and the molten

agarose poured into a 7.6 cm x 5.1 cm gel caster with an 8 well slot former (4.1 x 0.8 mm wells). After the gel had set, the slot former was removed and the gel placed in the tank with 500 ml of 1xTAE.

Depending on the time available and the level of resolution required, the DNA was separated by electrophoresis for 30-60 mins with an applied voltage of 2-10 V/cm. 1/10 volume loading buffer (section **2.4.1**e) was added to the DNA samples before they were loaded onto the gel. The separated DNA molecules were visualised on a 302 nm UV transilluminator.

b) Large gels

200 ml gels were also used to ensure good separation of DNA fragments for accurate sizing and/or Southern analysis. They were made by pouring 200 ml of molten agar, containing 200 mg EtBr, into a 16.5 x 23 cm gel former with a 14 space slot former. The gels were run overnight at 20 V in 1 litre of 1x TAE buffer in gel tanks with a capacity of 3 litres. DNA samples were mixed with 1/10 volume of 10x loading buffer (section **2.4.1**e) before loading onto the gel.

2.4.11 Photography of agarose gels

Gels stained with ethidium bromide were viewed on a 302 nm UV transilluminator and photographed using Polaroid type 67 land film or using a Pentax 35 mm SLR loaded with Ilford HP5 film. Both cameras were fitted with Kodak Wratten filters (No. 23A).

2.4.12 Recovery of DNA from agarose

a) Using SPIN-X tubes

SPIN-X tubes were obtained from Costar UK Ltd. These tubes contain a cellulose acetate membrane, which allows the passage of buffer and DNA, but stops agarose from passing through.

The DNA band was excised from the gel and placed in the upper chamber of the SPIN-X tube. This was then placed at -70^{o}C for 5 minutes, thawed by placing at 37^{o}C for 5 minutes, followed by centrifugation at 12000g for 5 mins. The filtrate contained the DNA, in a state which can be used immediately.

b) Using Qiaex-suspension

Qiaex gel extraction kit was obtained from Qiagen, Germany. The gel slice containing the excised DNA fragment was mixed with 3 vols QX1 solubilisation buffer and Qiaex-suspension (10 µl per 5 µg DNA). The gel slice was incubated at 50^{o}C for 10 mins with frequent mixing to solubilise the gel and allow the DNA to absorb to the Qiaex-suspension. The suspension was pelleted, washed twice with 500 µl QX2 wash buffer, and twice with 500 µl QX3 wash buffer. The suspension was pelletted and air-dried for 10-15 mins, followed by resuspension in 1x TE buffer, 5 mins incubation at room temperature and pelletted again. The DNA was contained in the supernatant, in a condition ready for any subsequent manipulations.

2.4.13 Techniques with *E.coli* bacteriophage λ

2.4.13.1 Preparation of plating bacteria for infection with bacteriophage λ

5 ml of L-broth was inoculated with a single colony of *E.coli* LE392 and grown overnight at 37°C on an orbital shaker. 1ml of this overnight culture was used to inoculate 100 ml L-broth and incubated at 37°C for approximately 3 hrs, until the $O.D._{600}$=0.5. The cells were pelleted by centrifugation (4,000 g, 4°C, 10 mins) and resuspended in sterile, ice cold 10 mM $MgSO_4$ to $O.D._{600}$=2 (approx. 10^{10} cells/ml). The cells were stored at 4°C and remained viable for 3 weeks.

2.4.13.2 Infection of bacteriophage λ, plating and titre

Serial 10-fold dilutions of λ phage stock were prepared in phage buffer (section **2.4.1b**). Bacteriophage λ infection was achieved by adding 10 µl aliquots of each dilution to 200 µl (1.5 x 10^8 cells) of a suspension of plating bacteria. The samples were incubated at 37°C for 20 mins. 3ml of NZCYM top agar (section **2.3.2e**) at a temperature of 45°C was added and the mixture was poured onto plates containing bottom L-agar. The plates were left to stand for 5minutes at room temperature to allow the top agar to harden and then incubated at 37°C overnight. The plaques were counted and the titre determined for each dilution assayed.

2.4.13.3 Isolation of bacteriophage particles from a plaque

The plaque of interest was stabbed out of the plate using the narrow end of a sterile glass Pasteur pipette to form a plug of agar. The plug was left in 1 ml of phage buffer containing 70 µl of DMSO (to kill any cells) for two hours at room

remperature (or overnight at 4°C) to allow bacteriophage particles to diffuse out of the agar. An average plaque yielded 10^6-10^7 infectious bacteriophage particles, which could be stored indefinitely at -70°C in phage buffer/DMSO without loss of viability.

2.4.14 Techniques for handling E.coli bacteriophage M13

2.4.14.1 Preparation of single-stranded DNA from M13

The single-stranded M13 templates were prepared as described in the "M13 Cloning/Dideoxy sequencing Instruction Manual" published by Bethesda Research Laboratories.

a) Minipreparations

A single M13 plaque was used to infect 1.5 ml of 2xYT broth containing 15 µl of an overnight culture of *E.coli* TG1. This culture was grown at 37°C for 5-6 hours with vigorous shaking, then transferred to a microfuge tube and harvested by centrifugation at room temperature for 5 mins. The supernatant, containing the phage particles was recovered and respun. The remaining supernatant was mixed with 200 µl of a solution of 20% (w/v)PEG, 2.5 M NaCl and left to stand at room temperature for 15 mins to precipitate the phage particles. These were recovered by centrifugation at room temperature in a microfuge for 15 mins.

The supernatant was discarded, the pellet respun and all traces of supernatant removed. The pellet was then resuspended in 100 µl dH_2O and extracted twice with phenol/chloroform and twice with chloroform. The DNA was then precipitated from

the aqueous phase with sodium acetate and ethanol, and recovered by centrifugation in a microfuge. The single stranded (ss) DNA was then washed with 70% (v/v) ethanol, dried in *vacuo* and redissolved in 20µl dH_2O.

b) Maxipreparations

The same overall procedure was followed as for the "minipreparations", except that all the volumes were scaled up 20-fold. The cells from a 30 ml culture were spun out (14,000 g for 2 mins), the supernatant recovered and respun as before. The supernatant (20 ml) that remained was then precipitated with 5 ml of 20% (w/v) PEG, 2.5 M NaCl for 10 mins at room temperature and the phage harvested by centrifugation (14,000 g at 20^{o}C for 15 mins). The phage pellet was resuspended in 1 ml of dH_2O, then reprecipitated and processed as for the minipreparations with the volumes scaled up accordingly.

2.4.15 Labelling of DNA with γ-^{32}P or Digoxygenin

2.4.15.1 Labelling of DNA using the "random priming" technique

The method of "random primed" DNA labelling is based on the hybridisation of a mixture of all possible hexanucleotides to the DNA to be labelled. The complementary strand is synthesised from the 3'-hydroxyl termini of the random hexanucleotide primer, with Klenow enzyme (DNA polymerase I). A random priming kit from Boehringer Mannheim, UK., was used for this purpose. Protocols were followed as per the manufacturer's instructions.

2.4.15.2 Removal of unincorporated radionucleotide

Unincorporated label from random priming reactions was removed by gel filtration chromatography on a 20x1 cm Sephadex G-50 column. Sephadex G-50 was hydrated in 1x TE and poured columns equilibrated in 1x TE. The reaction mixture was mixed with an equal volume of Blue Dextran dye in 1x TE and loaded directly on top of the column. The radioactivity was monitored as it passed down the column and as it approached the bottom, fractions were collected manually. Labelled probe was co-eluted first with the Blue Dextran dye, followed by a trough of radioactivity, and then a second peak corresponding to the unincorporated label.

2.4.15.3 Labelling of DNA with digoxygenin-dUTP (DIG-dUTP) using the "random priming" technique

The method of "random primed" DIG DNA is essentially the same as for labelling with ^{32}P (section **2.4.15.1**), except that the labelling reaction was carried out overnight. A random priming kit from Boehringer Mannheim, UK., was used for this purpose. Protocols were followed as per the manufacturer's instructions.

2.4.16 Southern blotting

After electrophoresis and photography, the resolved DNA fragments were transferred under alkali conditions to nylon membrane (adapted from Southern, 1975), as described in "Blotting and hybridization protocols for Hybond-N^{TM} membranes" (published by Amersham International plc).

2.4.16.1 Reagents

a) Denaturing solution

1.5 M NaCl, 0.5 M NaOH.

b) Alkali transfer buffer

1.5 M NaCl, 0.25 M NaOH.

c) 20x SSC

3 M NaCl, 0.3 M tri-sodium citrate.

2.4.16.2 Procedure

The gel was rinsed in distilled water, placed in enough denaturing solution to immerse it completely and left for 30 mins to chemically denature the DNA contained within the agarose. The gel was then equilibrated for 10-15 mins in alkaline transfer buffer. The DNA was transferred to the nylon membrane in transfer buffer, by capillary action (disposable nappies proved a particularly useful absorbent material for driving the transfer process).

After blotting for at least 4 hours (but usually overnight), the membrane was washed briefly in 2x SSC to remove any adhering agarose. The DNA was fixed to the membrane either by UV crosslinking using a Stratagene "Stratalinker", or by baking at $80^{o}C$ for 2 hours.

2.4.17 Hybridisation of DNA probe to filter-bound nucleic acid

2.4.17.1 Prehybridisation

Nylon filters were not pre-wetted, but were placed directly into a hybridisation tube containing prehybridisation solution (6x SSC, 0.05% (w/v) sodium pyrophosphate, 200 μg/ml 0.5% (w/v)heparin, 0.05% (w/v) SDS); 50 ml of the prehybridisation solution was used per cm^2 of filter surface area. The tube was placed in a hybridisation oven and the filter prehybridised for at least 4 hours at 68^oC.

Alternatively, filters were prehybridised in QuickHyb solution (Stratagene), for 1 hr at 68^oC. A minimum of 33 μl of Quickhyb solution was used per cm^2 of filter surface area.

2.4.17.2 Hybridisation and washing

After prehybridisation, the prehybridisation solution was replaced by hybridisation solution. The hybridization solution was the same as the prehybridization solution, except that it contained 0.5% (w/v) SDS and the salt concentration was varied according to the conditions required (see results of individual hybridisations).

50 μl of the solution was used per cm^2 of filter surface area. 15 ng of labelled probe (section **2.4.15**) was added to the tube. Hybridisation conditions were as described for each individual hybridisation.

Alternatively, when Quickhyb was used, 15 ng of the labelled probe was denatured in a boiling water bath for 10 mins with salmon sperm DNA , which acts as a

blocking reagent, then added directly to the prehybridisation solution and hybridisation allowed to take place at 68°C for at least 2 hrs.

Hybridised filters were washed twice in large volumes of buffer at ionic strengths appropriate to the experimental conditions. The temperature of this buffer was also varied experimentally. After washing, the filters were left damp, wrapped in Saran wrap and exposed to Fuji RX film using intensifying screens at either -70°C (for radiolabelled probes) or room temperature (for DIG-labelled probes). Films were developed by a Kodak X-OMAT processor.

2.4.18 Screening of plasmid clones by colony hybridisation

Recombinant pUC-based clones were screened using a modification of the method described in the Hybond-N protocol manual. Nylon filters were placed on duplicate agar plates containing the selective antibiotic. Bacterial colonies were crossed onto a master plate (containing antibiotic) and then onto the nylon filters (onto which alignment marks had been made). The plates were inverted and grown overnight at 37°C.

The filters were removed and placed colony side up on a pad of absorbent filter paper soaked in denaturation solution (1.5 M NaCl, 0.5 M NaOH) and left for 7 mins. The filters were then transferred, colony side up, to a pad of filter paper soaked in neutralising solution (1.5 M NaCl, 0.5 M Tris-HCl pH 7.2, 1 mM EDTA), and left for 3 mins. This step was repeated with a fresh pad soaked in the same solution. The filters were then washed briefly in 2x SSC, transferred to dry filter paper and allowed to dry in air, colony side up. Finally, the filters were baked at 80°C for 2 hrs, or UV-

crosslinked. The filter was then hybridised with a nucleic acid probe (section **2.4.17**).

2.4.19 Screening of a bacteriophage λ library

2.4.19.1 Primary screening

Cells from a prepared bacterial suspension (section **2.4.13.1**) were infected with phage from the bacteriophage λ library at a multiplicity of 10^4 pfu/10^8 cells. $4x10^3$ pfu were plated onto 10x10 cm petri dishes using 8 ml of 0.6% (w/v) top agar. The plates were incubated overnight at 37°C.

Up to six impressions could be taken from one plate onto nylon membranes, provided that sufficient time was allowed for fresh phage to diffuse to the top agarose surface (Sambrook *et al.*, 1989). Alignment marks were made on the plate and on the filter using a syringe needle. Filters were removed and treated as described for colony filters (section **2.4.18**).

2.4.19.2 Secondary screening

Single plaques were isolated as described in section **2.4.13.3.** A lawn of bacteria (10^8 cells) were plated onto a 9 cm circular petri dish and once the top agar had set, phage particles which had been soaked from the agar plug were streaked onto the surface of the plate. The plates were then grown up overnight at 37°C. Single plaques were isolated from each plate (section **2.4.13.3**) and the phage particles were streaked in duplicate onto Hybond-N. The filters were baked at 80°C for 2 hrs to fix the DNA and then hybridised with a radiolabelled probe (section **2.4.17**).

2.4.20 Hybridisation, and detection of DIG-labelled DNA hybrids

Prehybridisation and hybridisation of filters with DIG-labelled random primed DNA was carried out as described in the Digoxygenin protocols manual from Boehringer Mannheim, UK.

DIG-bound DNA was detected using AMPPD as a chemiluminescent substrate for alkaline phosphatase. Again, all protocols were followed as per the manufacturer's instructions.

2.4.21 Stripping filters

Filters were boiled for approximately 30 mins in 0.1% SDS solution, and the efficiency of stripping detected by autoradiography overnight. If the autoradiogram was clear when developed, the filter was suitable for reprobing.

2.4.22 DNA Sequencing techniques

Single-stranded sequencing

Dideoxy sequencing (Sanger *et al.*, 1977) was carried out on single-stranded M13 templates using a Sequenase^{TM} kit (supplied by United States Biochemical Corporation). All sequencing strategies used deoxy-7-deazaguanosine triphosphate (dc^7GTP) as a replacement for dGTP, to reduce sequence compressions (Mizosawa *et al.*, 1986).

Sequencing with T7 polymerase

Extension, labelling and termination reactions were performed as suggested by the manufacturers. ^{35}S-dATP was used for labelling.

2.4.23 Denaturing polyacrylamide gel electrophoresis for DNA sequencing

A BRL sequencing unit (Model S2) was used for high voltage polyacrylamide gel electrophoresis.

2.4.23.1 Preparation of polyacrylamide gels

6% (w/v) denaturing polyacrylamide gels were used for sequencing. The gels were prepared from the following stock solutions:

40% (w/v) acrylamide stock	9 ml
urea	30 g
10x TBE	6 ml
dH2O	21 ml

The urea was dissolved by heating the mix to 37^{o}C and then cooled to room temperature. The sequencing gel solution could be stored at 4^{o}C for several weeks without loss of resolution. Before pouring the gel, 300 µl of freshly prepared 10% (w/v) ammonium persulphate and 50 µl of TEMED were added to 50 ml of the stock solution.

2.4.23.2 Preparation of glass plates and pouring of the gel

The plates (40 cm x 33 cm) were cleaned thoroughly with water and alcohol, and assembled using spacers (0.4 mm thick) along the vertical sides and 3MM Whatman paper along the bottom of the gel. The entire assembly was held in place by clamps. The gel solution was poured from a beaker down one edge of the plates while tilting the plates at an angle of approximately 30 degrees. The plates were then laid at an angle of 5 degrees and the sharks tooth combs inserted in an inverted fashion. The gel polymerized usually within 30 mins at room temperature.

2.4.23.3 Electrophoresis of sequencing gels

The gel was pre-electrophoresed for 30 mins at a constant power of 60 W. Prior to loading, the samples containing sequencing loading buffer were heated to 95^{o}C for 5 mins, placed on ice and loaded onto the gel. The gels were run for approximately 2 hrs to read the first 100 nucleotides, 3.5 hrs to read the next 100 nucleotides and 5.5 hrs to read up to 350 nucleotides.

2.4.24 Purification of Lambda DNA

Lambda DNA was purified using Wizard^{TM} lambda preps DNA purification system (supplied by Promega). Protocols were followed as per the manufacturers instructions.

2.4.24.1 Lysate preparation

Phage lysates were prepared from liquid cultures. 500 µl of a fresh overnight culture of *E.coli* LE392 grown in NZCYM medium, supplemented with 0.2% (w/v) maltose and 0.01 M $MgSO_4$. was mixed with 10^7 pfu of lambda bacteriophage and incubated at 37°C for 20 mins. The infected culture was transfered to a 250 ml Erlenmeyer flask containing 100 ml NZCYM medium supplemented with 1 M $MgSO_4$. This was incubated at 37°C with agitation until cell lysis occured (approx 5 hours) which could be recognised by the medium clearing. 500 µl chloroform was then added and incubated for a further 15 mins.

The lysate was centrifuged at 8,000 g for 10 mins to remove cellular debris and the supernatant stored at 4°C for up to 3 months. The number of lambda bacteriophage in the lysate was checked to be at least 10^{10} pfu/ml (section **2.4.13.2**) before proceeding with DNA purification.

2.4.24.2 Removal of the Lambda Phage Coat

40 µl of nuclease mixture (supplied in the kit) was added to 10 ml lysate and incubated at 37°C for 15 mins. 4 ml of provided phage precipitant was added and gently mixed and placed on ice for 15 mins. This mixture was then resuspended in 500 µl of phage buffer and centrifuged at 12,000 g for 10 secs. The supernatant was then mixed thoroughly with 1 ml of purification resin.

2.4.24.3 Lambda DNA purification

The lysate/resin mix was passed through a Wizard™ minicolumn. The minicolumn was then washed with 2 ml of 80% isopropanol. The resin was dried by centrigugation of the minicolumn in a microcentrifuge tube for 30 secs at 12,000 g. 100 µl of water, preheated to 80°C, was applied to the minicolumn and centrifugation at 12,000 g for 20 secs was carried out to elute the DNA. The purified lambda DNA was stored at 4°C until use.

2.4.25 Polymerase chain reaction

DNA was amplified from a number of sources, but all following the same general procedure, detailed below:

Reaction cocktail:

Unamplified DNA (~100 ng genomic DNA, ~5 ng plasmid)

5 µl 10X promega PCR buffer

5 µl 2.5 mM $MgCl_2$

5 µl DMSO (10% v/v)

8 µl dNTPs (1.25 mM)

Primers (100 pmol each)

water to 50 µl

Protocol:

1.Denature DNA (95°C for 4 mins)

2.Add 2.5 U Promega Taq polymerase

3.Anneal primers (45-65^{o}C, depending on the annealing temperature of the primers used; typically for 30 secs).

4.Extension of the DNA (72^{o}C for 30 sec-2 mins, depending on the length of DNA being amplified).

5.Repeat steps 3 and 4 20-30 times.

6.After last cycle, denature DNA (95^{o}C for 1 min)

7.Anneal primers for 2 mins and carry out extension for 4 times the usual time.

Typically, a tenth of the PCR reaction was run on a 1% agarose gel to determine if amplification of the DNA had occured. DNA could be extracted and purified as described in section **2.4.13.**

2.4.26 Dynabead™ PCR Sequencing

This method was used to prepare ssDNA for sequencing, from a double-stranded template.

2.4.26.1 Preparation of the Dynabeads™

The beads were pulled down from a 30 µl suspension using the Magnetic Particle Concentrator(MCP) and the supernatant removed. The beads were washed 3 times with 0.5 ml 0.1 M NaCl/TE pH 8.0 and finally resuspended in 30 µl 0.1 M NaCl/TE pH 8.0.

2.4.26.2 Preparation of the PCR product

The PCR reactions were set up using a single biotinylated primer (forward or reverse) in each. The PCR product was cleaned by gel purification using Qiaex (section **2.4.12**). The PCR product was spun through Sephadex G50 for 12 mins at 15,000 g to remove the primers. The DNA was ethanol precipitated (section **2.4.5**) and resuspended in 50 µl 1M NaCl. This was then added to the prepared Dynabeads™ and incubated at room temperature with agitation for 30 mins. The beads were separated using the magnet and the supernatant removed. The DNA was denatured by incubation of the beads for 5 mins in 100 µl 0.1 M NaOH and repeated once. The beads were then washed 3 times in water and finally resuspended in 14 µl of water.

2.4.26.3 Sequencing Reactions

7 µl of the beads plus DNA suspension was added to 2 µl of 5x sequencing mixture and 1 µl sequencing primer. The miture was heated to 85°C in a water bath for 5 mins and then slowly cooled to allow annealing of the primers. 1 µl DTT, 2 µl Gmix, 0.5 µl ^{35}SdATP and 2 µl sequenase (diluted 1:5) were added to the mixture and incubated at room temperature for 5 mins. 3.5 µl of this mixture was added to each termination mix and incubated at 37°C for 10 mins. The beads were removed and the supernatant added to 4 µl stop solution and incubated at 37°C for 10 mins before adding 3 µl H_2O. 2-3 µl was loaded on the sequencing gel (section **2.4.23**).

CHAPTER 3

Results and Discussion

3.1 Introduction

Chapman (1994) cloned and sequenced the gene encoding isocitrate lyase from *Streptomyces coelicolor.* The 5.4 kb fragment that was cloned, not only contained the entire gene encoding ICL, but also 1152 bp of the gene encoding malate synthase downstream of *icl* and 2.8 kb upstream of the start codon of *icl* (Fig.3.1). The clones and subclones created by Chapman (1994) were used in this study.

In *E.coli* the genes encoding ICL and MS, *aceA* and *aceB* respectively, are found next to each other on the genomic map, with *aceB* upstream of *aceA*. They are transcribed as an operon which also includes *aceK*, the gene encoding isocitrate dehydrogenase kinase phosphatase, which is situated downstream from *aceA*.

In *S.coelicolor* the gene order of *icl* and *ms* is different from that in *E.coli* (Chapman, 1994). It would now be interesting to see if *icl* and *ms* form an operon as they do in *E.coli* and also to investigate whether a gene encoding ICDHK/P forms part of this operon, although Taylor (1992) suggested that the *S.coelicolor* ICDH is not phosphorylated in the manner found in *E.coli.* So, an ICDHK/P enzyme may not be present.

It was decided that in order to investigate the glyoxylate bypass operon in *S.coelicolor* further, it would be most useful to clone the entire *ms* gene within a fragment which would also include the region downstream of *ms* in order that this may also be investigated.

It was also of interest to sequence the region upstream of *icl*. There may be regulatory genes such as *iclR* or a gene encoding ICDHK/P(as in *E.coli*) which could

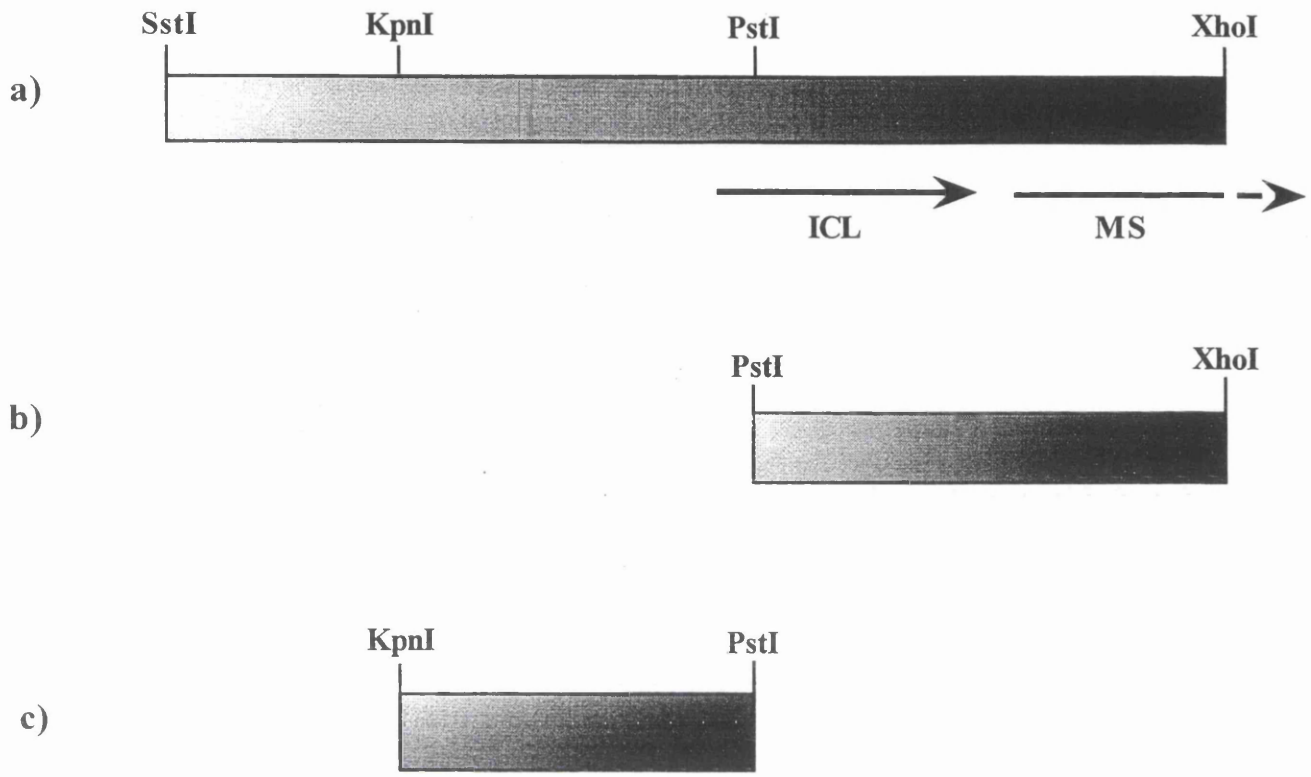


Figure 3.1 Plasmid clones and subclones used in this study

Figure 3.1 Plasmid clones and subclones used in this study

a) ICL92 (5.4kb). The arrows represent the ICL and MS ORFs.
b) ICL92B (2.4kb)
c) ICL92C (1.8kb)

All the clones are contained within pBluescript II (ks+) and were created by Chapman (1994). ICL92B and ICL92C were cloned from ICL92. pICL92 was obtained by screening a *S.coelicolor* plasmid sublibrary created in pBluescript II (ks+) by colony hybridisation. Only the relevant restriction sites are shown.

be located close to the *icl* and *ms* on the genome and be involved in control of expression of these genes. Genes encoding important metabolic proteins could also be located upstream of *icl*.

This chapter describes the attempts to clone the entire gene encoding malate synthase and the region of DNA downstream and the sequencing of the region of DNA upstream from the gene encoding isocitrate lyase in *S.coelicolor*.

3.2 Cloning of the gene encoding malate synthase

ICL92B (Chapman 1994) (Fig 3.1) was used as a probe, since this is homologous to 5' end of *ms*. This probe was used to screen *S.coelicolor* genomic DNA and a genomic library of *S.coelicolor* DNA prepared in bacteriophage λ.

3.2.1 Screening of genomic DNA

S.coelicolor genomic DNA was digested to completion using *Pst*I and a selection of other restriction enzymes. *Pst*I was used in all the digests because this restriction enzyme cuts at the 5' end of the probe, ICL92B, and therefore a positive signal will indicate that the DNA is downstream of this *Pst*I site and contains part of *ms*.

The digested DNA was separated on a 0.8% (w/v) agarose gel and visualised by ethidium bromide staining (Fig 3.2a). The DNA was transferred to Hybond-N nylon membrane and the DNA fixed to the membrane by UV cross linking (section **2.4.16**). The filter was then subjected to prehybridisation after which, the DIG labelled ICL92B probe was added to the hybridisation solution (section **2.4.20**).

Figure 3.2 Restriction digests and Southern hybridisation of *S.coelicolor* 1174

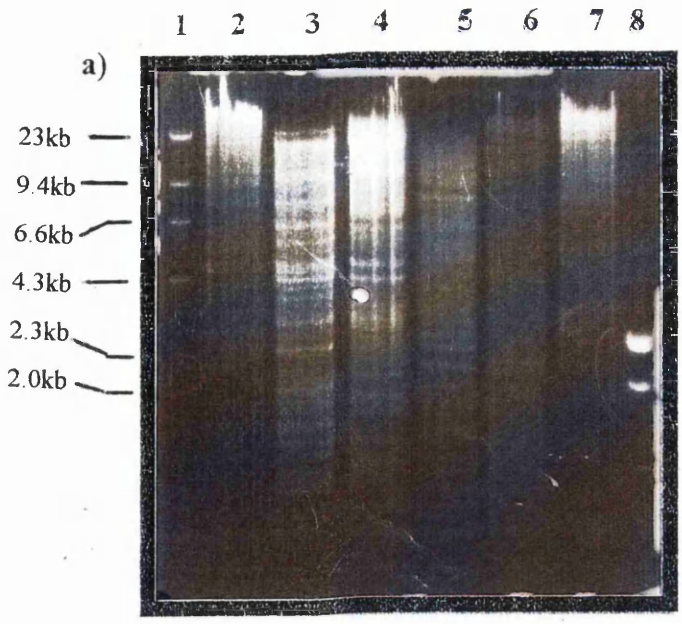
a) Genomic DNA was digested for 2 hours at 37°C using 10 units of each of the restriction enzymes shown below. Fragments were separated by gel electrophoresis through a 0.8% (w/v) agarose gel overnight and photographed under UV illumination.

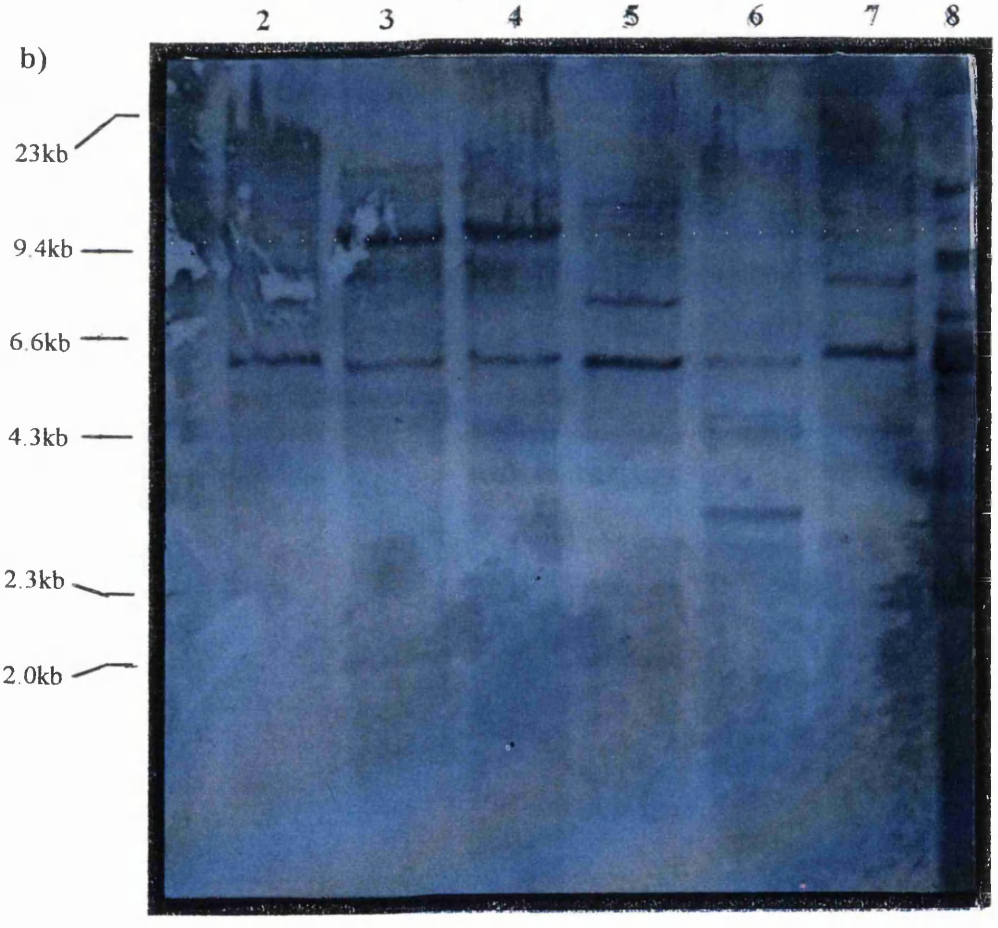
b) The DNA was transferred to Hybond-N for hybridisation with the DIG-labelled ICL92B probe. Hybridisation was carried out at 65°C in 0.1xSSC, 0.1%SDS. Chemiluminescent detection with DIG was carried out and autoradiography was at room temperature for 5 hours.

Several bands can be seen in each lane suggesting that conditions were not stringent enough to allow specific hybridisation to the *ms* gene alone, or that there was incomplete digestion of the genomic DNA. The appearance of a 5.5 kb *Pst*I hybridising band for all the digests suggests that this is specific.

The sizes of the bands are shown below. The size of each DNA fragment was calculated by comparison to the mobility of the λ*Hind*III DNA markers using a calibration graph.

1. λ*Hind*III markers	
2. *Pst*I	5.5 kb
3. *Pst*I/*BamH*I	5.5 kb, 10 kb
4. *Pst*I/*Sst*I	5.5 kb, 10 kb
5. *Pst*I/*Kpn*I	5.5 kb, 7 kb
6. *Pst*I/*Sma*I	5.5 kb, 3.2 kb
7. *Pst*I/*Xba*I	5.5 kb, 8 kb
8. *Pst*I/*Xho*I digest pICL92B	


1 2 3 4 5 6 7 8
a)
23kb
9.4kb
6.6kb
4.3kb
2.3kb
2.0kb


2 3 4 5 6 7 8
b)
23kb
9.4kb
6.6kb
4.3kb
2.3kb
2.0kb

Hybridisation was carried out at 65°C in 1x SSC, 0.1% SDS for 16 hours. The filter was then washed four times (section **2.4.17.2**) in 0.1x SSC, 0.1% SDS, twice at room temperature and twice at 65°C.

Chemiluminescent DIG detection was then carried out according to the manufacturer's instructions. The filter was subjected to autoradiography for 5 hrs and developed (Fig 3.2b).

Malate synthase from other sources has been found to have a molecular weight of around 62 kDa. Assuming an average MW for an amino acid of 100 Da, the maximum size of the *ms* gene is 1.6-1.9 kbp. Chapman (1994) cloned 1152 bp of the *ms* gene, therefore approximately 0.45-0.75 kb remains to be cloned. To ensure that the whole gene is cloned, the fragment must be at least 0.75 kb larger than the 2.4 kb probe, i.e. at least 3.15 kb, and preferably larger in order to investigate the region downstream of *ms*.

Hybridisation studies revealed a 5.5 kb *Pst*I hybridising band for all the digests (Fig 3.2b). Other larger bands were observed, probably due to incomplete digestion of the genomic DNA. These were *PstI/BamHI* (10 kb), *PstI/Sst*I (10 kb), *PstI/Kpn*I (7 kb), *PstI/Xba*I (8 kb). A 3.2 kb *PstI/Sma*I band was also revealed, but this fragment was only 0.8 kb larger than the probe and therefore it was thought that this may not contain the entire gene encoding MS and would certainly contain very little of the region downstream.

It was predicted that the 5.5 kb *Pst*I fragment was large enough to potentially contain the whole of the gene encoding MS and the region downstream and yet is a

convenient size for introduction into a plasmid vector. This fragment was therefore chosen for further investigation.

3.2.2 Purification of size selected hybridising DNA

In a scaled up reaction, 20 units of *Pst*I was used to digest 10 μg of *S.coelicolor* genomic DNA to completion. Fragments were separated by agarose gel electrophoresis using a 0.8% (w/v) TAE gel. Slices of gel that contained DNA fragments in the size range of 4.8 kb to 6.5 kb were then excised and the DNA extracted by centrifugation through Spin-X tubes (section **2.4.12**).

3.2.3 Construction of a plasmid sub-library

A ligation was set up containing approximately 100 ng of the size-selected DNA and 50 ng of linearised pUC18 vector (section **2.4.8**). This vector was supplied by Pharmacia, already linearised by *Pst*I and treated with calf intestinal alkaline phosphatase. This treatment removes the phosphate group from the 5' ends of the linearised molecule, preventing both the intramolecular and intermolecular ligation of vector molecules. The efficiency of removal of the phosphate from the 5' ends of the vector meant that 99% of the colonies obtained were expected to be recombinants.

Varying proportions of the ligation mix were used to transform a culture of *E.coli* TG1 that had been made competent by the method of Hanahan (section **2.3.8.1**). Colonies that contained recombinant plasmids, detected by the white colour of the colonies plated on media containing X-gal (section **2.3.8.4**), were then picked for subsequent screening by hybridisation to radiolabelled ICL92B.

3.2.4 Screening of a partial plasmid library

Two hundred of the transformants were picked and plated onto square L-plates containing ampicillin, as a 10 x10 array on 2 plates (100 colonies per plate). Duplicate arrays were also made on Hybond-N and the colonies grown overnight before being treated prior to hybridisation (section **2.4.16**). Prehybridisation was carried out in Quickhyb at 65^{o}C. After 30 mins the radiolabelled probe was added and hybridisation was carried out for 2 hrs at 65^{o}C (section **2.4.17**). Washing was carried out 4 times in 0.1x SSC, 0.1% SDS, twice at room temperature and twice at 65^{o}C. Autoradiography at -70^{o}C overnight revealed no colonies that hybridised to ICL92B.

This entire procedure was carried out a number of times. In total 600 transformants were screened, but none were found to hybridise to the probe, ICL92B, so this method was eventually abandoned.

The transformation frequencies for the control DNA were much greater than those for the recombinant DNA which suggested that there was a problem with the cutting and ligation procedure and this may have been the reason for the lack of success with this method.

3.2.5 Screening of an amplified lambda library

Concurrently with the screening of genomic DNA and construction of a partial plasmid library, a *S.coelicolor* lambda library was also screened as a contingency that the entire gene may not be cloned from the partial plasmid library.

A lambda library of *S.coelicolor* genomic DNA had been constructed previously (Taylor, 1992). This amplified library was screened using the homologous ICL92B product (Fig.3.1) as a probe, in an attempt to clone a DNA fragment containing the entire gene encoding MS.

It was calculated, assuming a genome size of 8 Mb (Hopwood and Keiser, 1990) and an average insert size of 11 kb, that 3347 plaques had to be screened to have a 99% probability that the required DNA sequence would be present, using the equation below.

$$\text{n.o. clones required} = \frac{\ln(1-p)}{\ln[1-\text{insert size/genome size}]}$$

where p = the probability of a particular DNA sequence being present.

Approximately 8000 plaques were grown on a lawn of *E.coli* LE392 and the DNA transfered to Hybond-N nylon membrane (section **2.4.16**). This is more than twice the number required to have a 99% chance of picking a particular sequence.

Plaque lifts were treated as described in section **2.4.19**. Hybridisation of duplicate filters in Quikhyb at 680C was carried out for 16 hrs with radiolabelled ICL92B, prepared by random priming (section **2.4.15**). The filters were washed twice at room temperature in 2x SSC, 0.1% SDS and twice at 68^{o}C in 0.1x SSC, 0.1% SDS (section **2.4.17.2.**) and after autoradiography the films were developed and at least 14 positive signals were obtained (Fig 3.3).

The plaques were too densely packed and the signals obtained too large to be assigned to individual plaques. Therefore 5 of the positive signals were investigated

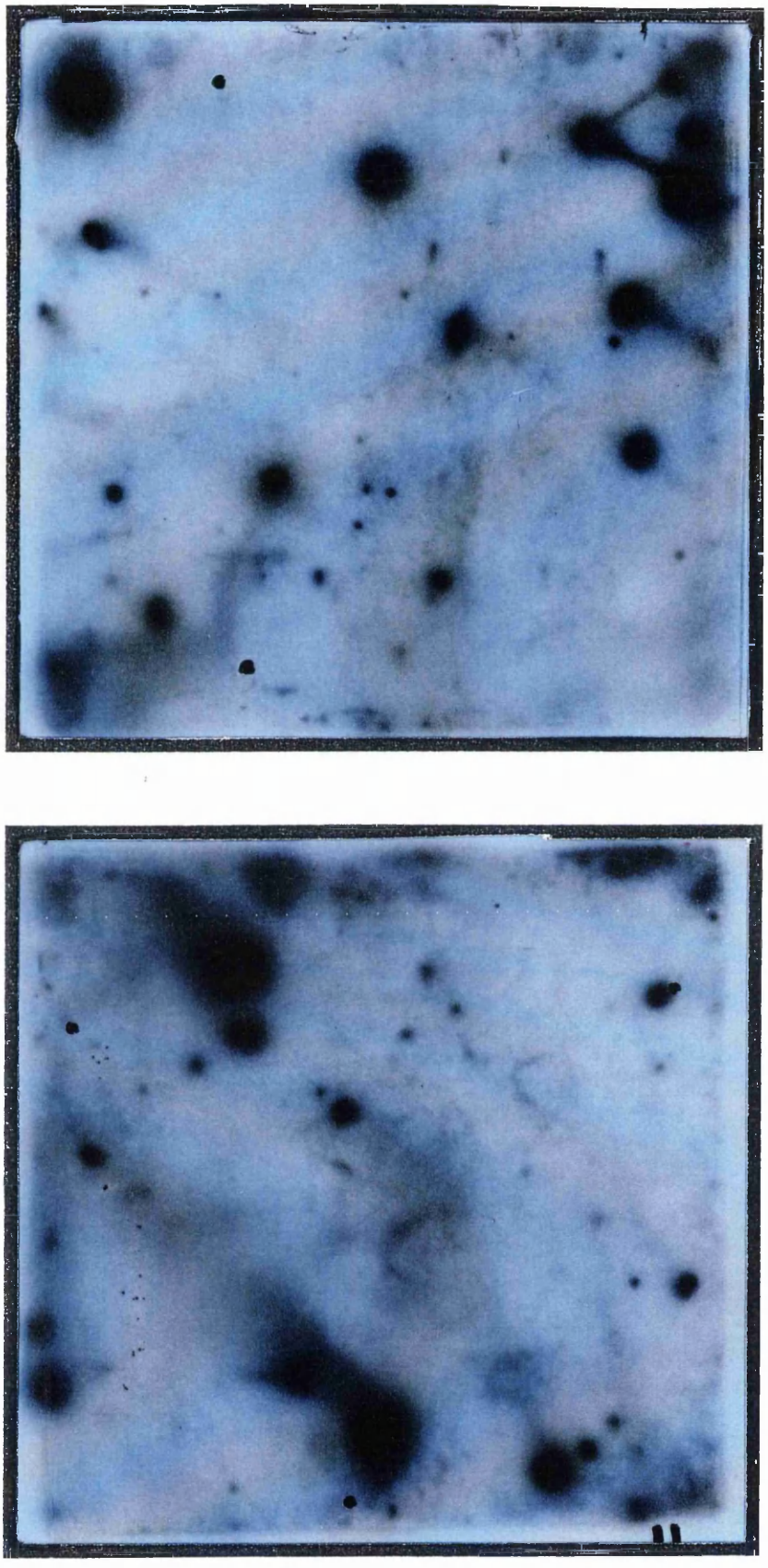


Figure 3.3 Plaque hybridisation with ICL92B

2 plates (with approximately 4000 plaques) were screened by hybridisation to ICL92B. Duplicate filters were prepared by taking plaque lifts. All 4 filters were subjected to hybridisation at 68°C in Quikhyb and washed twice at room temp in 2 x SSC and twice at 68°C in 0.1 x SSC. Only 1 of each of the duplicate filters are shown.

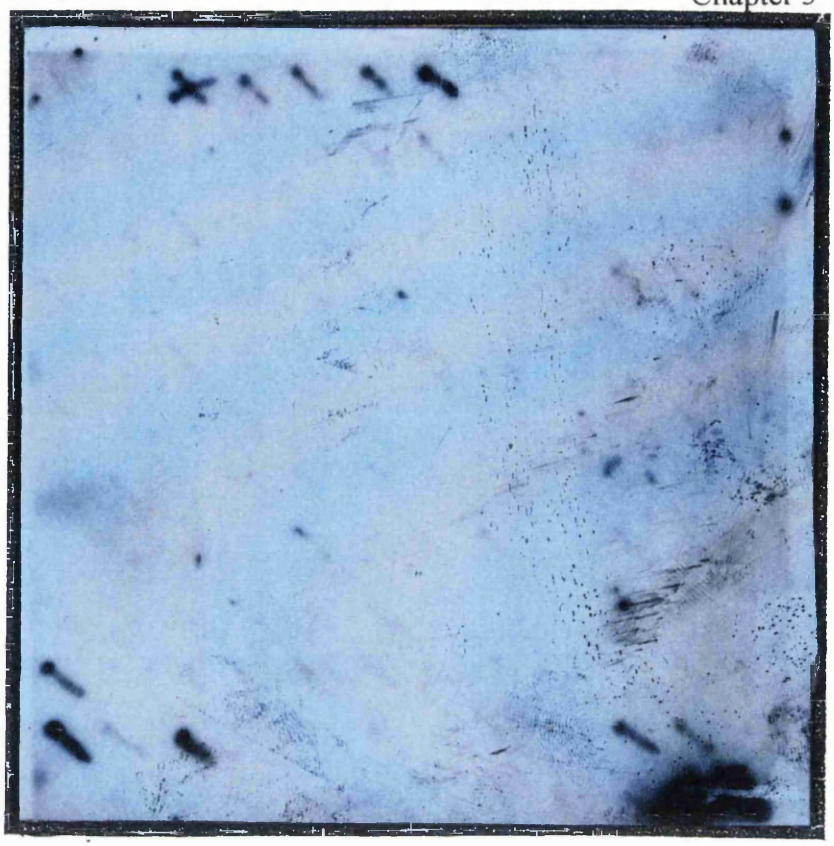


Figure 3.4 Phage hybridisation with ICL92B

Plaques from the regions of positive signals (Fig3.3) were picked and the phage streaked onto a lawn of *E.coli* LE392 in order to obtain distinct plaques. Phage from these plaques were then streaked onto Hybond-N in duplicate and probed with radiolabelled ICL92B. Only one of the duplicate filters is shown.

by picking between 3 and 6 plaques from the area of each signal (section **2.4.13.2**). Phage from each plaque were streaked on a lawn of *E.coli* LE392 to produce distinct single plaques. 5 single plaques were isolated from each plate and the phage streaked onto Hybond-N in duplicate. The duplicate membranes were probed using radiolabelled ICL92B using the conditions described above. 3 of the original 5 plaques gave positive signals (Fig 3.4).

The 3 recombinant phage which gave the strongest signals were chosen for further analysis (λ1-3). They were from 3 different plaques which gave positive signals in the original screening.

DNA was prepared from each of the λ clones (section **2.4.24**) and restriction digests showed that λ1 and λ2 contained the same insert (Fig3.5) whilst λ3 contained a different insert. Further analysis was then carried out to investigate whether the 2 clones (λ2 and λ3 only) contained the required fragment of DNA.

Earlier studies show that a 5.5 kb *Pst*I fragment should contain the required DNA (section **3.2.1**). The polylinker of the lambda vector contains an *Sst*I site on each side of the insert, so this enzyme is used to remove the insert from the vector. It is known that no *Sst*I site exists within the probe, ICL92B, or for at least 4 kb downstream (Chapman, 1994). Therefore if the λ clones are cut with *Pst*I and *Sst*I, only if the probe hybridises to a band greater than 3.15 kb will it contain the entire *ms* gene.

DNA from the 2 λ clones was digested with the restriction enzymes *Pst*I, *Sst*I and *Pst*I and *Sst*I (Fig3.6a) and probed with radiolabelled ICL92B. This should give enough information to indicate whether or not a large enough region of DNA had

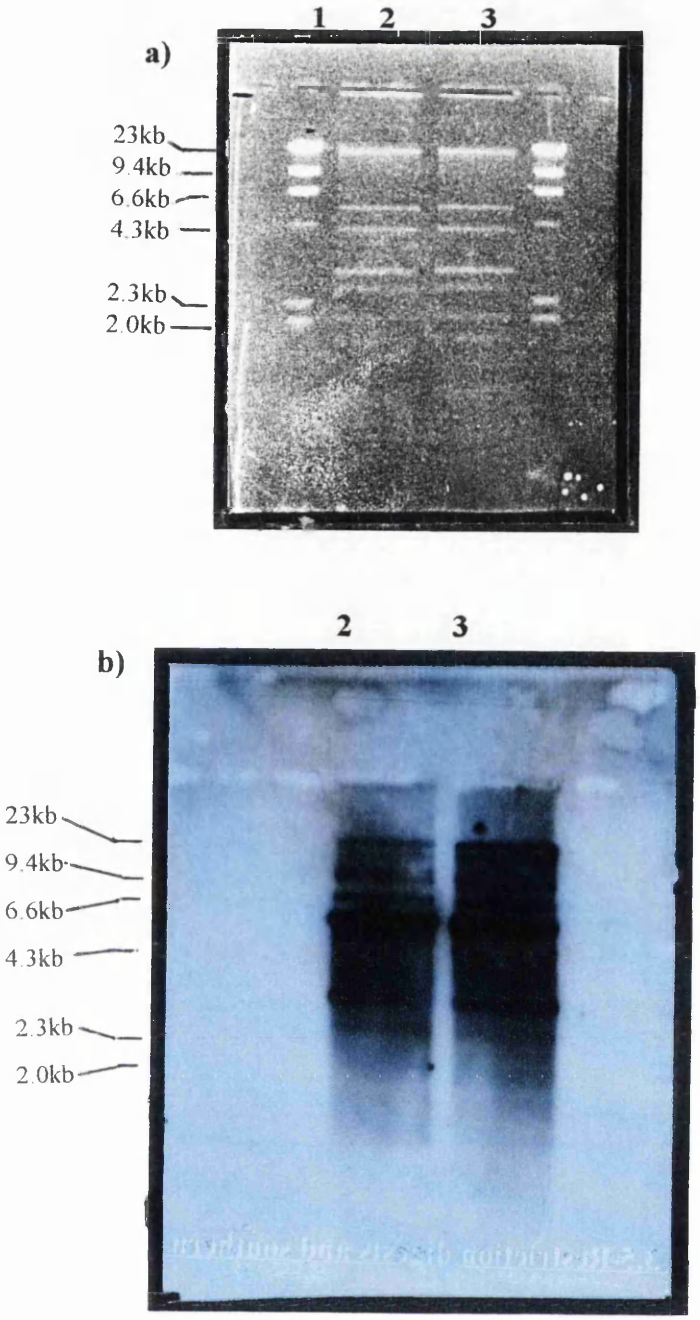


Figure 3.5 Restriction digests and southern hybridisation of recombinant phage λ1 and λ2

a) DNA prepared from recombinant phage λ1 and λ2 were digested for 2 hours at 37^{0}C using 2 units of restriction enzyme *PstI*. Fragments were separated by gel electrophoresis through a 0.8%(w/v) agarose gel for 2 hours, stained with ethidium bromide and photgraphed under UV illumination (2.4.11).

b) The DNA was transferred to Hybond-N (2.4.16) for hybridisation with radiolabelled ICL92B used as a probe. Hybridisation was carried out at 65^{O}C in Quikhyb and the filter washed twice at room temperature and twice at 65^{O}C in 0.1 x SSC, 0.1% SDS.

It can be seen that both recombinant phage have the same digest pattern when digested with *PstI* and the same bands hybridise to ICL92B.

1) λ*HindIII* markers
2) λ1
3) λ2

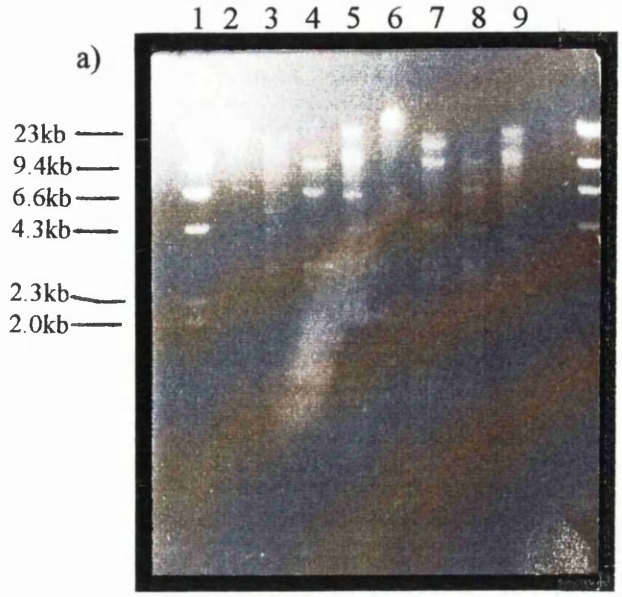
Figure 3.6 Restriction digests and Southern hybridisation of recombinant phage λ2 and λ3 DNA

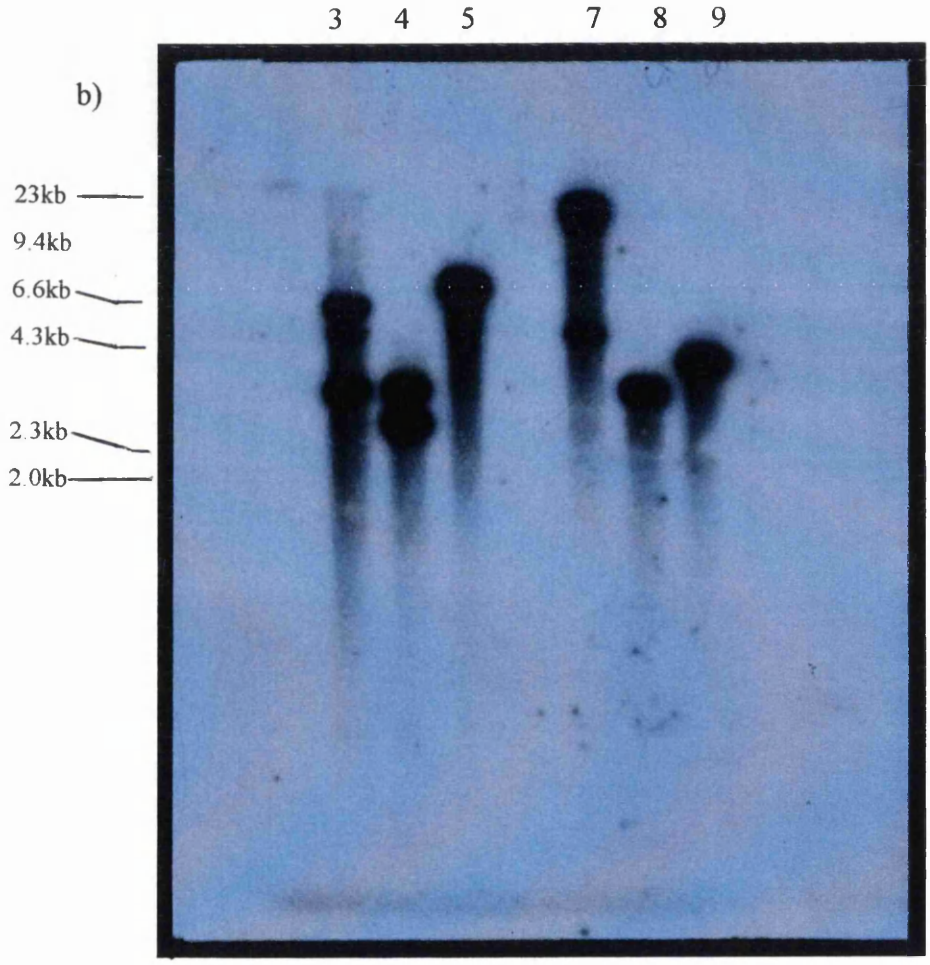
a) DNA prepared from recombinant phage λ2 an λ3 was digested with *PstI, SstI* and *PstI* and *SstI*. Fragments were separated by gel electrophoresis, stained with EtBr and photographed under UV illumination.

b) The DNA was transferred to Hybond-N for hybridisation with radiolabelled pICL92B. Hybridisation was carried out at 65^{0}C in Quikhybe and the filter washed as described in section 3.2.1.

The size of each hybridising band is shown below. The size was calculated by comparison with the mobilities of the λ*HindIII* DNA markers.

1) λ*HindIII* markers	
2) λ2	
3) λ2 *PstI*	5.8 kb, 3kb
4) λ2 *PstI/SstI*	3 kb, 2.8 kb
5) λ2 *SstI*	10 kb
6) λ3	
7) λ3 *PstI*	20 kb, 4.8 kb
8) λ3 *PstI/SstI*	3 kb
9) λ3 *SstI*	3.5 kb


1 2 3 4 5 6 7 8 9
a)
23kb
9.4kb
6.6kb
4.3kb
2.3kb
2.0kb


3 4 5 7 8 9
b)
23kb
9.4kb
6.6kb
4.3kb
2.3kb
2.0kb

been cloned.

λ2 gave a 2.4 kb band when cut with *PstI* and *SstI* which hybridised to ICL92B (Fig 3.6b) suggesting that this clone did not contain any of the region downstream of the probe. λ3 when cut with *PstI* and *SstI* produced a band of 3 kb (Fig 3.6b) which could contain the entire gene encoding MS. However it would not include any of the region downstream of *ms* and discrepancies with previous results suggested that the entire gene may not even be included within this clone.

The presence of the 3 kb fragments in lanes 2 and 3 could not be easily explained. This and previous anomolous results (not shown) suggested that the probe may not be clean or may contain regions upstream of the *PstI* site which would also cast doubt on previous experiments. Due to the inconsistency of the results and the doubts about this clone it was decided not to pursue with either of these recombinant λ phage.

3.3 Sequencing upstream of *icl*

pICL92 (Fig3.1) contains *icl*, part of *ms* and 2.8 kb upstream of the start codon of *icl*. This section describes how part of this region was sequenced in order to investigate the presence of additional ORFs in the region.

3.3.1 Construction of sub-clones for DNA sequencing

To simplify the task of sequencing, sub-clones were constructed in M13, from which single stranded DNA for sequencing is easily prepared.

The plasmid pICL92C (Chapman, 1994) (Fig3.1) was digested with a number of restriction enzymes in order to obtain a restriction map of the clone. *Sal*I was found to be the most useful restriction enzyme since there are 3 sites at which it cuts within the fragment. The plasmid was digested with *Sal*I and *Kpn*I which yielded 4 bands. A *Kpn*I/*Sal*I fragment of 0.2 kb, 2 *Sal*I fragments of 0.9 kb and 0.45 kb and the linearised plasmid including 0.35 kb of the insert which had already been sequenced (Chapman, 1994).

The DNA was separated on a 0.8% (w/v) agarose gel, the 3 smaller bands excised and the DNA purified using Spin-X tubes (section **2.4.12**) The 0.9 kb and 0.45 kb *Sal*I fragments, S1 and S2, were initially cloned into plasmid pUC18 in order to determine their orientation before subcloning into M13 in a directional manner. Ligations were set up containing S1 or S2 and linearised pUC18 vector. This vector was supplied by BRL already linearised by *Sal*I and treated with calf intestinal alkaline phosphatase. Varying proportions of the ligation mix were used to transform a culture of *E.coli* TG1 that had been made competent by the method of

Hanahan (section **2.3.8.1**). Colonies that contained recombinant plasmids were picked and the plasmid DNA purified (section **2.4.2**).

The S1 and S2 inserts were cut out of pUC18 by digesting with *EcoR*I and *Hind*III and the 0.9 kb and 0.45 kb bands purified using SpinX tubes (section **2.4.12**). The M13 vectors, mp18 and mp19 were digested with *Sal*I/*Kpn*I and *Hind*III/*EcoR*I and ligations were set up with the 0.2 kb *Sal*I/*Kpn*I fragment, S3, and S2 and S1 respectively. The use of different ends in each ligation resulted in the cloning of the inserts in both orientations with respect to the -40 primer.

No recombinants were obtained for the S1 fragment and despite a number of attempts to clone it into M13 using varying conditions, this was not achieved and an alternative method of preparing single stranded (ss)DNA for sequencing was employed (section **3.3.2**)

The sub-cloning yielded a number of M13 clones containing fragments S2 and S3 which were used for sequencing (Fig 3.7).

3.3.2 Preparation of ssDNA using PCR

PCR reactions (section **2.4.15**) were set up using the pUC18S1 template. 2 reactions were set up using a single biotinylated primer (forward or reverse) in each (section **2.4.26.2**). The ssDNA was then separated by binding the biotinylated primer to Dyna beads™ and denaturing the dsDNA (section **2.4.26.2**), the ssDNA was then collected and used in the sequencing reactions.

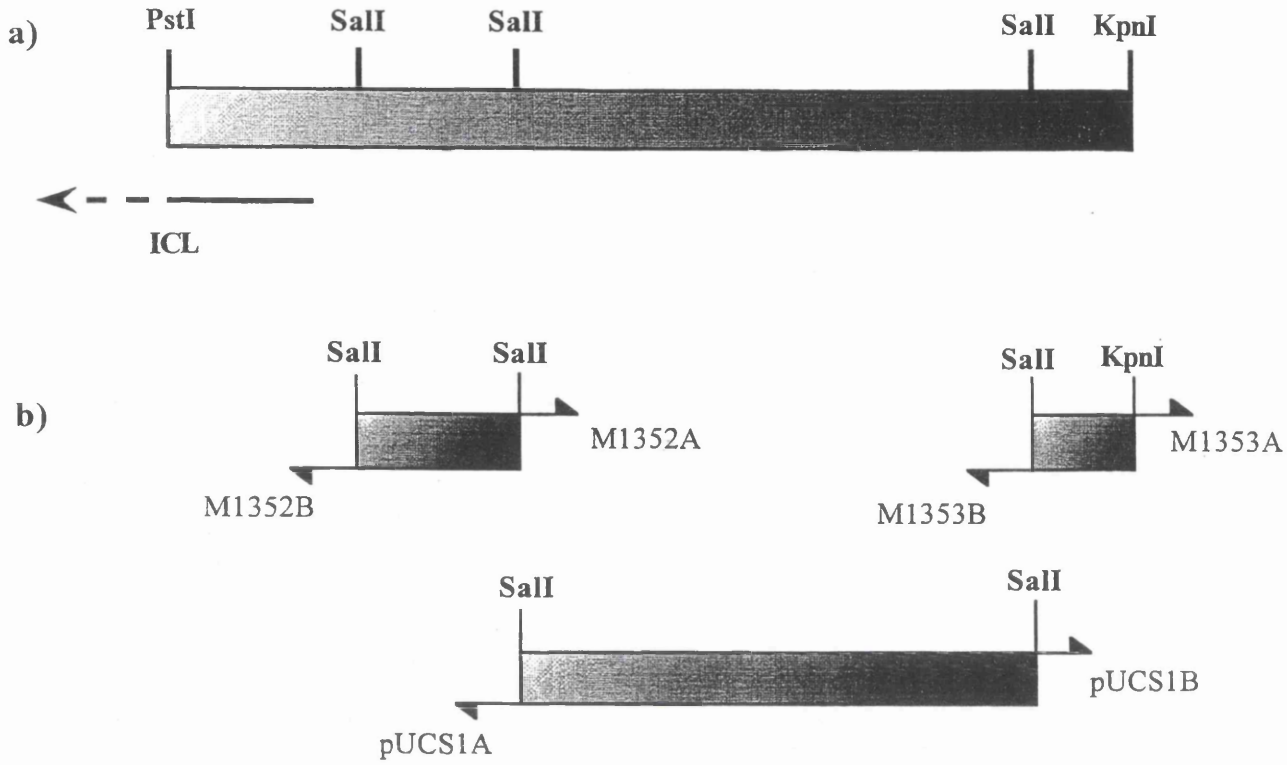


Figure 3.7 Summary of the subclones used for sequencing ICL92C

a) This shows the ICL92C subclone (Chapman, 1994), only the relevant restriction sites are shown.

b) Fragments of DNA were cloned into pUC18 or M13mp18 or 19 for sequencing. The labelled arrows indicate the name of the subclone and the direction in which the DNA was sequenced.

3.3.3 Sequencing Strategy

The most reliable, easiest and most used method of sequencing DNA is the di-deoxy chain termination method of Sanger *et al* (1977). Many commercial kits are available that contain all the reagents required for performing the reactions, making this method both reliable and convenient. The reactions involve the strand complementary to the one being sequenced is synthesised in vitro by extension from a DNA primer using a DNA polymerase.

In this study a kit supplied by USB Biochemicals was used which makes use of the "sequenase", a T7 polymerase which has been modified to remove its 5' to 3' exonuclease activity.

The protocols were followed as per the manufacturers instructions except for the termination reactions which were carried out at 42^{o}C rather than the recommended 37^{o}C. Streptomycete DNA has a high G+C content and this leads to the formation of secondary structures during sequencing reactions. These structure cause the polymerase enzyme to stop and results in a four track stop when the reactions are run on a gel. A higher incubation temperature for the sequencing reactions results in the melting of the intramolecular secondary structures and reduces the risk of problems.

The high G+C content of Streptomycete DNA can also result in a higher frequency of compressions. Four track stops are caused by the formation of secondary structures during the sequencing reactions, whilst compressions are a result of secondary structures during electrophoresis and therefore the DNA runs aberrantly. Two bands may appear to have the same mobility and therefore this makes it difficult to read the gel accurately.

Deaza nucleotide anologues are available to overcome this problem and were routinely used in sequencing reactions. They help to prevent G+C base pairing which is the cause of much of the secondary structure formation and therefore prevent compressions (Mizosawa *et al.;* 1986).

Single stranded DNA template generated from M13 (section **2.4.14**) or PCR (section **3.3.2**) was used for sequencing which routinely gave 250-300 nucleotides of sequence. SsDNA templates have been found to be more reliable than dsDNA templates at producing clear, distinct bands.

SsDNA was sequenced using the -40 primer which anneals just outside of the multiple cloning site. S3 was sequenced on both strands, whilst S2 was sequenced only on one strand, S1 was partially sequenced on one strand (Fig 3.8). Sequencing of both strands was unnecesary for the preliminary investigation. If sequence of interest had been recognised, the entire region would have been sequenced on both strands and across the restriction sites used for cloning. This would also have been repeated 2-3 times in order to eliminate mistakes. The sequence obtained is shown in Figure 3.9.

3.3.4 Identification of Open Reading Frames

The high G+C content of *Streptomyces* DNA imposes an extreme bias on the distribution of nucleotides in each of the 3 positions of a codon within a coding sequence.

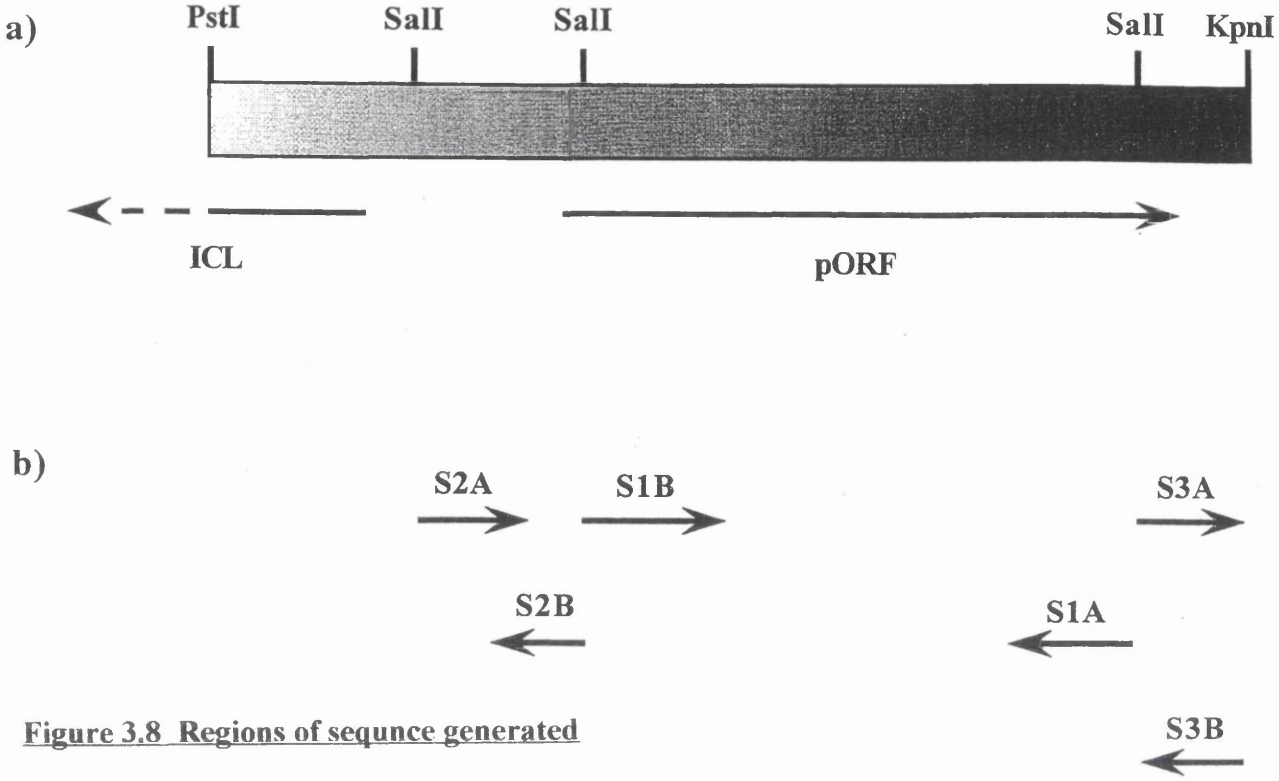


Figure 3.8 Regions of sequnce generated

a) This shows the ICL92C subclone (Chapman, 1994). The 2 arrows represent the gene encoding ICL and the potential ORF.

b) Each arrow represents the direction and relative amount of DNA sequenced from each subclone.

Figure 3.9 DNA sequence and potential peptide sequence of sections of pICL92C

The regions of DNA sequence (A) obtained from sequencing the subclones of pICL92C (Fig 3.8). Potential amino acid sequences (B) were obtained by Bibb analysis.

a) S2
b) S1B
c) S1A
d) S3

a) S2

```
A)1     atcttgttgccgtactttgccaactttgcgaagccctcg

A)40    ggagccggtgcgtcccctacgctgaccgggtcgggttac

A)79    cggtagggaggagcggtcggtgagcaagacgtacgcggg

A)118   tgcgcggctgcggcgctgcgcgaggagtcggctgagcca

A)157   ggccgcactcgcccgcgtgctcggtatctccccgagcta

A)196   cctgaaccagatggagcacgactcccgcccgctcaccgt

A)235   gcccgttctcctccggctgacggagacccttggtgtcga

A)274   c
```

b) S1B

```
A)1     gtcgacgccgccttcttctccgagcgggacaccgcccgc
B)      V  D  A  A  F  F  S  E  R  D  T  A  R

A)40    ctcttggccgacctgcgcgaggcactgaccggtgagctg
B)      L  L  A  D  L  R  E  A  L  T  G  E  L

A)79    gcggtcgcccgggtctccgcgtccgacctcgccgaactg
B)      A  V  A  R  V  S  A  S  D  L  A  E  L

A)118   gcgtcacggatgccggcggtcgcccaggtgctgatcgac
B)      A  S  R  M  P  A  V  A  Q  V  L  I  D

A)157   ctgggccgccgcaaccagctgctgtccgagcgactggcc
B)      L  G  R  R  N  Q  L  L  S  E  R  L  A

A)196   gggcgacggacggaccgatgggaccga
B)      G  R  R  T  D  R  W  D  R
```

c) S1A

```
A)1     gtgcaggacttcgacccggttcccgcccacgcgttggcc
B)      V  Q  D  F  D  P  V  P  A  H  A  L  A

A)40    cgcatcggcatcgccaactacttcgcggccgtgttgatc
B)      R  I  G  I  A  N  Y  F  A  A  V  L  I

A)79    ctgccgtacacggccttccacgcggcggtcgaggagttc
B)      L  P  Y  T  A  F  H  A  A  V  E  E  F

A)118   cgctacgacatcgagcgcctcaccgaccattacggcctg
B)      R  Y  D  I  E  R  L  T  D  H  Y  G  L

A)157   ggctacgagactgtgggccaccgtctgagcacccttca
B)      G  Y  E  T  V  G  H  R  L  S  T  L
```

d) S3

```
A)1     gtcgaccgggccggcaacatgtccaaacgccagtccgcg
B)        R  P  G  R  Q  H  V  Q  T  P  V  R  D

A)40    accgggttccacttcagccgggccggcggcacctgcccg
B)        R  V  P  L  Q  P  G  R  R  H  L  P  A

A)79    ctctggaacgtatacgagtccttcgccactcccggccgc
B)        S  G

A)118   atccacgtccagctgtccgagatgcccgacggacagcgg

A)157   tacc
```

Analysis of the third postion G+C% bias is a particularly useful way to identify potential coding regions within *Streptomyces* genes. Approximately 90% of all codons within an open reading frame contain a G or C in this position. The first and second postions are also distinctive. 70% of codons contain a G or C in the first postion and 50% in the second (Bibb *et al.*; 1984). In non-coding regions of DNA the distribution is more random, thus it is relatively easy to distinguish an open reading frame in a region of DNA by looking for this non-random distribution of nucleotides.

The computer program 'CODONPREFERENCE' (Devereaux *et al.*; 1984) from the 'GCG' package is used to perform this kind of analysis. The program is used to help identify protein coding regions, the frame which encodes the proteins and the direction of translation. It is also useful in predicting the location of DNA sequencing errors.

The program finds regions of each reading frame, in a DNA sequence that show either a strong codon preference (Gribskov *et al.*; 1984) or a strong G+C% bias in the third position of each codon (Bibb *et al.*; 1984).

The codonpreference is calculated from a codon usage table that was compiled from 67 *Streptomyces genes.* This is used to calculate a statistic which shows the similarity to the usage table. The 'CODONPREFERENCE' program also draws a plot that provides a measure of how closely the codon usage of the predicted frame correlates with the codon usage table. Non-coding regions are more random and therefore have lower values for both the third position G+C bias and the codon preference.

The program calculates the third position G+C content and the codon preference in a given window starting at each of the three postions. The window is moved

sequentially along the DNA sequence and the results plotted as three separate graphs representing each frame.

Analysis of the 'CODONPREFERENCE' plot (Fig 3.10) for each of the sections of sequence revealed extreme G+C bias and a codon usage similar to that expected for *Streptomyces* genes in some areas of the sequence. These areas also contain far fewer rare codons in the open reading frames than in non-coding regions. This indicates that the region may include an ORF. This is known as a 'Bibb' analysis.

3.3.5 Database Searching

'CODONPREFERENCE' analysis indicated the presence of a possible ORF in the regions of pICL92C that have been sequenced. The potential amino acid sequence was deduced (Fig 3.9) using a Bibb analysis to analyse the G+C% composition of the bases. Streptomyces codons show a distinctive MEDIUM-LOW-HIGH G+C% composition for the 3 bases within each codon of a coding region. The correct reading frame can therefore be deduced and the amino acid sequence predicted.

The amino acid sequences S1A, S1B and S3 were predicted and used to search against the GENEMBL database. S2, on analysis, did not appear to be a coding region.

The 3 peptide sequences all produced a good match with a potential ORF on the *Mycobacterium tuberculosis* genome. The *M. tuberculosis* sequencing project has cloned part of the genome into cosmid SCY22G8. The insert has been sequenced and the function of some of the genes have been deduced. Other potential ORFs have

Figure 3.10 Analysis of the sequenced DNA for open reading frames

The computer program 'CODONPREFERENCE' was used to analyse the DNA sequence for ORFs. The results are shown as a graphical plot which is divided into 3 sections representing the 3 possible reading frames. Each section shows 2 traces.

The top trace shows the G+C% for the third nucleotide of the codon in that frame. In a coding region of *Streptomyces* DNA this value is usually >90%.

The lower trace shows the codon preference when compared against a codon frequency table which has been compiled from previously sequenced streptomyces genes. A P value above 1 indicates that the given codons are used more frequently that would be expected in a random sequence.

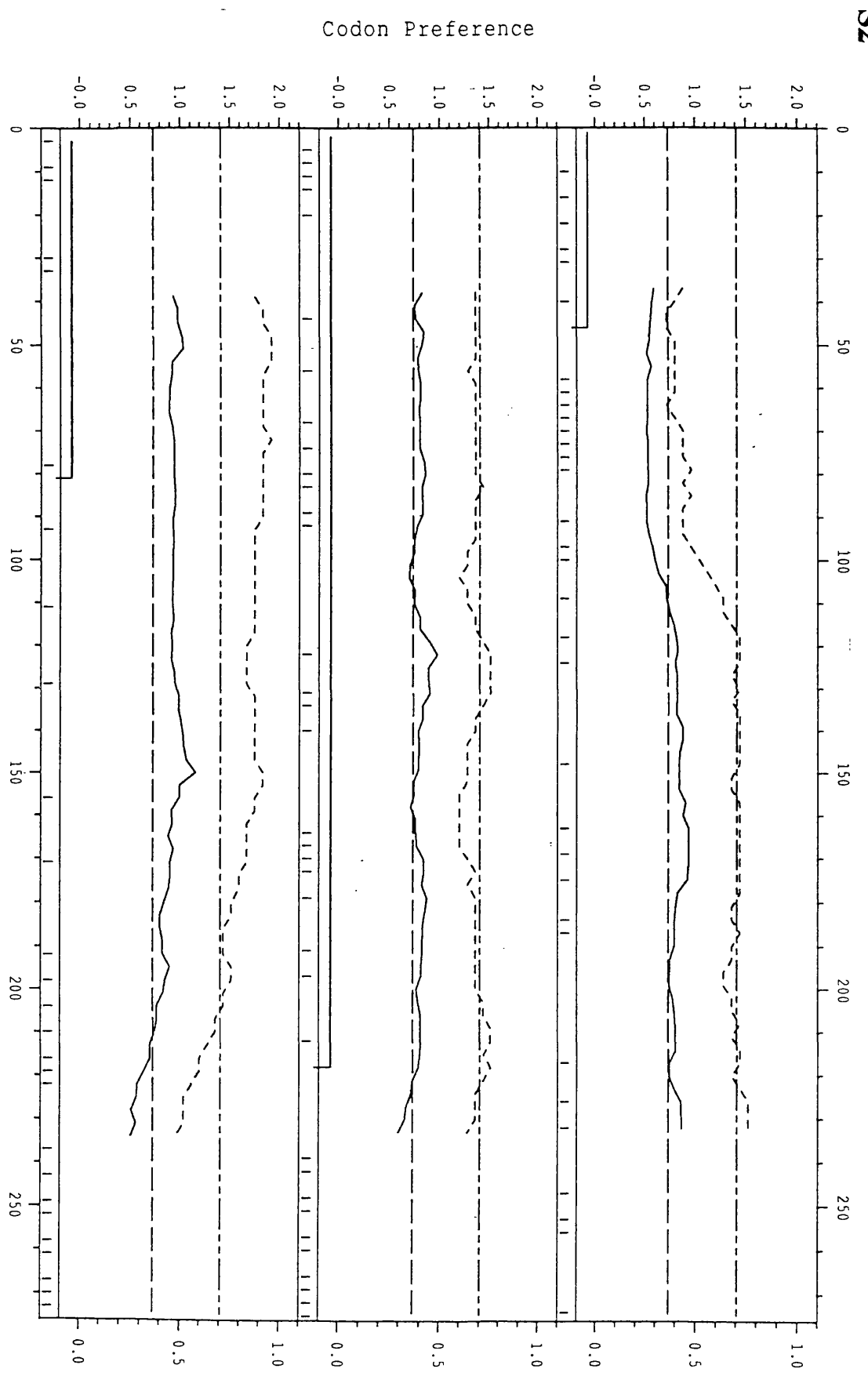
Rare codons are also shown on the trace as a horizontal mark at the bottom of each section.

4 plots are shown of the separate sections of sequence

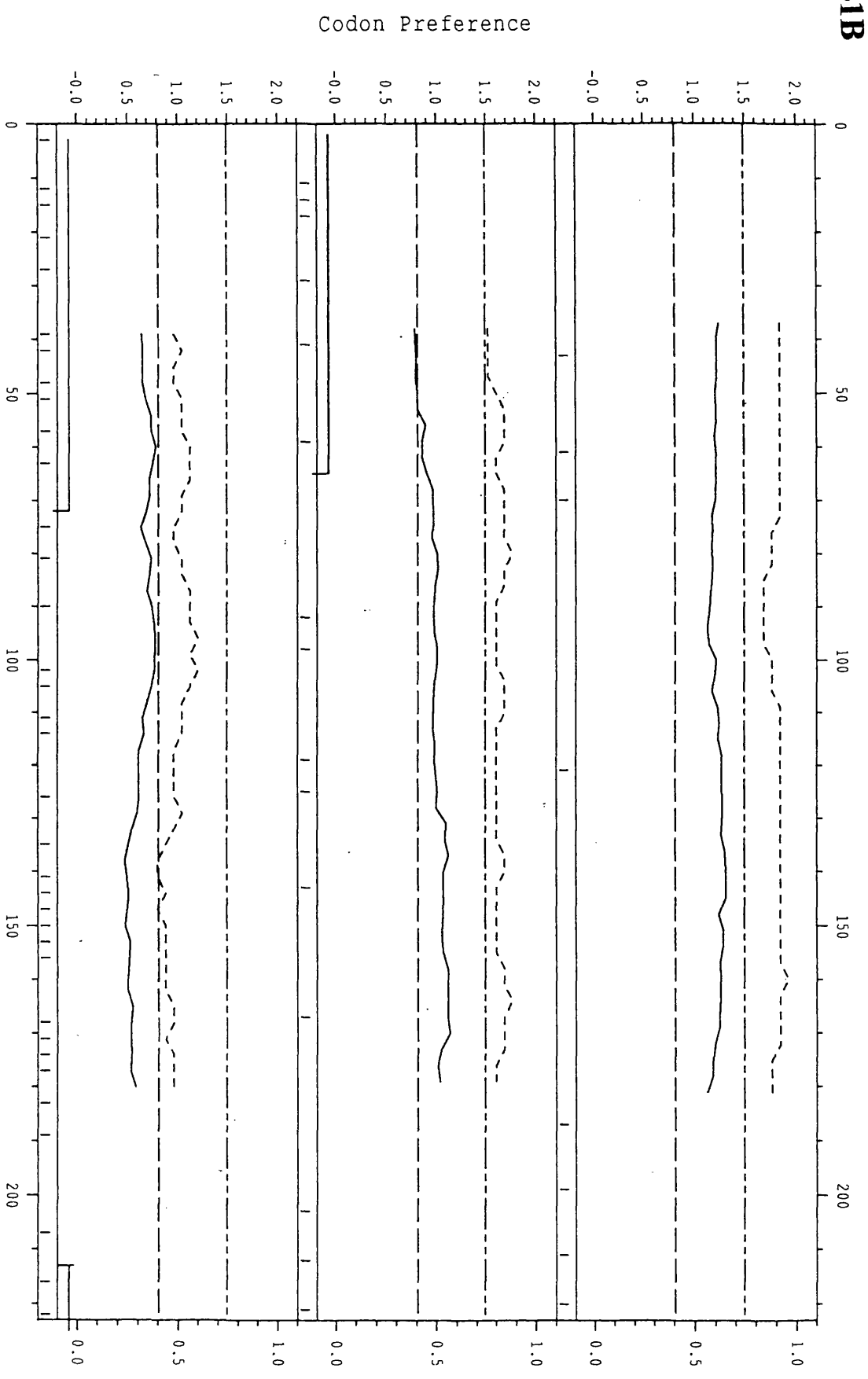
a) S2
b) S1B
c) S1A
d) S3

All the plots indicate the presence of an ORF in at least part of the sequence.

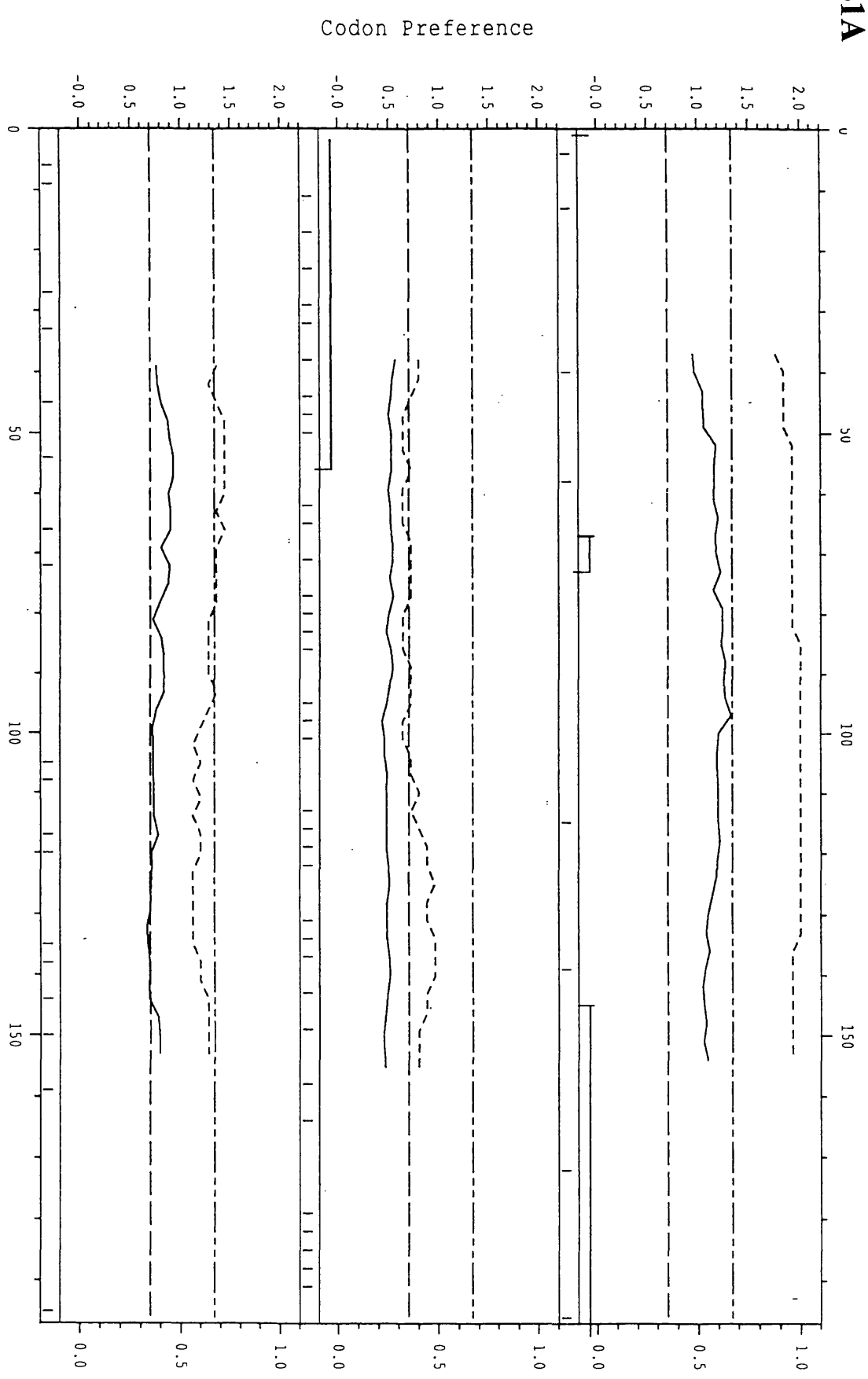
S2


Codon Preference
Third Position GC Bias
0
50
100
150
200
250
-0.0
0.5
1.0
1.5
2.0
0.0
0.5
1.0

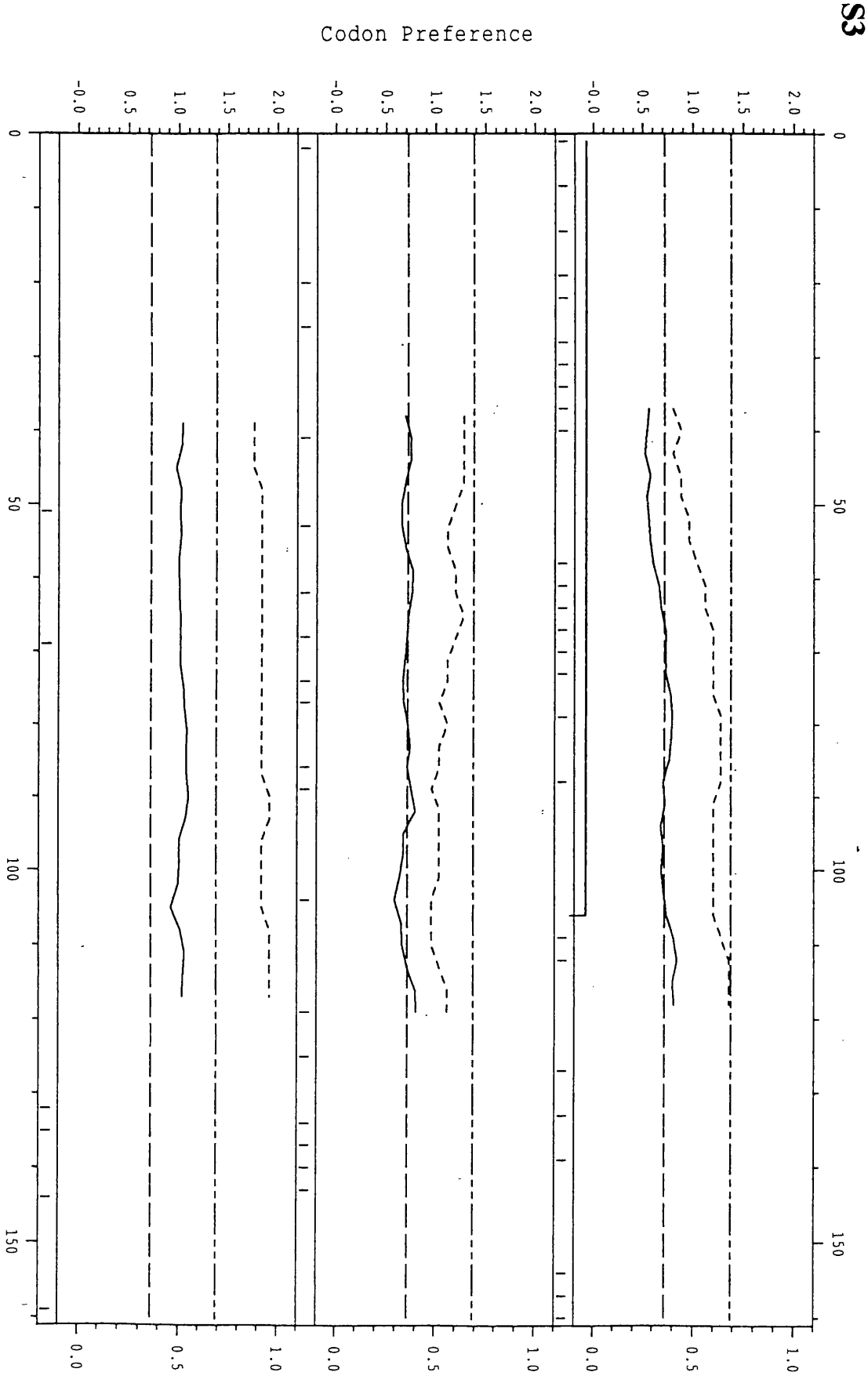
S1B



S1A



S3



been identified but their function is, as yet, unknown. It was with one of these potential ORFs that the deduced peptide sequences from ICL92C matched (Fig 3.11)

Figure 3.11 shows that good matches have been obtained between S1B and S1A and the *M.tuberculosis* ORF. A very close match was obtained for S3 but in a different frame from that predicted by the Bibb analysis. This analysis is very consistent for Streptomycete ORFs and therefore suggests that there may be a sequencing error in the *M.tuberculosis* ORF causing the ORF to be translated in the wrong frame in this section of DNA. The Bibb analysis did not indicate the presence of an ORF in S2 but a possible amino acid sequence matches the *M tuberculosis* in the region upstream of S1B suggesting that this region of *S.coelicolor* DNA may indeed be a coding region

The fact that potential ORFs (pORF) from both *M.tuberculosis* and *S.coelicolor* have been found to be very similar suggests that the these regions are coding regions. The role of these genes have yet to be determined.

In *M.tuberculosis* the pORF is situated upstream of the gene encoding citrate synthase (separated by 1600 bp and another unidentified pORF). Citrate synthase catalyses the formation of citric acid from oxaloacetate and acetyl CoA in the first step of the TCA cycle. This may be significant because in *S.coelicolor* the ORF is situated upstream of ICL and MS and therefore regulation of the TCA cycle and glyoxylate bypass may be linked in some way that we have yet to understand.

Figure 3.11 Comparison of pORFs from *S.coelicolor* and *M.tuberculosis*

The sections of potential amino acid sequences from pICL92C have been aligned with the potential polypeptide sequence from *Mycobacterium tuberculosis*.

The top line shows the complete *M.tuberculosis* amino acid sequence, below are the aligned sections of *S.coelicolor* sequence, S2, S1B, S1A and S3.

Those sections of *S.coelicolor* polypeptide sequence which are underlined are those of S2 and S3 which show a close match with the *M.tuberculosis* sequence but were not the amino acid sequences predicted by the Bibb analysis. The sequence for S3 has been translated in a different frame from that predicted by the Bibb analysis. This analysis suggested that S2 did not appear to be a coding region. A close match with the *M.tuberculosis* sequence suggests otherwise.

M. t. MTRSNVLPVARTYSRTFSGARLRRLRQERGLTQVALAKALDLSTSYVNQL

S. c. VSKTYAGARLRRG LSQAALARVLGISPSYLNQM

M. t. ENDQRPITVPVLLLLTERFDLSAQYFSSDSDARLVADLSDVFTDIGVEHA

S. c. EHDSRPLTVPVLLRLTETLGVDAAFFSERDTARLLADLREALTGELAVAR

M. t. VSGAQIEEFVARMPEVGHSLVAVHRRLRAATEELEGYRSRATAETELPPA

S. c. VSASDLAELASRMPAVAQVLIDLGRRNQLLSERLAGRRTDRWDR

M. t. RPMPFEEVRDFFYDRNNYIHDLDMAAERMFTESGMRTGGLDIQLAELMRD

M. t. RFGISVVIDDNLPDTAKRRYHPDTKVLRVAHWLMPGQRAFQIATQLALVG

M. t. QSDLISSIVATDDQLSTEARGVARIGLANYFAGAFLLPYREFHRAAEQLR

S. c. VQDFDPVPAHALARIGIANYFAAVLILPYTAFHAAVEEFR

M. t. YDIDLLGRRFGVGFETVCHRLSTLQRPRQRGIPFIFVRTDKAGNISKRQS

S. c. YDIERLTDHYGLGYETVGHRLSTL VDRAGNMSKRQS

M. t. ATAFHFSRVGGSCPLWVVHDAFAQPERIVRQVAQMPDGRSYSWVAKTTAA
S. c. ATGFHFSRAGGTCPLWNVYESFATPGRIHVQLSEMPDGQRY

M. t. DGLGYLGPHKNFAVGLGCDLAHAHKLVYTGVVLDDPSTEVPIGAGCKICC

M. t. NRRTSCAQRFAPYLGGRVAVDENAGSSLPYSSTEQSV

CHAPTER 4

General Discussion

4.1 Introduction

Streptomycetes produce a wide range of secondary metabolites many of which are of commercial and medical importance (particularly antibiotics). Many of these secondary metabolites are synthesised from intermediates of the TCA cycle. Polyketides, a particularly important group of antibiotics, are synthesised from acetyl-CoA units. If the TCA cycle is to continue, intermediates must be replaced. A variety of anaplerotic pathways exist in prokaryotes for this purpose, one of which is the glyoxylate bypass.

When *E.coli* is grown on acetate the glyoxylate bypass is induced, which allows the metabolism of acetyl-CoA units derived from acetate or fatty acids.. The enzymes involved are Isocitrate lyase and Malate synthase. ICL and MS form part of the *ace* operon.

Icl has been cloned from *S.coelicolor* and the gene encoding MS was found immediately downstream (Chapman, 1994) indicating that the glyoxylate bypass may be functional in this organism. Nothing is known about the transcription of these genes or the presence of other genes involved in the glyoxlate bypass close-by on the genome.

In *E.coli*, isocitrate dehydrogenase is important in controlling the division of flux between the TCA cycle and the glyoxylate bypass. The enzyme is regulated by phosphorylation/dephosphorylation reactions catalysed by ICDH kinase/phosphatase. The gene encoding ICDH in *S.coelicolor* has been cloned and the enzyme characterised (Taylor, 1992) but its regulation and role in controlling the glyoxylate bypass are unknown.

4.2 Attempts to clone *ms* from *S.coelicolor*

The gene encoding malate synthase was partially cloned by Chapman (1994) during the cloning of *icl*. Two methods were employed during this study in an attempt to clone the entire gene.

Genomic DNA was digested and probed with the section of the malate synthase gene already cloned. A plasmid sub-library was constructed in pUC18 which was used to transform *E.coli*. Despite the screening of 600 transformants no positives were obtained. Controls suggested that there may have been a problem with the cutting and ligation procedure which was never solved.

A lambda library of *S.coelicolor* genomic DNA was also screened using the homologous probe and this proved to be more successful. A number of positives were obtained and the DNA was purified. Restriction digests and a number of anomalous results suggested that the 3 clones chosen for further analysis did not contain the entire gene encoding malate synthase and certainly none of the region downstream which is also of interest. However, other positives were obtained in the initial screening of the lambda library and since the techniques were established it would have been relatively easy to carry out further screening in order to obtain the entire gene encoding malate synthase, had time been permitting.

4.3 Sequencing upstream of *icl* in *S.coelicolor*

The region upstream of *icl* was cloned by Chapman (1994) but was not sequenced. In *E.coli* other genes involved in the glyoxylate bypass (*aceK* and *iclR*) are situated

close to the genes encoding ICL and MS and it was therefore of interest to investigate this region.

The region was partially sequenced and analysis of this DNA segment suggested the presence of a potential open reading frame. Searching of a database found a strong match between the amino acid sequence and that of a potential open reading frame from *Mycobacterium tuberculosis*. This pORF was recognised as part of the *M.tuberculosis* sequencing project but its function is, as yet, unknown.

The pORF of *M.tuberculosis* is situated just downstream of the gene encoding citrate synthase which is an important enzyme of the TCA cycle. In *S.coelicolor* the pORF is situated upstream of the genes encoding ICL and MS, also important in central metabolism, so it could be possible that the pORF that has been identified may play some role which links these important pathways.

The pORF appears to be expressed in the opposite direction to *icl* and *ms* so it may be unconnected, however, there is only a short region of DNA separating the two genes and sometimes the genes encoding regulator proteins are located on the opposite strand from the main structural genes.

4.4 Future work

A number of options are open for further investigation of the pORF that has been identified. The entire gene could now be sequenced and the pORF isolated in order that it could be over-expressed. The protein could then be purified and analysed in order to determine its function. Also the gene could be disrupted in order to assess its function.

The entire gene encoding malate synthase remains to be cloned although this should be relatively easy using the *S.coelicolor* genomic DNA library in lambda. If this is achieved the region downstream could also be investigated to see if there are any other genes associated with glyoxylate bypass located in this region, as there are in *E.coli*. This could include genes encoding ICDHK/P or ICLR or similar regulatory genes. If they are not located near-by, then their presence in *S.coelicolor* or an alternative mechanism for regulation of ICDH and the glyoxylate bypass still needs to be investigated.

The expression of *ms* and *icl* could be investigated by S1 nuclease mapping to identify promoters and determine whether they are expressed as an operon. Chapman (1994) suggested that a stem-loop structure between the two genes may be a terminator or may affect the level of *ms* transcription.

Growth studies could be carried out to investigate the expression of ICDH, ICL and MS on different carbon sources which may give some indication of the control of these enzymes. It would also be of interest to investigate why growth of *S.coelicolor* on acetate was so poor.

Disruption of genomic copies of ICDH and ICL have been considered and preliminary work was carried out by Chapman (1994). It is possible that a second ICDH may exist, as it does in some prokaryotes. The disruption of the gene encoding ICL should prevent *S.coelicolor* from growing on Tween. However, an alternative pathway to the glyoxylate pathway has been identified in *S.collinus* and it is therefore possible that an ICL-negative *S.coelicolor* may be capable of employing this alternative pathway for growth on fatty acids and C_2 compounds.

Ultimately, it may be possible to have controllable copies on plasmid vectors so the effects of altering levels of the enzymes could be studied. This could eventually result in the ability to control the flux between central metabolism and secondary metabolite production.

REFERENCES

Adams, C.W., Fornwald, J.A., Schmidt, F.J., Rosenberg, M. and Brawner, M.E. (1988). Gene organisation and structure of the *Streptomyces lividans gal* operon. J. Bacteriol. **170,** 203-212.

Bennet, P.M. and Holms, W.H. (1975). Reversible inactivation of the isocitrate deydrogenase of *Escherichia coli* ML308 during growth on acetate. J. Gen. Microbiol. **87**, 37-51.

Berg, P. (1956). Acyl adenylates: an enzymic mechanism of acetae activation. J. Biol. Chem. **222**, 991-1013.

Bibb, M.J., Findlay, P.R. and Johanson, M.W. (1984). The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. Gene **30**, 157-166.

Binstock, J.F., Pramanik, A. and Schultz, H. (1977). Isolation of a multienzyme complex of fatty acid oxidation from *Escherichia coli* Proc. Natl. Acad. Sci. USA. **74,** 492-495.

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

Borthwick, A.C., Holms, W.H. and Nimmo, H.G. (1984a). Isolation of active and inactive forms of isocitrate dehydrogenase from *Escherichia coli* ML308. Eur. J. Biochem. **141,** 393-400.

Borthwick, A.C., Holms, W.H. and Nimmo, H.G. (1984b). The phosphorylation of *Escherichia coli* isocitrate dehydrogenase in intact cells. Biochem J. **222**, 797-804.

Borthwick, A.C., Holms, W.H. and Nimmo, H.G. (1984c). Amino acid sequence round the site of phosphorylation in isocitrate dehydrogenase from *Escherichia coli* ML308 FEBS Letters. **174,** 112-115.

Brice, C.B. and Kornberg, H.L. (1968). Genetic control of isocitrate lyase activity in *Escherichia coli.*, J. Bacteriol. **96**, 2185-2186.

Byrne, C., Stokes, H.W. and Ward K.A. (1988). Nucleotide sequence of the *aceB* gene encoding malate synthaseA in *E.coli.* Nucleic Acids Res. **16**, 9342.

Campbell, J.J.R., Smith, R.A. and Eagles, B.A. (1983). A deviation from the conventional tricarboxylic acid cycle in *Pseudomonas aeruginosa.* Biochem. Biophys. Acta. **11**, 594.

Celenza, J.L.and Carlson, M. (1986). A yeast gene that is essential for release from glucose repression encodes a protein kinase. Science **233**, 1175-1180.

Chapman, P.G. (1994). Studies of the glyoxylate bypass of *Streptomyces coelicolor* A3(2): purification of isocitrate lyase and cloning of the *icl* gene. Ph.D. Thesis. University of Glasgow.

Chater, K.F. (1984). Morphological and Physiological differentiation in Streptomyces. In: "Microbial Development." Losick, R and Shapiro. (eds). Cold Spring Harbour Laboratory Press, New York. 89-115

Chell, R.M., Sundaram, T.K. and Wilkinson, A.E. (1978). Isolation and characterisation of isocitrate lyase from a thermophilic *Bacilus* sp. Biochem. J. **173**, 165-177.

Chen, C.W., Lin, Y.S., Kieser, H.M. and Hopwood, D.A. (1993). The chromosomal DNA of *Streptomyces-lividans* 66 appears to be linear. J. Cell. Biochem. No.S17E, 291.

Chung, T., Klumpp, D.J. and LaPorte, D.C. (1988). Glyoxylate bypass operon of *Escherichia coli*: cloning and determination of the functional map. J. Bacteriol. **170**, 386-392.

Chung, T.W., Resnik, E., Stueland, C. and Laporte, D.C. (1993). Relative expression of the products of glyoxlate bypass operon - contributions of transcription and translation. J. Bacteriol, **175**, 4572-4575.

Cioni, M., Pinzauti, G. and Vanni, P. (1981). Comparative biochemistry of the glyoxylate cycle. Comp.Biochem Physiol. **70**, 1-26.

Cleveland, D.W., Fischer, S.G., Kirschner, M.W. and Laemmli, U.K. (1977). Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J. Biol. Chem. **252,** 1102-1106.

Cortay, J.C., Bleicher, F; Rieul, C, Reeves, H.C and Cozzone, A.J. (1988). Nucleotide sequence and expression of the *ace*K gene coding for isocitrate dehydrogenase kinase phosphatase in *Escherichia coli:* J. Bacteriol. **170**, 89-97.

Cortay, J.C., Bleicher, F., Duclos, B., Cenatiempo, Y., Gautier, C., Prato, J.L. and Cozzone, A.J. (1989). Utilisation of acetate in *Escherichia coli*: structural organisation and differential expression of the *ace* operon. Biochimie **71**, 1043-1049.

Cortay, J.C., Negre, D., Galinier, A., Duclos, B., Perriere, G. and Cozzone, A.J. (1991). Regulation of the acetate operon in *Escherichia-coli* - purification and functional characterization of the *iclR* Repressor. EMBO. **10**, 675-679.

Dean, A.M., Lee, M.H.I. and Koshland, D.E. (1989). Phosphorylation inactivates *Escherichia coli* isocitrate dehydrogenase by preventing isocitrate binding. J. Biol. Chem. **264**, 20482-20486

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

Diehl, P. and McFadden, D.A. (1993). Site-directed mutagenesis of lysine 193 in *Escherichia coli* isocitrate lyase by use of unique restriction enzyme site elimination. J. Bacteriol. **175**, 2263-2270.

Dirusso, C.C. and Nunn W.D. (1985). Cloning and characterization of a gene (*fadR*) involved in regulation of fatty-acid metabolism in *Escherichia-coli* J. Bacteriol. **161**, 583-588.

Durchschlag, H., Biedermann, G. and Eggerer, H. (1981). Large-scale purification and some properties of malate synthase from bakers yeast. Eur. J. Biochem. **114**, 255-262.

Ebbighausen, H, Weil, B and Kramer, R. (1991). Carrier-mediated acetate uptake in *Corynebacterium glutamicum*.. Arch Microbiol. **155**, 505-510.

El-Mansi, E.M.T., Nimmo, H.G. and Holms, W.H. (1985). The role of isocitrate in control of the phosphorylation of isocitrate dehydrogenase in *Escherichia coli* ML308 FEBS letts **183,** 251-255.

El-Mansi, E.M.T., Mackintosh, C., Duncan, K., Holms, W.H. and Nimmo, H.G. (1987). Molecular cloning and over-expression of the glyoxylate bypass operon from *Escherichia coli* ML308. Biochem. J. **242**, 661-665.

Fornwald, J.A., Schmidt, F.J., Adams, C.W., Rosenberg, M. and Brawner, M.E. (1987). Two promoters, one inducible and one constitutive, control transcription of the *Streptomyces lividans* galactose operon. Proc. Natl. Acad. Sci. USA **84** 2130-2134.

Frerman, F.E. (1973). The role of acetyl-coenzyme A in the transferase uptake of butyrate by isolated membrane vesicles of *Escherichia coli.* Arch. Biochem. Biophys. **159**, 444-452.

Galinier, A., Bleicher, F., Negre, D., Perriere, G., Duclos, B., Cozzone, A.J. and Cortay, J.C. (1991). Primary structure of the intergenic region between *ace*K and *icl*R in the *Escherichia coli* chromosome. Gene. **97,**149-150.

Garland, D. and Nimmo, H.G. (1984). A comparison of the phosphorylated and unphosphorylated forms of isocitrate dehydrogenase from *Escherichia coli* ML308. FEBS Letts, **165**, 259-264.

Garnak, M. and Reeves, H.C. (1979a). Purification and properties of phosphorylated isocitrate dehydrogenase of *Escherichia coli.* Science **203,** 1111-1112.

Garnak, M. and Reeves, H.C. (1979b). Purification and properties of phosphorylated isocitrate dehydrogenase of *Escherichia coli.* J. Biol. Chem. **254**, 7915-7920.

Gibson, T.J. (1984). Studies on the Epstein Barr virus genome. PhD Thesis, Cambridge, England.

Gribskov, M., Devereux, J. and Burgess, R.R. (1984). The codon preference plot: graphic analysis of protein coding sequences and prediction of gene expression. Nucleic Acids Res. **12**, 539 549.

Han, L. andReynolds, K.A. (1997). A novel anaplerotic pathway to the glyoxylate cycle in Streptomycetes. J.Bacteriol. **179**, 5157-5164.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**, 557-580.

Holms, W.H. (1987). Control of flux through the citric acid cycle and the glyoxylate bypass in *Escherichia coli.* Biochem. Soc. Symp. **54**, 17-31

Holms, W.H. and Bennet, P.M. (1971). Regulation of isocitrate dehydrogenase activity in *Escherichia coli* on adaptation to acetate. J. Gen. Microbiol. **65**, 57-68.

Hopwood, D.A., Bibb, M.J., Chater, K.F., Kieser, T., Bruton, C.J., Kieser, H.M., Lydiate. D.J., Smith, C.P., Ward, J.M. and Schrempf, H. (1985). Genetic manipulation of *Streptomyces* a laboratory manual. The John Innes Foundation, Norwich England.

Hopwood, D.A. and Kieser, T. (1990). The *Streptomyces* genome. In "The Bacterial Chromosome." Drlica, K. and Riley, M. (ed). American Society for Microbiology, Washington DC. 147-162.

Hoyt, J.C., Robertson, E.F., Berlyn, K.A. and Reeves, H.C. (1988). *Escherichia coli* isocitrate lyase: properties and comparison. Biochem Biophys. Acta. **966**, 367-371.

Hoyt, J.C. and Reeves, H.C. (1992). Phosphorylation of *acinetobacter* isocitrate lyase. Biochem Biophys Res Comm. **182**, 367-371.

Hoyt, J.C., Lin, H.P.P. and Reeves, H.C. (1994). *In-vivo* phosphorylation of isocitrate lyase in *Escherichia coli* and *acinetobacter calcoaceticus*. Curr. Microbiol. **28**, 67-69.

Hurley, J.H., Dean, A.M., Sohl, J.L., Koshland, D.E. and Stroud, R.M. (1990). Regulation of an enzyme by phosphorylation at the active site. Science. **249**, 1012-1016.

Huttner, S., Mecke, D. and Frohlich, K.U. (1997). Gene cloning and sequencing and enzyme purification of the malate synthase of *Streptomyces arenae*. Gene. **188** 239-246.

Ishii, A., Ochiai, T., Imagawa, S., Fukunaga, N., Sasaki, S., Minowa, O., Mizuno, Y. and Shiokawa, H. (1987). Isozymes of isocitrate dehydrogenase from an obligately psychrophilic bacterium, *Vibrio* sp. strain ABE-1: purification and modulation of activities by growth conditions. Biochem. J. **102**, 1489-1498.

Jenkins, L.S. and Nunn, W.D. (1987). Genetic and molecular characterization of the genes involved in short chain fatty acid degradation in *Escherichia coli* - the *ato* system. J. Bacteriol. **169**, 42-52.

Kieser, H.M., Kieser, T. and Hopwood, D.A. (1992). A combined genetic and physical map of the *streptomyces-coelicolor* A3 (2) chromosome. J. Bacteriol. **174**, 5496-5507.

Klumpp, D.J., Plank, D.W., Bowdin, L.J., Stueland, C.S., Chung, T. and Laporte, D.C. (1988). Nucleotide sequence of *ace*K, the gene encoding isocitrate dehydrogenase kinase phosphatase. J. Bacteriol. **170**, 2763-2769.

Kornberg, H.L. (1966a). Anaplerotic sequences and their role in metabolism. Essays Biochem. **2**, 1-31.

Kornberg, H.L. (1966b). The role and control of the glyoxylate cycle in *Escherichia coli*. Biochem. J. **99**, 1-11.

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

LaPorte, D.C. (1993). The isocitrate dehydrogenase phosphorylation cycle; regulation and enzymology. J. Cell. Biochem. **51**, 14-18.

LaPorte, D.C., Thorsness, P.E. and Koshland, D.E. (1985). Compensatory phosphorylation of isocitrate dehydrogenase - a mechanism for adaptation to the intracellular environment. J. Biol. Chem. **260,** 563-568

LaPorte, D.C. and Koshland, D.E.Jr. (1982). A protein with kinase and phosphatase activities involved in the regulation of tricarboxylic acid cycle. Nature (London) **300**, 458-460.

LaPorte, D.C. and Koshland, D.E.Jr. (1983). Phosphorylation of isocitrate dehydrogenase as a demonstration of enhanced sensitivity in covalent regulation. Nature **305**, 286-290.

LaPorte, D.C. and Chung, T. (1985). A single gene codes for the kinase and phosphatase which regulate isocitrate dehydrogenase. J. Biol. Chem. **260**, 15291-15297.

Lemarie, H.G. and Muller-Hill, B. (1986). Nucleotide sequences of the *gal*E and the *gal*T gene of *E. coli*. Nucleic Acids Res. **14**, 7705-7711.

Mackintosh, C. and Nimmo, H.G. (1988). Purification and regulatory properties of isocitrate lyase from *E.coli*. ML308 Biochem. J. **250**. 25-31.

Malloy, P.J., Reeves, H.C. and Spiess, J. (1984). Amino-acid sequence of the phosphorylation site of isocitrate dehydrogenase from *Escherichia coli*. Curr. Microbiol. **11**, 37-42.

Maloy, S.R., Bohlander, M. and Nunn, W.D. (1980). Elevated levels of glyoxylate shunt enzymes in *Escherichia coli* strains constitutive for fatty acid degradation. J. Bacteriol. **143**, 720-725.

Maloy, S.R. and Nunn, W.D. (1981). Role of gene *fad*R in *Escherichia coli* acetate metabolism. J.Bacteriol. **148**, 83-90.

Maloy, S.R. and Nunn, W.D. (1982). Genetic regulation of the glyoxylate shunt in *Escherichia-coli* K-12. J. Bacteriol. **149**, 173-180.

Matsuoka, M. and Mcfadden, B.A. (1988). Isolation, hyperexpression, and sequencing of the *ace*A gene encoding isocitrate lyase in *Escherichia coli*. J. Bacteriol. **170**, 4528-4536.

Mattern, S.D., Brawner, M.E. and Westpheling, J. (1993). Identification of a complex operator for *gal*P1, the glucose sensitive galactose-dependant promoter of the *Streptomyces* galactose operon. J. Bacteriol. **175**, 1213-1220.

McFadden, B.A., Rao, G.R., Cohen, A.L. and Roche, T.E. (1968). Isocitrate lyase from *Pseudomonas indigofera*. V. Subunits and terminal residues and the relation to catalytic activity. Biochem **7**, 3574-3582.

Mizosawa, S., Nichimora, S. and Seda, F. (1986). Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. Nucleic Acids Res. **14**, 1319-1324.

Murray, N.E., Brammer, W.J. and Murray, K. (1977). Lambda phage that simplify the recovery of *in vitro* recombinants. Mol. Gen. Genet. **160**, 53-61

Negre, D., Cortay, J.C., Old, I.G., Galinier, A., Richaud, C., Saintgirons, I. and Cozzone, A.J. (1991). Overproduction and characterization of the *ic/R* gene-product of *Escherichia coli* K-12 and comparison with that of *Salmonella-typhimurium* LT2. Gene **97**, 29-37.

Negre, D., Cortay, J.C., Galinier, A., Sauve, P. and Cozzone, A.J. (1992). Specific interactions between *ic/R* repressor of the acetate operon of *E. coli* and its operator. J Mol. Biol. **228**, 23-29.

Nimmo, G.A. and Nimmo, H.G. (1984). The regulatory properties of isocitrate dehydrogenase kinase and isocitrate dehydrogenase phosphatase from *Escherichia coli* ML308 and the roles of these activities in the control of isocitrate dehydrogenase. Europ. J. Biochem., **141**, 409-414.

Nimmo, H.G., Borthwick, A.C., El-Mansi, E.M.T., Holms, W.H.; Mackintosh, C. and Nimmo, G.A. (1987). Regulation of enzymes at the branchpoint between the citric acid cycle and the glyoxylate bypass in *Escherichia coli.* Biochem. Soc. Symp. **54**, 93-101.

Nimmo, H.G., Douglas, F., Kleanthos, C., Campbell, D.G. and Mackintosh, C. (1989). Identification of a cysteine residue at the active site of *Escherichia coli* isocitrate lyase. Biochem. J. **261**, 431-435.

Nunn, W.D. (1986). A molecular view of fatty acid catabolism in *Escherichia coli.* Microbiol. Reviews **50**, 179-192.

Olson, J.A. (1954). The D-isocitric lyase system: the formation of glyoxylic and succinic acids from D-isocitric acid. Nature **174**, 695-696.

Omura, S. (1986). Studies on macrolides and new antibiotics. Yakugaku Zasshi. J. Pharmaceut. Soc. Japan, **106**, 729-757.

Overath, P., Pauli. G. and Schairer, H.U. (1969) . Fatty acid degradation in *Escherichia coli.* An inducible acyl-CoA synthetase, the mapping of *old* mutations and the isolation of regulatory mutants, Eur. J. Biochem. **7**, 559-574.

Pawar, S. and Schultz, H. (1981). The structure of the multienzyme complex of fatty acid oxidation from *Escherichia coli.* J. Biol. Chem. **256**, 137-142.

Pramanik, A., Pawar, S., Antonian, E. and Schultz, H. (1979). Five different enzymic activities are associated with the multienzyme complex of fatty acid oxidation from *Escherichia coli*. J. Bacteriol. **137**, 469-473.

Rieul, C., Bleicher, F., Duclos, B., Cortay, J.C. and Cozzone, A.D. (1988). Nucleotide sequence of the *ace*A gene coding for isocitrate lyase in *Escherichia coli*. Nucleic Acids Res. **16**, 5689-5690.

Robertson, E.F., Hoyt, J.C. and Reeves, H.C. (1987a). Evidence of histidine phosphorylation in isocitrate lyase from *Escherichia-coli*. Federation Proceedings. **46**, 1952.

Robertson, E.F., Hoyt, J.C. and Reeves, H.C. (1987b). *In vitro* phosphorylation of *Escherichia coli* isocitrate lyase. Current Microbiol. **15**, 103-105.

Robertson, E.F. and Reeves, H.C. (1987). Purification and characterization of isocitrate lyase from *Escherichia coli*. Current Microbiol. **14**, 347-350.

Ruchti, M. and Widmer, F. (1986). Isocitrate lyase in germinating soybean cotyledons: effects of metabolites *in vitro* J. Plant Physiol. **123,** 507-511.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular Cloning, 2nd edition. Cold Spring Harbor Press.

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

Saz, H.J. (1954). The enzymic formation of glyoxylate and succinate from tricarboxylic acids. Biochem. J. **58**, 112

Schumperli, D., Mckenney, K., Sobieski, D.A. and Rosenberg, M. (1982). Translational coupling at an intercistronic boundary of the *Escherichia coli* galactose operon. Cell **30**, 865-871.

Seno, E.T. and Chater, K.F. (1983). Glycerol catabolic enzymes and their regulation in wild type and mutant strains of *Streptomyces coelciolor* A(3)2. J. Gen. Microbiol. **129,** 1403-1413.

Seno, E.E., Bruton, C.J. and Chater, K.F. (1984). The glycerol utilisation operon of *Streptomyces coelicolor*: genetic mapping of *gyl* mutations and the analysis of cloned *gyl* DNA . Mol. Gen. Genet. **193**, 119-218.

Shapiro, J.A. and Adhya, S.L. (1969). The galactose operon of *E.coli* K-12II. A deletion analysis of operon structure and polarity. Genetics **62**, 249-264.

Shine, J. and Dalgarno, L. (1974). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc Nat. Acad. Sci. **71**, 1342-1346.

Simons, R.W., Egan, P.A., Chute, H.T. and Nunn, W.D. (1980). Regulation of fatty acid degradation in *Escherichia coli*: isolation and characterisation of strains bearing insertion and temperature-sensitive mutations in gene *fad*R. J. Bacteriol. **142**, 621-632.

Smith, C.P. and Chater, K.F. (1988). Structure and regulation of controlling sequences for the *Streptomyces coelicolor* glycerol operon. J. Mol. Biol. **204,** 569-580.

Smith, C.P. and Chater, K.F. (1988). Cloning and transcription analysis of the entire glycerol utilization (*gy*/abx) operon of *Streptomyces coelicolor* A3(2) and identification of a closely associated transcription unit . Mol. Gen. Genet. **211**, 129-137.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J.Mol. Biol. **98**, 503-517.

Spratt, S.K., Black, P.N., Ragozzino, M.M. and Nunn, W.D. (1984). Cloning, mapping, and expression of genes involved in the fatty acid degradative multienzyme complex of *Escherichia coli*. J. Bacteriol. **158**, 535-542.

Stueland, C.S., Ikeda, T.P. and LaPorte, D.C. (1989). Mutation of the predicted ATP binding-site inactivates both activities of isocitrate dehydrogenase kinase/phoaphatase. J. Biochem. 264, 13775-13779.

Sunnarborg, A., Klumpp, D., Chung, T. and Laporte, D.C. (1990). Regulation of the glyoxylate bypass operon - cloning and characterization of *icl*R. J. Bacteriol. **172**, 2642-2649.

Taylor, R.D. (1992). Purification and characterisation of the isocitrate dehydrogenase from *Streptomyces coelicolor* and cloning of its gene. Ph.D. Thesis. University of Glasgow.

Thorsness, P.E. and Koshland, D.E. Jr. (1987). Inactivation of isocitrate dehydrogenase by phosphorylation is mediated by the negative charge of the phosphate. J. Biol. Chem. **262**, 10422-10425.

Uchida, M., Ueda, M., Matsuki, T., Okada, H., Tanaka, A. and Fukui, S. (1986). Properties of isocitrate lyase from an alkane utilisable yeast, *Candida tropicalis* Agric. Biol. Chem. **50**, 127-134.

Vanderwinkel, E., Liard, P., Ramos, F. and Wiame, J.M., (1963). Genetic control of the regulation of isocitritase and malate synthase in *Escherichia coli* K-12. Biochem. Biophys. Res. Comm. **12**, 157-162.

Vanderwinkel, E. and DeVliegher, M. (1968). Physiologie et genetique de L'isocitrase et des malate synthases chez *Escherichia coli.* Eur. J. Biochem. **5**, 81-90.

Wilson, R.B. and Maloy, S.R. (1987). Isolation and characterization of *Salmonella typhimurium* glyoxylate shunt mutants. J. Bacteriol. **169**, 3029-3034.

Wong, D.T.O. and Ajl, S.J. (1956). Conversion of acetate and glyoxylate to malate. J. Am. Chem. Soc. **78**, 3230.

Yang, B., Gerhardt, S.G. and Larson, T.J.; (1997). Action at a distance for *glp* repressor control of *glpTQ* transcription in *Escherichia coli* K-12. Molec Microbiol. **24**, 511-521.