

ANALYSIS OF CLONED GENES FOR AROMATIC
CATABOLISM: THE HYDROXYPHENYLACETATE
GENES OF *Escherichia coli* AND
Klebsiella pneumoniae

SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE DEPARTMENT OF BIOCHEMISTRY,
UNIVERSITY OF LEICESTER.

TONY FAWCETT.

MAY 1989

UMI Number: U022608

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U022608

Published by ProQuest LLC 2015. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346



4168478914

41449414x251594414

ACKNOWLEDGEMENTS

I am grateful to Ron Cooper for allowing me to do this work in his laboratory and for providing a stimulating environment in which my scientific skills could develop . I would also like to thank my contemporaries in Lab. 108 for making my stay a happy one, especially Suzanne, Sudjadi and Dave. I am indebted to many people in Leicester, in many departments, for helpful discussions - I thank them all.

Finally, I thank my parents for continual encouragement and assistance, without which none of this would have been possible.

CONTENTS—TEXT.

PAGE

1	INTRODUCTION:
2	1.1 AROMATIC CATABOLISM
3	General Mechanisms of Aromatic Catabolism
6	Convergence of Aromatic Catabolic Pathways
9	1.2 ORGANISATION OF CATABOLIC OPERONS
11	1.3 REGULATION OF CARBON CATABOLIC PATHWAYS
11	Global Regulation
14	Regulation of Transcription
15	Regulation by RNA Processing
16	Translational Regulation
17	1.4 EVOLUTION OF METABOLIC PATHWAYS
19	Transfer of DNA between Species
20	Conservation of Pathways between Species
21	1.5 GENETIC ANALYSIS OF CARBON CATABOLIC PATHWAYS
23	Chromosomally Encoded Catabolic Pathways
24	1.6 THE HYDROXYPHENYLACETATE AND HOMOPROTOCATECHUATE PATHWAY
26	Regulation of the HPA Pathway
30	Analysis of Cloned <i>hpc</i> Genes
37	AIMS
38	MATERIALS AND METHODS
39	2.1 BIOCHEMICAL METHODS
39	Growth Media and Conditions
42	Cell Free Extracts
42	Protein Estimations
42	Washed Cell Suspensions and Measurement of Oxygen Evolution
43	Preparation of Intermediates of the HPC Pathway
44	Enzyme Assays
46	SDS PAGE Analysis
47	Fast Protein Liquid Chromatography (FPLC)
47	Amino Acid Sequencing
47	Oligonucleotide Synthesis
48	2.2 GENETIC PROCEDURES
48	Preparation of Chromosomal DNA
48	Preparation of Plasmid DNA

50	Agarose Gel Electrophoresis
50	Restriction Enzyme Digestion
51	Phosphatase Treatment
51	Isolation of DNA Fragments from Agarose Gels
52	Ligation of DNA
52	Transformation of <i>E. coli</i> with Plasmid DNA
53	Preparation of Total RNA
54	Formaldehyde/Agarose Gel Electrophoresis
55	2.3 BLOTTING AND RADIOLABELLING TECHNIQUES
55	Transfer of Nucleic Acids to Nitrocellulose
55	Colony Hybridisation
57	Radiolabelling of Oligonucleotide Probes
57	Radiolabelling of DNA Restriction Fragments
58	Hybridisation of Radiolabelled Probes to Nucleic Acids
60	2.4 DNA SEQUENCING
61	Analysis of DNA Sequence Data
63	ORGANISATION AND EXPRESSION OF <i>Escherichia coli</i> <i>hpc</i> GENES
64	3.1 FURTHER INVESTIGATIONS INTO THE ORGANISATION OF THE <i>hpc</i> GENES
64	Southern Blot Analysis to Identify the Position of the <i>hpcG</i> Gene
68	DNA Sequencing of the 5' End of the <i>hpcG</i> gene
71	Subclones of pJJ801 and pJJ200 to Identify <i>hpc</i> Gene Order
76	Localisation of the 5' End of the <i>hpcC</i> Gene
79	Location of the <i>hpcF</i> Gene
82	Ordering the <i>hpc(EF)</i> Genes
84	Ordering the <i>hpc(BD)</i> Genes
	3.2 EXPRESSION OF THE <i>hpc</i> GENES
86	Direction of Transcription of the <i>hpc</i> Genes
86	Transcription of <i>hpc(BD)</i> Genes
89	Evidence for Two Transcriptional Units on pJJ801
91	Northern Blot Analysis of <i>hpc</i> Transcripts
98	3.3 MODEL OF ORGANISATION AND EXPRESSION OF <i>hpc</i> GENE OF <i>E. coli</i> C
98	The Gene Order
101	Expression of the <i>hpc</i> Genes
102	Direction of Transcription

104	INVESTIGATIONS INTO THE RELATEDNESS OF <i>hpc</i> PATHWAY GENES IN DIFERENT BACTERIA
105	4.1 PURIFICATION AND CHARACTERISATION OF THE CHMS DEHYDROGENASE FROM <i>E. coli</i> C
106	Characterisation of Purified CHMS Dehydrogenase
119	4.2 HOMOLOGIES BETWEEN THE CHMS DEHYDROGENASE ENZYMES OF <i>E. coli</i> AND <i>K. pneumoniae</i>
121	Comparison of Purified CHMS Dehydrogenases
124	Southern Blot Analysis of CHMS Dehydrogenase Genes Using a Probe Derived From the Amino Acid Sequence
126	4.3 CLONING AND PRIMARY ANALYSIS OF <i>hpc</i> GENES FROM <i>Klebsiella pneumoniae</i> M5a1
126	Isolation of Clones
132	Complementation of <i>E. coli</i> C Mutants of the HPC Pathway with Putative Clones
135	Enzyme Activities Associated with the Clones
137	Comparison of pTF100 and pTF102
139	Number of Copies of <i>hpc</i> Genes on the Chromosome
141	Oxidation of HPA and HPC by Whole Cells of 5K(pTF100)
141	Model of Regulation of <i>Klebsiella hpc</i> Genes
143	Comparison of Cloned <i>Escherichia</i> and <i>Klebsiella hpc</i> Genes
146	Future Prospects
147	REFERENCES

CONTENTS—FIGURES.

PAGE	FIGURE	
4	1.1	Fate of Aromatic Ring Substituents Prior to Ring Cleavage
5	1.2	Aromatic Ring Hydroxylation Reactions
7	1.3	Ring Cleavage Reactions
8	1.4	<i>meta</i> -Cleavage Pathway Encoded by TOL Plasmid pWWO
10	1.5	The Convergence of Catabolic Intermediates in the β -Ketoadipate Pathway
12	1.6	Locations and Organisation of known Arabinose Transport and Catabolic Genes
25	1.7	Restriction Map of the <i>A. calcoaceticus ben-cat</i> Region Showing Supraoperonic Clustering
27	1.8	Feasible Mechanisms Involved in the Catabolism of HPC (Dagley, 1975)
28	1.9	Pathway for the Degradation of HPC in <i>E. coli</i> C (Jenkins, 1987)
31	1.10	Primary clones Isolated by Jenkins Containing <i>hpc</i> Genes
33/34	1.11	Summary of the Subclones of pJJ200/pJJ801 Used to Determine the Order of the <i>hpc</i> Genes
36	1.12	Model of the Organisation of <i>hpc</i> Genes as Proposed by Jenkins
41	2.1	The Plasmid Vectors pUC18 and pUC19
56	2.2	Apparatus for the Transfer of Nucleic Acids to Nitrocellulose
65	3.1a	Sequence from the First 34 Residues of the Amino Terminus of OHED Hydratase
65	3.1b	Sequence of the Synthetic Oligonucleotides Corresponding to the First Six amino Acid Residues of OHED Hydratase
67	3.2a	DNA-DNA Dot Blot Hybridisation of Various Plasmids with Probe G
67	3.2b	Southern blot Analysis of Restriction Fragments of Various Plasmids with Probe G.
69	3.3	Localisation of the 5' End of the <i>hpcG</i> Gene
70	3.4	Sequence strategy Used to Identify the <i>hpcG</i> Gene
72	3.5	Interpretation of the DNA Sequence Information
74b		Diagram of Plasmids Represented in Table 3.1
77	3.6a	Sequence of the First 34 Residues from the Amino Terminus of CHMS Dehydrogenase
77	3.6b	Sequence of the Synthetic Oligonucleotide Corresponding to Residues 5-10 of the Amino terminus of CHMS Dehydrogenase

78	3.7	Southern Blot Analysis of pJJ801 with Probe C
80	3.8	Construction of the Subclones pTF305 and pTF306
93	3.9	Northern Blot Analysis of Total RNA from <i>E. coli</i> C and <i>E.coli</i> 5K (pJJ200) grown in the Presence and Absence of HPA
95	3.10	Laser Densitometric Scans of Northern Hybridisations
99/1003.11		Summary of Constructs used to Identify the <i>hpc</i> Gene Order
103	3.12	Model for the Organisation and Expression of <i>E. coli</i> C <i>hpc</i> Genes
107	4.1	Elution of Proteins from a HR 10/10 Mono Q Column
108	4.2	Elution of CHMS Dehydrogenase from a HR 5/5 Phenyl Superose Column
109	4.3	SDS-PAGE gel Showing Steps in the Purification of CHMS Dehydrogenase from <i>E. coli</i> C
110	4.4	SDS-PAGE Gel Showing Purity of CHMS Dehydrogenase
112	4.5	A Graph Showing the calibration Curve of the SDS-PAGE Gel in Figure 4.4
114	4.6	Calibration Graph of the Superose 12 Gel Filtration Column
117	4.7	Graph Showing the Variation in Enzyme Specific Activity with pH
118	4.8	Sequence from the First 34 Residues of the Amino Terminus of <i>E. coli</i> CHMS Dehydrogenase
120	4.9	Amino Terminal Amino acid Sequences of the CHMS Dehydrogenases of <i>E. coli</i> C and <i>K. Pneumoniae</i> M5a1
122	4.10	Kinetic Plots of v Against S for CHMS Dehydrogenases of <i>E. coli</i> C and <i>K. pneumoniae</i> M5a1
125	4.11	Southern Blot of <i>Bam</i> H1 Digested DNAs Hybridised with Probe C
127	4.12	Southern Hybridisation of M5a1 DNA with Probe C
129	4.13	Gel Showing the Size Fractionation of Chromosomal DNA
131	4.14	Autoradiogram of a Secondary Screen with Probe C
133	4.15	Restriction Analysis of Putative Positive Clones
138	4.16a	Restriction Digest Patterns of pTF100 and pTF102
138	4.16b	Southern Blot Analysis of pTF100 and pTF102 with Probe C
140	4.17	Southern Blot Analysis of M5a1 Chromosomal DNA Hybridised with pTF100 and pTF102 Insert DNA
145	4.18	Restriction Digests of pTF100 and pTF102

CONTENTS—TABLES.

PAGE	TABLE	
40	2.1	Bacterial Strains and Plasmids
62	2.2	Dideoxy-Deoxynucleotide Sequencing Mixtures
74	3.1	Activities of <i>E. coli</i> 5K Harboursing Various Subclones of pJJ801
81	3.2	Enzyme Activities with Subclones of pJJ801 When Harboured in 5K
85	3.3	Enzyme Activities Detected in <i>Ba</i> 31 Deletion Subclones
88	3.4	Activities of HPC Catabolic Enzymes of Various Subclones in pUC Vectors: the Effect of IPTG
90	3.5	Enzyme Activities Associated with Subclones of pJJ801 when Harboured in 5K and JJ221
96	3.6	To Show the Percent Intensities of Hybridisation of pJJ200 insert DNA to Total RNA from Various Sources
96b	3.7	To Show the Increase in Enzyme Specific Activity with Increased Optical Density of a Culture of 5K(pJJ200)
111	4.1	A Summary of the Purification of CHMS Dehydrogenase from <i>E. coli</i> C
115	4.2	Molar Extinction Coefficients of CHMS at Different pH Values
123	4.3	Summary of Known Characteristics of Purified CHMS Dehydrogenases
134	4.4	Complementation of Various Strains by Cloned <i>Klebsiella</i> DNA
136	4.5	Activities Associated with M5a1 and 5K(pTF100/pTF102) Under Non-Inducing and Inducing Conditions
142	4.6	Oxidation of HPA and HPC by Whole Cells

STATEMENT.

This thesis submitted for the degree of Doctor of Philosophy entitled: Analysis of cloned genes form aromatic catabolism: the hydroxyphenylacetate genes of *Escherichia coli* and *Klebsiella pneumoniae* is based upon work conducted by the author mainly during the period between 1985 and 1988.

All work recorded in this thesis is original unless otherwise acknowledged in the text or by references.

None of the work has been submitted for another degree in this or any other University.

Signed Tawcett

Date 06/06/89.

1: INTRODUCTION.

2) the 'design' of microorganisms by genetic engineering techniques to degrade otherwise recalcitrant compounds. In principle, it is possible to create a hybrid catabolic plasmid by cloning chromosomally and /or plasmid encoded genes from different pathways into a suitable vector, such that the result would be a novel catabolic sequence. Although the design of biodegradable aromatic-compound containing products is now a regular event, the design of specific microorganisms to degrade otherwise recalcitrant compounds is only just beginning.

General Mechanisms of Aromatic Catabolism.

Man-made and naturally occurring aromatic compounds often contain substituted side chain groups which are either changed enzymically so that they can be catabolised via a mainstream ring fission compound or they remain intact and are catabolised via a substituted ring fission compound. The simplest example of this is when a single methyl group is present on the ring, as in toluene (fig. 1.1). The methyl group can be sequentially oxidised to a carboxyl group giving benzoate which is then converted into the ring fission substrate catechol or converted directly into the less common ring fission compound methylcatechol.

The return of carbon locked within aromatic molecules to the environment depends upon the ability of microorganisms to cleave the chemically stable benzene ring. De-stabilisation of the ring is achieved by the enzymic addition of hydroxyl groups either singly, by the action of hydroxylases (mono-oxygenases or mixed function oxygenases) or by the addition of two hydroxyl groups together in a single step by a dioxygenase (dihydroxylase). Figure 1.2 shows two examples of such hydroxylation. For a compound to be acceptable as a ring fission substrate two hydroxyl groups must be present, either *ortho* (1,2-dihydroxy) as in catechol or *para* (1,4-dihydroxy) as in gentisate. Ring cleavage occurs via a dioxygenase which donates an atom of oxygen to each of the two

FIGURE 1.1: FATE OF RING SUBSTITUENTS PRIOR TO RING CLEAVAGE.

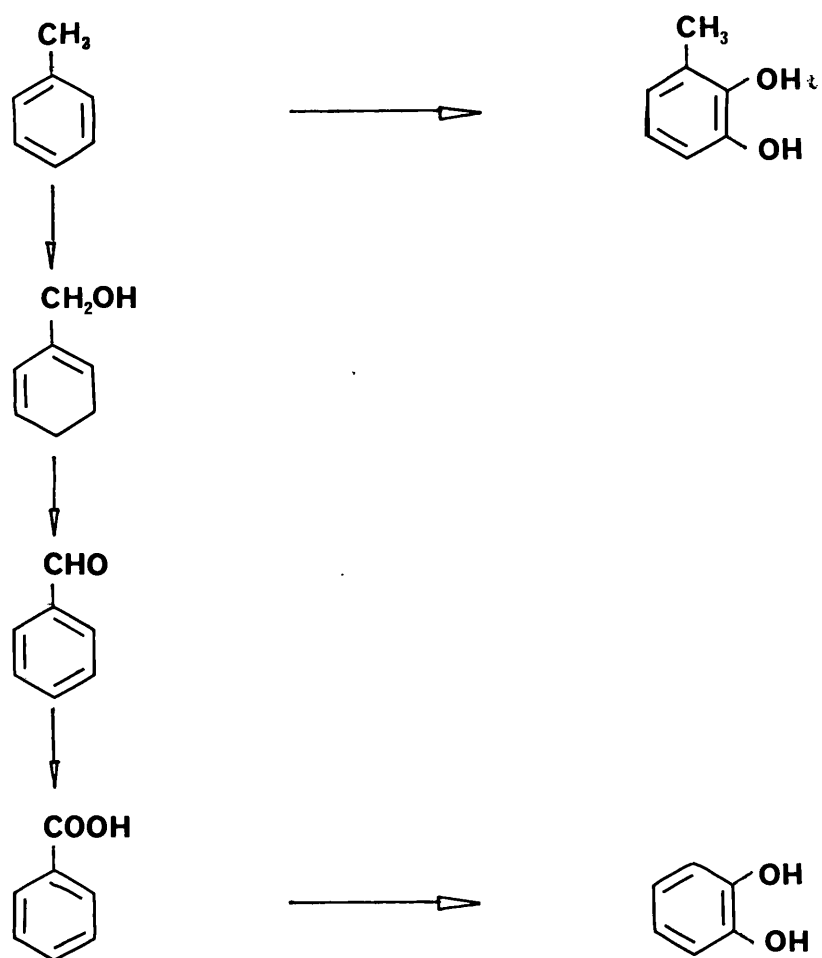
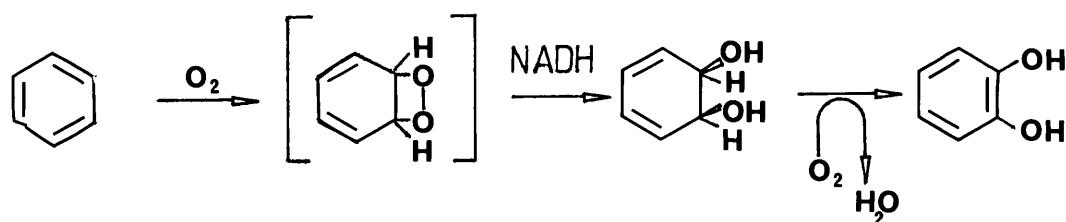
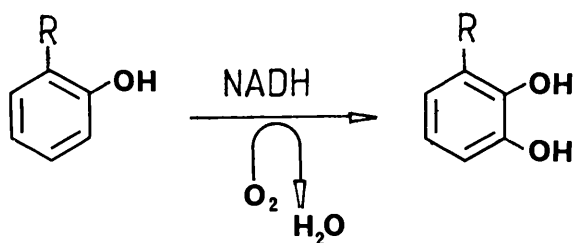


FIGURE 1.2: AROMATIC RING HYDROXYLATION REACTIONS.

a) Dihydroxylation of the benzene nucleus via 1,2 dioxetane and the *cis*-1,2-dihydrodiol.



b) Mono-hydroxylation of a phenolic compound.



carbon atoms on either side of the break. When the ring is *ortho*-dihydroxylated it can undergo two types of cleavage: a 1,2-dioxygenase can cleave the ring between adjacent hydroxyl groups (*ortho* cleavage) and a 2,3-dioxygenase can cleave the ring to one side of the two adjacent hydroxyls (*meta* cleavage). If the hydroxyl groups are present in a 1,4- (*para*) configuration the ring is cleaved between one of the hydroxyl groups and a side chain (fig.1.3). The opened structure can serve as a substrate for subsequent reactions in which it is converted to central metabolites.

Alternate reaction sequences are available for the subsequent reaction series in a *meta* cleavage pathway. Usually only one of the alternative pathways is available, however, sometimes the alternatives are encoded along side each other in the same set of genes. The TOL plasmid-specified *meta*-cleavage pathway for the oxidative catabolism of benzoate and toluates branches at the ring cleavage product of catechol and converges later at 2-oxopent-4-enoate or its corresponding substituted derivatives. The hydrolytic branch of the pathway involves the direct formation of 2-oxopent-4-enoate, with the release of formate, whereas the oxalocrotonate branch involves three enzymic steps effected by a dehydrogenase, an isomerase and a decarboxylase (fig 1.4). Only *m*-toluate is dissimilated via the hydrolytic branch and *p*-toluate and benzoate are catabolised almost exclusively by the oxalocrotonate branch. (Harayama *et al.*, 1987) The biochemical basis of the selectivity was found to reside in the inability of the dehydrogenase for ring fission products to attack the ring fission product derived from *m*-toluate.

Convergence of Aromatic Catabolic Pathways.

Although microorganisms possess a variety of different enzymes for the initial attack on diverse compounds, their catabolic pathways are usually convergent and channel substrates to a limited number of key intermediates, which are then metabolised further by central pathways. For example, a large proportion

carbon atoms on either side of the break. When the ring is *ortho*-dihydroxylated it can undergo two types of cleavage: a 1,2-dioxygenase can cleave the ring between adjacent hydroxyl groups (*ortho* cleavage) and a 2,3-dioxygenase can cleave the ring to one side of the two adjacent hydroxyls (*meta* cleavage). If the hydroxyl groups are present in a 1,4- (*para*) configuration the ring is cleaved between one of the hydroxyl groups and a side chain (fig.1.3). The opened structure can serve as a substrate for subsequent reactions in which it is converted to central metabolites.

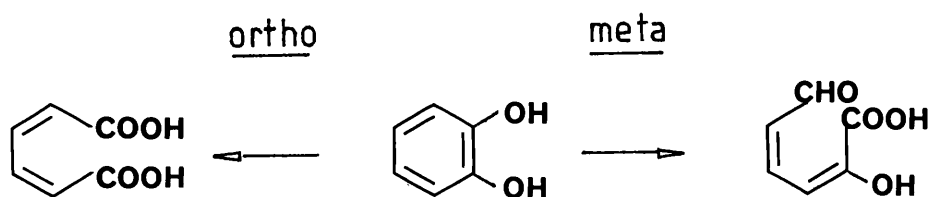
Alternate reaction sequences are available for the subsequent reaction series in a *meta* cleavage pathway. Usually only one of the alternative pathways is available, however, sometimes the alternatives are encoded along side each other in the same set of genes. The TOL plasmid-specified *meta*-cleavage pathway for the oxidative catabolism of benzoate and toluates branches at the ring cleavage product of catechol and converges later at 2-oxopent-4-enoate or its corresponding substituted derivatives. The hydrolytic branch of the pathway involves the direct formation of 2-oxopent-4-enoate, with the release of formate, whereas the oxalocrotonate branch involves three enzymic steps effected by a dehydrogenase, an isomerase and a decarboxylase (fig 1.4). Only *m*-toluate is dissimilated via the hydrolytic branch and *p*-toluate and benzoate are catabolised almost exclusively by the oxalocrotonate branch. (Harayama *et al.*, 1987) The biochemical basis of the selectivity was found to reside in the inability of the dehydrogenase for ring fission products to attack the ring fission product derived from *m*-toluate.

Convergence of Aromatic Catabolic Pathways.

Although microorganisms possess a variety of different enzymes for the initial attack on diverse compounds, their catabolic pathways are usually convergent and channel substrates to a limited number of key intermediates, which are then metabolised further by central pathways. For example, a large proportion

FIGURE 1.3: RING CLEAVAGE REACTIONS.

a) *Ortho* and *meta* fission of catechol by different dioxygenases.



b) *Meta* fission of gentisate.

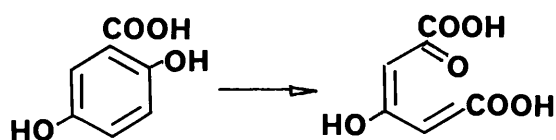
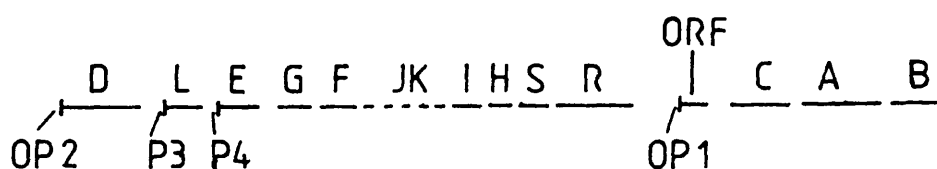
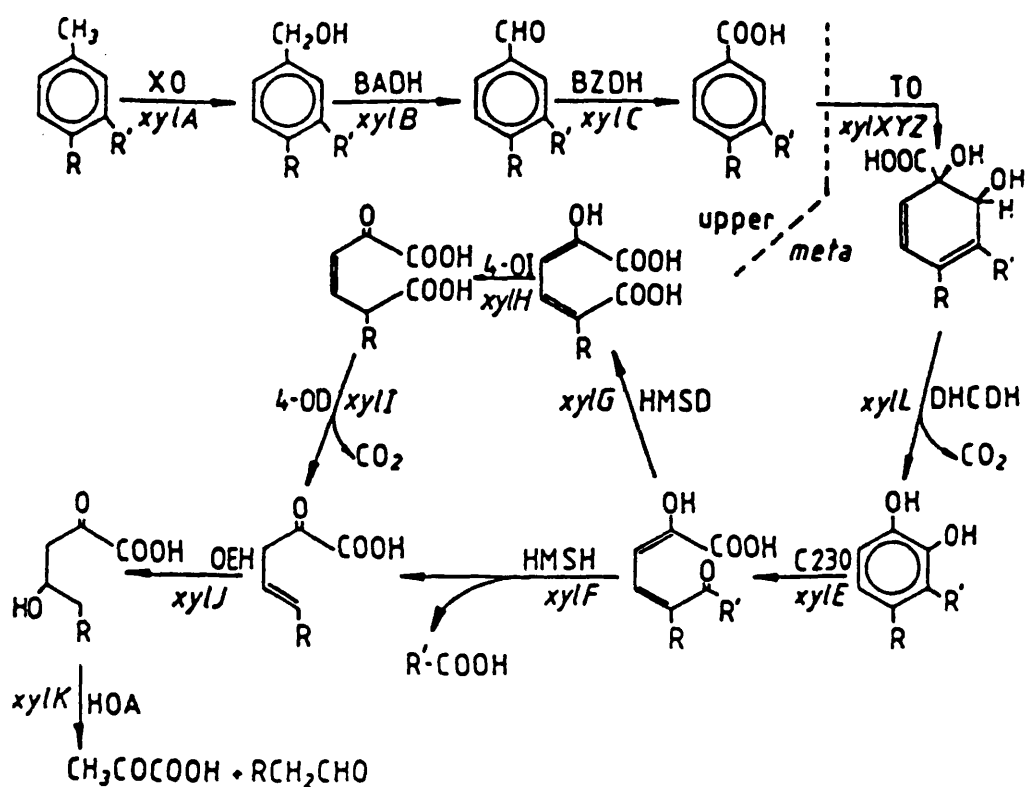


FIGURE 1.4: TOL plasmid specified pathway for the degradation of toluene and xylenes: Enzyme abbreviations: XO, xylene oxygenase; BADH, benzyl alcohol dehydrogenase; BZDH, benzaldehyde dehydrogenase; TO, toluate dioxygenase; DHCDH, dihydroxycyclohexadiene carboxylate dehydrogenase; C23O, catechol 2,3 dioxygenase; HMSH, hydroxymuconic semialdehyde hydrolase; HMSD, hydroxymuconic semialdehyde dehydrogenase; 4-OI, 4-oxalocrotonate isomerase; 4-OD, 4-oxalocrotonate decarboxylase; OEH, 2-oxopent-4-enoate hydratase; HOA, 2-oxo-4-hydroxy-pentolate aldolase. Initial compounds: R,R'=H, toluene; R=H, R'=CH₃, *m*-xylene; R=CH₃, R'=H, *p*-xylene. *xylA* to *xylZ*, genes encoding the pathway enzymes.

The organisation of TOL pathway genes is taken from Keil *et al.*, 1987.

FIGURE 1.4: *meta*-CLEAVAGE PATHWAY ENCODED BY TOL PLASMID pWWO.

(from Harayama *et al.*, 1987; Keil *et al.*, 1987).



0 5 kbp

of different aromatic compounds are converted to one of a few aromatic ring cleavage substrates, such as catechol, gentisate and protocatechuate or their derivatives (Ribbons and Eaton, 1982). The advantages of channelling diverse compounds into a few central pathways include a reduced genetic load and a simplification of regulatory circuits. The most extensively studied example of pathway convergence is the β -ketoadipate pathway (Stanier and Ornston, 1973). Various compounds converge on either catechol or protocatechuate for ring cleavage: following ring cleavage reactions the steps in each branch of the pathway are chemically similar and after β -ketoadipate-enol-lactone the intermediates are the same in both branches of the pathway (fig 1.5). The existence of such convergent pathways implies a high degree of pathway intergration throughout evolution.

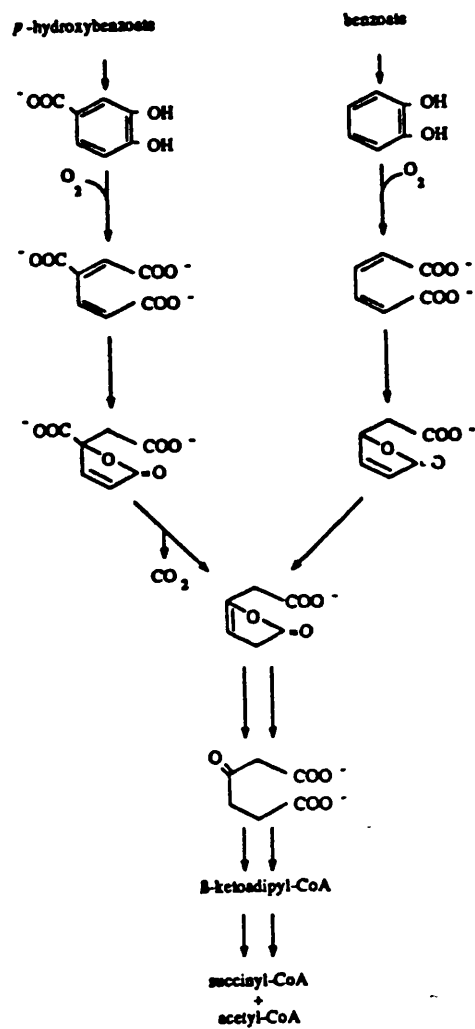
1.2: ORGANISATION OF CATABOLIC OPERONS.

The model for operons, proposed by Jacob and Monod, defined them as units of transcriptional activity and suggested that the repressor-operator mode of control might account for genetic regulation in general. Today we know that the control mechanism elaborated by the *lac* operon is only one of the many different ways in which bacterial genes are regulated. The concepts of a single promoter site, a single operator site and a single transcript, associated with operons, have been eroded with the understanding of a greater number of systems. Indeed, general principles in the organisation of genes within an operon, the number of operator and promoter regions and size of transcripts do not seem to exist (Cove, 1985).

It is quite common for genes encoding positive or negative regulatory elements to be unlinked to any other genes of the pathway, for example the gene encoding the Gal repressor, *galR*. In the case of the arabinose utilisation operon, the *araC* regulatory gene is spatially close to the *araBAD* gene cluster,

FIGURE 1.5: The β -ketoadipate pathway of *Acinetobacter calcoaceticus*: the two branches of the pathway converge at the intermediate β -ketoadipate enol-lactone.

FIGURE 1.5: THE CONVERGENCE OF CATABOLIC INTERMEDIATES IN THE β -KETOADIPATE PATHWAY.



but is transcribed separately in the opposite direction to the structural genes (fig.1.6). Indeed, the known arabinose transport and catabolic genes are distributed around the *E. coli* genome. Divergent promoters within operons appear to be a common form of gene organisation. A recent review by Beck and Warren (1988) cites over 60 regions of divergent transcription which occur in procaryotes and their viruses and in the nuclear genomes, mitochondria, chloroplasts and viruses of eucaryotes (over 20 have been found on the Chromosome of *E. coli* K-12 alone). The wide spread of such regions indicates that they are a general type of gene organisation rather than accidental rarities. Three types of promoter arrangement have been identified: back-to-back, with intervening DNA between the promoters; overlapping; and face to face. In the last case, the 5' termini for transcripts may not overlap. In the divergent arrangement, binding sites for regulatory proteins are often centrally located within the control region and proteins bound there may regulate transcription in both directions.

Genes which are transcribed in the same direction may also have separate promoter regions. The *InfC* gene, which codes for translation initiation factor 3 of *E. coli*, is situated just downstream of the termination codon for *thrS*, but is capable of being transcribed from a promoter region within the *thrS* coding sequence (Pramanik *et al.*, 1986).

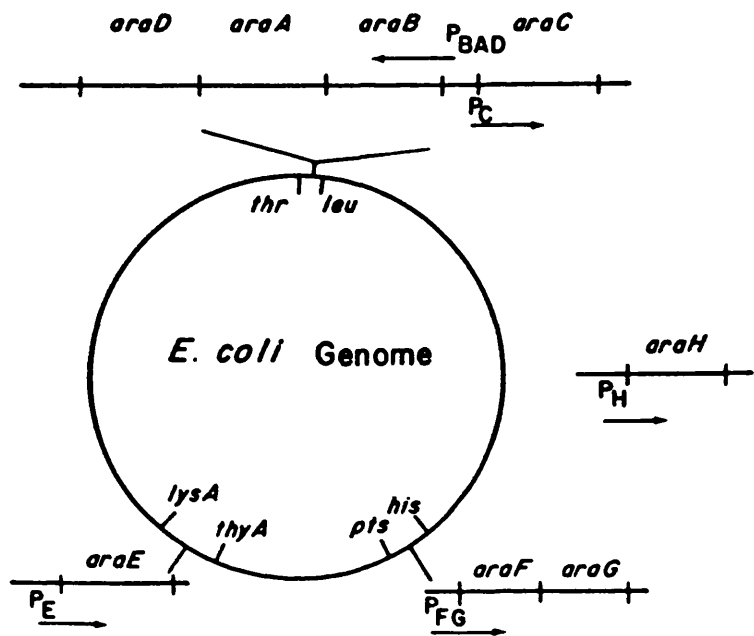
1.3 REGULATION OF CARBON CATABOLIC PATHWAYS.

Global Regulation.

"Enteric bacteria have evolved mechanisms that permit rapid growth under favourable conditions and foster survival when conditions are unfavourable for growth. The potential conflict in meeting these two requirements (simplicity and efficiency for rapid growth and complex protective devices for survival) may be what has led to the extraordinary ability of these cells to modulate

FIGURE 1.6: Location on the *E. coli* genetic map of the positions of the known arabinose transport and catabolic genes and their direction of transcription as indicated by the arrows.

FIGURE 1.6: LOCATIONS AND ORGANISATION OF KNOWN ARABINOSE TRANSPORT AND CATABOLIC GENES.



the expression of individual genes" (Neidhardt, 1987). The understanding of the control of separate but functionally related pathways, or regulons, (Grossman *et al.*, 1984) such as carbon utilisation, nitrogen utilisation and the heat shock response is becoming possible and is a step towards understanding global control of gene expression.

Perhaps the best understood regulon is the global carbon-energy (Cer) regulon. When growing in minimal media, enteric bacteria prefer glucose as the carbon source (judging from growth rates) and, irrespective of the carbon source present under such conditions, the enzymes for the catabolism of glucose are always present. The enzymes necessary for the utilisation of other carbon sources are generally present in very low concentrations. In addition the synthesis of enzymes capable of degrading other carbon-containing compounds is usually reduced in the presence of glucose (the glucose effect). The glucose effect has at least three components. One of these is inducer exclusion whereby glucose prevents the entry of inducers by inactivation of permeases by an intermediate phosphate carrier in the system responsible for the uptake and phosphorylation of glucose. The other components are transient repression, causing severe inhibition of enzyme synthesis for about one generation and catabolite repression. These phenomena have been reviewed by Postma, (1986).

Analysis of catabolite sensitive operons has shown that they contain promoter regions which require the binding of cAMP-CRP complex for the initiation of transcription (Ullmann and Danchin, 1983). The cAMP-CRP regulon is a global regulon consisting of operons which are regulated at the transcriptional level, in part at least, by the intracellular concentration of cAMP and is a component of the larger Cer regulon. The cAMP-CRP complex binds to a number of catabolic operons including *gal*, *ara*, *mal* and *lac* upstream of the RNA polymerase binding site in the promoter region and is essential for the initiation of transcription of these promoters. In the case of *lac* the binding of cAMP-CRP increases the binding constant of RNA polymerase (Malan *et al.*, 1984). As such

the binding of the cAMP-CRP complex to the promoters of catabolic operons acts as a positive activating signal for the potential of using carbon sources other than glucose in the media. In the absence of glucose and in the presence of specific inducers, transcription of individual operons occurs so that carbon and energy sources other than glucose can be utilised.

Regulation of Transcription.

The level of protein products of individual catabolic operons can be influenced by a variety of different regulatory events: one or more of which may affect any particular operon. The first possible step in the production of catabolic enzymes which can be regulated is the initiation of transcription of the genes which encode them. This is the most economical point at which to exert gross control of a pathway and has proved to be the one of which most examples have been found. Transcription initiation depends upon the correct alignment of RNA polymerase and a promoter sequence on the DNA and anything which perturbs this interaction can control the expression of the genes. Positive regulatory proteins enhance the binding of RNA polymerase to promoters whilst negative regulatory proteins reduce this binding. It is clear that operons controlled by positive regulatory elements are not transcribed in the absence of those proteins. An interesting example of this is the heat shock genes. The *rpoH* (also *htpR*) gene product is a sigma factor (Grossman et al., 1984) specific to heat shock gene promoters which combines with core RNA polymerase, to initiate transcription of operons in the heat shock regulon after stimulation by elevated temperature or some toxins, for example ethanol.

Elongation of RNA transcripts is another possible place for regulation: pausing of RNA polymerase on the DNA template has been suggested as a possible mechanism to synchronise transcription and translation (Winkler and Yanofsky, 1981) specifically in the case of the *trp* attenuator sequence where coordination is required for efficient function of the system. The *trp* operon

is an example of regulation by transcription termination (or attenuation), where the premature termination of transcription within an operon modulates the expression of downstream genes. (Review: Yanofsky, 1981).

Attenuation in the *trp* operon is relieved under conditions of low tryptophan concentrations, however, other attenuated systems are relieved by the action of protein factors. Such processes are termed antitermination and these work by either modifying the transcribing polymerase or the termination signal. The bacteriophage lambda (λ) utilises two antitermination proteins N and Q, to regulate early and late gene expression. Transcription which initiates at the major promoters P_L and P_R terminates at the ρ -dependent sites t_{L1} and t_{R1} respectively. The *N* gene is synthesised on the leftward transcript and once synthesised the N protein modifies RNA polymerase as it transcribes a specific 17bp region, the *nut* site. The modified RNA polymerase fails to recognise the termination signal and the remainder of the DNA can be transcribed.

A possible regulatory role for termination at the end of operons is less clear. However, the fact that tandem terminators have been identified at the end of the *trp* operon (Holmes *et al.*, 1983), each influenced by different factors, suggests that termination may be influenced by physiological control.

Regulation by RNA Processing.

Clearly, the amount of a particular mRNA species in a cell will depend upon the rate of degradation as well as the rate of synthesis: the stability of mRNA can be influenced by the secondary structure of the molecule, any interactions it may have with protein factors and ribosomes and cleavage by specific endonucleases.

Repetitive extragenic palindromic (REP) sequences are large inverted repeats first identified in the intergenic regions of a number of bacterial operons by Higgins *et al.* (1982). It seems likely that REP sequences act at the level of RNA as all of those identified are located in transcribed regions either

intergenically or in the 3' untranslated region of an operon but proximal to the terminator (Stern *et al.*, 1984). The role of REP sequences is, as yet, unclear: deletion of the REP sequence from the *hisJ-hisQ* intergenic region has little effect on upstream *hisJ* expression (Robinson and Higgins, cited in Higgins and Smith, 1985). REP sequences are also conspicuously absent from stable mRNA species. Whatever the true role of REP sequences, with such a high degree of conservation and spread (REP sequences are present in about 25% of all transcripts) they must play an important role in the cell.

Conflicting reports are present in the literature concerning the effect of ribosome binding on the stability of mRNA. Yates and Nomura, (1981) showed that in the absence of ribosomes the rate of mRNA degradation was increased. However, Har-El *et al.* (1979) reported that the presence of ribosomes can decrease the half life of mRNA. It has been suggested that the nuclease activity is associated with the ribosomes themselves, so that under certain conditions the presence of ribosomes enhances degradation.

Translational Regulation.

Translational efficiency depends upon the efficiency with which ribosomes bind to the mRNA and initiate translation and the rate at which translation proceeds. The ribosome binding site (Shine-Dalgarno sequence; the signal sequence which directs the ribosome to bind to the mRNA and initiate translation at the correct AUG codon) is a short purine rich sequence complementary to the 3' end of the 16S rRNA, which is located 5-10 bp upstream of the initiation codon. However, other signals, not yet understood, probably play a part as some sequences complementary to the 3' end of 16S rRNA do not act as sites for translation initiation. Genes expressed in a 1:1 ratio are often separated by very short intergenic regions so that the genes can be translationally coupled as has been shown to be the case with the *galT* and *galK* genes (Schumperli *et al.*, 1982). In operons where the genes are expressed

at different levels translational coupling must be relatively unimportant and expression of the various genes depends on the efficiencies of the individual ribosome binding sites.

Following the initiation of translation the rate at which a protein is synthesised may depend on codon usage. Highly expressed genes contain very few codons which are recognised by minor isoaccepting species of tRNA while poorly expressed genes tend to have a higher proportion of these rare codons (Konisberg and Godson, 1983). However the significance of codon usage in determining translational efficiency is yet to be demonstrated.

1.4: EVOLUTION OF METABOLIC PATHWAYS .

The metabolic capabilities of microorganisms have long been used as a basis for taxonomic schemes and such studies have also included the ability to degrade aromatic-containing compounds (Stanier *et al.*, 1966, for example). However as more exacting techniques, such as protein and DNA homologies, become available for the classification of microorganisms it is likely that the study of metabolic pathways will be more useful in the proposal and testing of theories concerning the evolution of novel enzymic steps and pathways.

One of the earliest models for the evolutionary growth of a metabolic pathway was that proposed by Horowitz (1945), which stated that pathways developed by a process of "retrograde evolution." The hypothesis suggested that the first primitive organism could find all of the nutrients it required from the environment. Upon depletion of a particular nutrient, a variant of the organism developed which possessed an enzyme that functioned to convert a chemically related compound to the required compound. Lewis, (1951), proposed a two step mechanism by which this new catabolic step could arise: new genetic material being created by duplication of an existing gene followed by "mutation to new functions" of one of the genes while the first retained its original function.

It was also suggested that functionally related but not tightly clustered genes also arose through gene duplication and subsequent dispersion.

Dagley, (1975), argued against this theory as it would require the new growth substrate to be stable and present in sufficient quantities to sustain growth and would also probably require the new enzyme to have a similar reaction mechanism to the original one. Analysis of consecutive enzymes in pathways provides conflicting evidence: the peptide maps of tryptophan synthase A and B from *E. coli* show no homology (Hegeman and Rosenberg, 1970), but comparison of the complete nucleotide sequence of the MetB- and MetC- proteins of methionine biosynthesis showed significant homology (Belfaiza *et al.*, 1986), 36% of the residues being identical or conservative substitutions. The *metB* (coding for cystathionine γ -synthase) and *metC* (coding for β -cystathionase) genes are located at 88 and 65 minutes respectively on the map of the *E. coli* chromosome and although each enzyme has evolved a specialised role the cystathionine γ -synthase is also capable of β -cystathionase activity at a very low frequency. Belfaiza *et al.* (1986) suggest that this vestigial activity may be evidence in support of the tandem duplication theory of evolution proposed by Horowitz (1945, 1965). New genetic material may also be produced by chromosome duplication as proposed by Riley and Anilionis (1978). They hypothesised that the *E. coli* chromosome had undergone two duplication events and that the additional DNA was available to mutate to produce potentially novel enzymic activities.

The recruitment of enzymes with required catalytic properties from other pathways has been demonstrated by Campbell *et al.* (1973) and may offer an explanation to why the genes for some pathways are scattered around the chromosome. They deleted the β -galactosidase (*lacZ*) gene from a strain of *E. coli* and then selected mutants capable of growing on lactose. The gene which restored growth (*ebg*) encoded a new β -galactosidase activity and mapped at minute 65 compared to the *lac* operon which normally maps at minute 8.

Investigation of proteins of the β -ketoadipate pathway by Ornston's group led them to suggest that regions of sequence similarity shared by enzymes of the pathway were acquired subsequent to their divergence from any common ancestor (Yeh and Ornston, 1980). The proposed mechanism for this was by mutations causing the transfer of relatively short regions of sequence information from one set of genes to another.

Transfer of DNA between Species.

The ability of DNA to be transferred between bacterial cells either of the same species or of widely different genera, by transformation, transduction or episome-mediated conjugation, provides another mechanism for the evolution of genetic material within microorganisms.

If the genes of a pathway are clustered together then the transfer of intact pathways becomes easier. If certain groups of genes are only occasionally required they could logically become incorporated into plasmids and reside in only a proportion of the population, which could act as a bank, reducing the genetic load of the population. When certain genes are required more frequently they could become integrated into the chromosome, providing a stimulus for the lateral movement and stable inheritance of genes within a mixed microbial population. The movement of sets of genes between plasmids and the chromosome can be illustrated by the recent discovery of a plasmid encoded example of the β -ketoadipate pathway harboured by *Acinetobacter calcoaceticus* (Winstanley *et al.*, 1987). This is the first report of a catabolic plasmid encoding genes for the dissimilation of catechol by the β -ketoadipate route: the genes on the plasmid appear to be related to those on the chromosome of *A. calcoaceticus*. However, significant differences are present between the two sets of genes in that the *catA* gene is closely linked to the *catBCEFD* operon on the chromosomal version but separated from them by at least 10kbp on the plasmid encoded version, suggesting that the transfer of genes from chromosome to plasmid

occurred in more than one step or the genes have been subjected to a translocation event on the plasmid. Interestingly, the ring cleavage enzyme encoded on the plasmid was found to be significantly more heat-labile than the chromosomally encoded version (Winstanley *et al.*, 1987). If the two enzymes shared a common ancestral gene it is possible that mutations could have occurred altering one of the genes when chromosomal and plasmid copies were present in different members of the same population.

In contrast to the β -ketoadipate pathway, the genes for the growth on toluene, *m*-xylene, *p*-xylene, and their corresponding alcohols, aldehydes and acids are most commonly found on large plasmids-the TOL plasmids (Williams and Murray, 1974; Wong and Dunn, 1974). Sinclair *et al.* (1986) showed that a soil isolate, *Pseudomonas putida* MW1000 contained a 56kbp fragment of DNA on its chromosome which was related to a region of the archaetypal TOL plasmid, pWWO, carrying all of the catabolic genes of the TOL pathway. Further they demonstrated that the fragment could move from the chromosome into a plasmid, supporting the theory of exchange of DNA between the chromosome and extrachromosomal elements.

Conservation of Pathways Between Species.

Acinetobacter and *Pseudomonas* are widely divergent bacterial genera, yet studies have revealed common ancestries for isofunctional enzymes of the β -ketoadipate pathway of the two genera. For example the muconolactone isomerase enzymes of the two organisms possess approximately 50% homologous residues over the first 46 amino-terminal amino acids (Yeh *et al.*, 1978). Similarly the amino-terminal amino acids sequence of the γ -carboxymuconolactone decarboxylase enzymes of the two organisms are 50% identical in the first 36 positions. (Yeh *et al.*, 1980). However, the *Acinetobacter* decarboxylase is induced by protocatechuate whereas the *Pseudomonas* decarboxylase is induced by β -ketoadipate. The β -ketoadipate pathway in the two genera has evolved from a

common ancestor such that some parts of the pathway are now controlled in the same way and some in different ways. The individual genes of the pathway have also diverged from their ancestral form as indicated by the amino-terminal amino acid sequences of their protein products (Yeh *et al.*, 1978; Yeh *et al.*, 1980 for example)

Although aromatic catabolic pathways of *Escherichia coli* and *Klebsiella pneumoniae* have not yet been studied in great detail, other catabolic pathways have been investigated. Buvinger and Riley (1985a, b) have cloned the *lac* genes of *K. pneumoniae* and compared them with those of *E. coli*. They found that the *lacZ*, *lacY* and *lacI* genes were 61%, 67% and 49% conserved respectively between the two species, but that the direction of transcription of the *lacI* gene of *K. pneumoniae* was inverted compared to that of *E. coli*. On more detailed analysis of the protein sequence, deduced from the DNA sequence, they found that conservation of residues was in regions of known structural or catalytic importance and that divergence had occurred in non-essential areas. An example of this is the first 26 residues of β -galactosidase, which can be removed without altering the enzyme activity (Muller-Hill and Kania, 1974); analysis of the enzyme from the two organisms showed no conservation of sequence in the first 16 residues (Buvinger and Riley, 1985b).

1.5: GENETIC ANALYSIS OF CARBON CATABOLIC PATHWAYS.

Most studies concerning the molecular genetics of aromatic catabolic pathways have been carried out on naturally occurring plasmids as the genes of interest are present on small, manipulable fragments of DNA. The most widely studied of these is the TOL plasmid, capable of degrading toluene, *m*-xylene, *p*-xylene and their derivatives via *meta*-cleavage of catechol. The techniques used in the elucidation of the organisation and regulation of the genes of the TOL plasmid will be described in some detail as they provide a model for the general

investigation of catabolic pathways.

Pseudomonads have been shown to degrade benzoate via catechol in the *ortho* fission pathway (Stanier, 1947), whereas, they catabolise phenol and methyl substituted benzene via benzoate and *meta*-cleavage of catechol, indicating the presence of two pathways for benzoate metabolism. Spontaneous mutants of *Pseudomonas arvilla* isolated by Nazakawa and Yokoto (1973) had lost the *meta*-cleavage enzymes but retained the ability to grow on benzoate using the *ortho*-cleavage route, suggesting that the *meta*-cleavage enzymes were plasmid encoded. Similar mutants were obtained by Williams and Murray (1974) who demonstrated that the *meta*-cleavage pathway was plasmid encoded by mating the plasmid into Tol⁻ mutants thereby restoring the ability of the mutants to grow on toluate and benzoate.

An insight into the regulation of the pathway was shown by the induction of the first series of enzymes, which catalyse the conversion of hydrocarbons to carboxylic acids, by toluene and *m*-/*p*- xylene (Worsey *et al.*, 1978). A regulatory mutant was isolated in which *m*-xylene and *m*-benzyl alcohol failed to induce the toluate catabolic enzymes. Toluene however, was still able to induce the enzymes necessary for its own degradation, indicating that the genes were present in two gene blocks, each controlled by their own regulatory gene.

Restriction maps of the TOL-plasmid (pWWO) were determined (Downing and Broda, 1979) and the positions of two gene clusters determined by cloning and transposon mutagenesis (Franklin *et al.*, 1981). The two gene clusters were separated by 14kbp; one was for the upper pathway (the conversion of a hydrocarbon to its corresponding carboxylic acid) and the other for the lower pathway (carboxylic acids to tricarboxylic acid cycle intermediates). The relative positions of the *xyIDGEF* genes coding for the lower pathway enzymes were determined by Inouye *et al.* (1981a, b) and these genes were shown to be non-inducible in the absence of the *xyIS* gene product, indicating positive

regulation. A precise map of the lower pathway genes (Harayama *et al.*, 1984) showed the gene order to be *xyIDLEGFJIH* and *xyIS*; the order of the genes is in general accord with the order in which the gene products act in the reaction sequence. The order of the upper pathway genes was determined by Harayama *et al.* (1986) using transposon mutagenesis and was found to be *xyICAB*, which is different from the order in which their gene products act in the reaction sequence.

A refined model for the regulation of the toluene catabolic pathway, based on complementation analyses, has been described (Inouye *et al.*, 1986). Benzoate (or toluate) is the inducer for the positive regulator gene *xyIS* the product of which, XylS, acts on the *xyIDEGF* operon. The product of the *xyIR* gene (induced by toluene, *m*-/*p*- xylene) was initially thought to act in a positive manner directly on both the *xyICAB* and *xyIDEGF* operons. However, XylR alone has been shown to be insufficient for the stimulation of expression of the lower pathway genes. Stimulation of expression of the lower pathway genes by the *xyIR* gene product occurs, however, in the presence of a functional *xyIS* gene, implying some interaction of these regulatory factors. A diagram of the organisation of the TOL pathway genes is given in figure 1.4.

Chromosomally Encoded Catabolic Pathways.

The detailed molecular analysis of chromosomally encoded pathway relies on either the genes of the pathway being closely linked, or a series of cloning experiments followed by *in vitro* construction of a vector containing all of the genes of the pathway. The first report of chromosomally-encoded aromatic catabolic genes being cloned for detailed molecular analysis was that of the *catBCDE* gene cluster, coding for four genes of the catechol branch of the β -ketoadipate pathway from *Acinetobacter calcoaceticus* (Shanley *et al.*, 1986). This report by Ornston and co-workers described the cloning of a part of the

pathway, which was subsequently used to probe a gene library for overlapping fragments of DNA. Four positive clones were identified, all of which contained the adjacent fragment of DNA which carried the *catA* gene (Neidle and Ornston, 1986). The same group have gone on to clone genes from the protocatechuate (*pca*) branch of the β -ketoadipate pathway (Doten *et al.*, 1987) and investigate their organisation. The fragment containing the *pca* genes hybridised strongly to the previously cloned *cat* genes, probably because the isofunctional *catDEF* and *pcaDEF* gene products show homology at the protein level (Yeh *et al.*, 1980). Cloning and analysis of further related genes by this group, the genes necessary for the conversion of benzoate to catechol (*benABCD*), has shown that they are clustered on the *A. calcoaceticus* chromosome with the independently regulated *cat* and *pca* genes for the dissimilation of catechol (Neidle *et al.*, 1987). In a 16kbp region of the chromosome there are 10 genes for benzoate catabolism in no fewer than three transcriptional units. This type of arrangement is termed supraoperonic clustering (fig. 1.7).

1.6: THE HYDROXYPHENYLACETATE AND HOMOPROTOCATECHUATE PATHWAY.

The pathway for the catabolism of 4-hydroxyphenylacetate (HPA) was first described by Spornins *et al.* (1974) for a species of *Acinetobacter* and two strains of *Pseudomonas*. All of these converted 4-HPA into carbon dioxide, pyruvate and succinate via the ring fission compound homoprotocatechuate (HPC). However, when 3-HPA served as the carbon source, homogentisate was the ring fission compound used by the *Acinetobacter* species. Adachi *et al.* (1964) identified HPC as an intermediate in the degradation of HPA by *Pseudomonas ovalis*: they also showed that the product of HPC ring fission was 5-carboxymethyl-2-hydroxymuconate (CHMS). This was confirmed by Spornins *et al.* (1974) who identified HPC, CHMS, 2-hydroxyhepta-2,4-diene-1,7-dioate (HHDD) and 4-hydroxy-2-ketopimelic acid (HKP) as intermediates in the conversion of

HPA to central metabolites. The sequence of enzyme-catalysed reactions proposed by Spornins was as follows: ring-fission by a 2,3-dioxygenase; nicotinamide adenine dinucleotide-dependent dehydrogenation; decarboxylation; hydration; aldol fission and oxidation of succinic semialdehyde (SSA). Dagley, (1975) proposed feasible mechanisms for the reaction sequence (fig 1.8) which suggested the presence of three possible electron rearrangement steps. Analysis of cloned *hpc* genes (Jenkins, 1987) showed that the tautomerisation step (g) was not necessary: aldol fission of the intermediate 2,4-dihydroxy-hepta-2,ene-1,7-dioate (HHED) producing pyruvate and succinic semialdehyde. The structure of HHED was confirmed by NMR analysis of the isolated intermediate to be the enol-form of HKP. Further HKP was chemically synthesised and was shown to have a similar λ -max (260nm) as the biologically produced HHED (Jenkins, 1987). The chemically synthesised HKP served as a substrate for the aldolase and from this evidence Jenkins suggested that the compound called HKP was infact HHED. A representation of the pathway for the conversion of HPA to SSA and pyruvate is given in figure 1.9 .

A similar pathway to the one found in *Acinetobacter* and *Pseudomonas* for the conversion of 4-HPA to central metabolites is also present in *E. coli* C when growing on 4-HPA and 3-HPA (Cooper and Skinner, 1980). That *E. coli* can degrade aromatic compounds related to the aromatic amino acids phenylalanine and tyrosine should not be surprising as it may meet such compounds in its intestinal-faecal environment (Spoelstra, 1978). The genes required for the degradation of HPC to SSA and pyruvate were cloned from *E. coli* C by Jenkins (1987).

Regulation of the HPA Pathway.

Growth of *P. ovalis* (Adachi et al., 1964), *P. putida* (Barbour and Bayly, 1977) and *E. coli* C (Skinner, 1981) on HPA induced the HPA hydroxylase and the HPC catabolic enzymes whereas growth on HPC induces the HPC catabolic enzymes but

FIGURE 1.8: The possible intermediates in the catabolism of HPC as proposed by Dagley, (1975).

- I) HPC, Homoprotocatechuate
 - a) dioxygenase
- II) CHMS, 5-carboxymethyl-2-hydroxymuconate semialdehyde
 - b) dehydrogenase
- III) CHM, 5-carboxymethyl-2-hydroxymuconate
 - c) isomerase
- IV) COHED, 5-carboxymethyl-2-oxo-hex-3-ene-1,6-dioate
 - d) decarboxylase
- V) HHDD, 2-hydroxyhepta-2,4-diene-1,7-dioate
 - e) isomerase
- VI) OHED, 2-oxo-hepta-3-ene-1,7-dioate
 - f) hydratase
- VII) HHED, 2,4-dihydroxy-hepta-2-ene-1,7-dioate
 - g) isomerase
- VIII) HKP, 4-hydroxy-2-ketopimilate
 - h) aldolase

FIGURE 1.8: FEASIBLE MECHANISMS INVOLVED IN THE CATABOLISM OF HPC.

(Dagley, 1975)

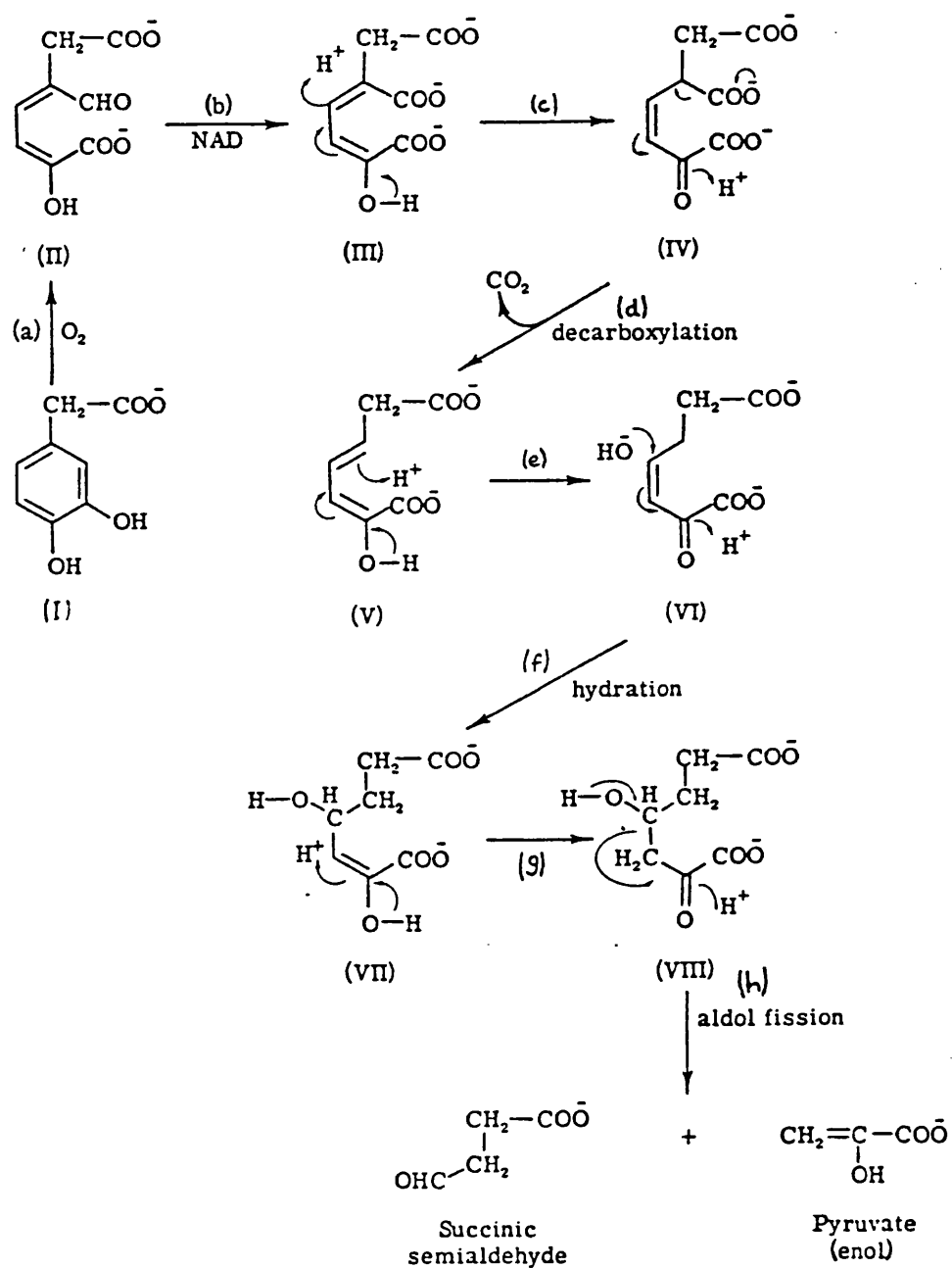
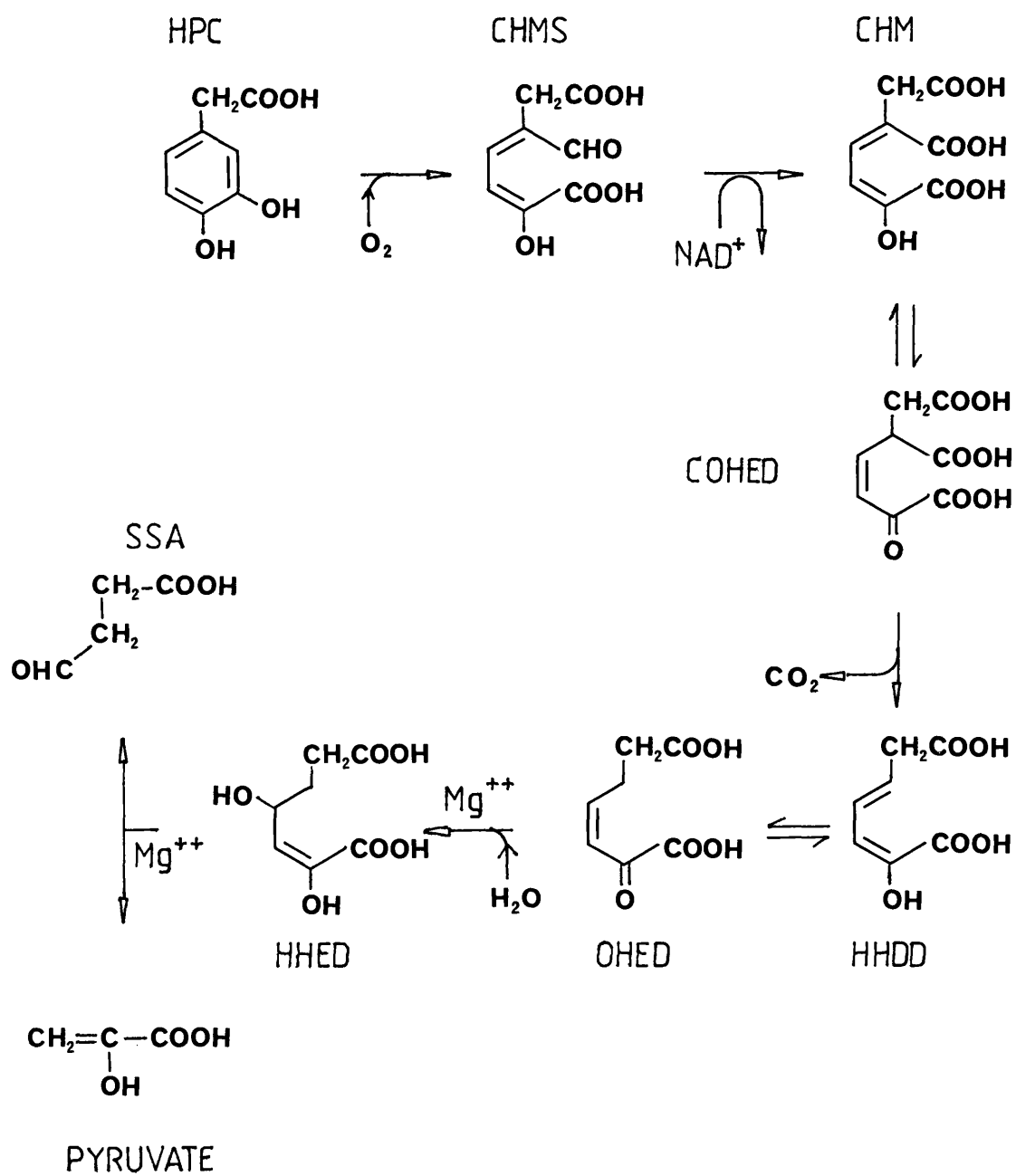


FIGURE 1.9: Intermediates in the catabolism of HPC:

HPC, homoprotocatechuate; CHMS, 5-carboxymethyl-2-hydroxymuconate semialdehyde; CHM, 5-carboxymethyl-2-hydroxymuconate; COHED, 5-carboxymethyl-2-oxo-hex-3-ene-1,6-dioate; HHDD, 2-hydroxyhepta-2,4-diene-1,7-dioate; OHED, 2-oxo-hepta-3-ene-1,7-dioate; HHED, 2,4-dihydroxy-hepta-2-ene-1,7-dioate; SSA, succinic semialdehyde.

FIGURE 1.9: PATHWAY FOR THE DEGRADATION OF HPC IN *E. coli* C.

(Jenkins, 1987).



not the HPA hydroxylase. These data suggest that the genes which encode these catabolic enzymes are divided into at least two regulatory groups. Mutants defective in a single enzyme, the CHMS dehydrogenase, grown in the presence of HPC were shown to contain all of the enzyme activities except the CHMS dehydrogenase indicating that HPC induced all of the enzymes responsible for its own catabolism (Barbour and Bayly, 1978; Skinner, 1981). The *hpa* and *hpc* genes of *E. coli* C were shown by Skinner (1981) to be 90% co-transducible in P1 transduction experiments, suggesting that the genes were closely linked but probably not contiguous. The set of genes responsible for the complete mineralisation of HPA can be considered in terms of three groups of reactions, namely:

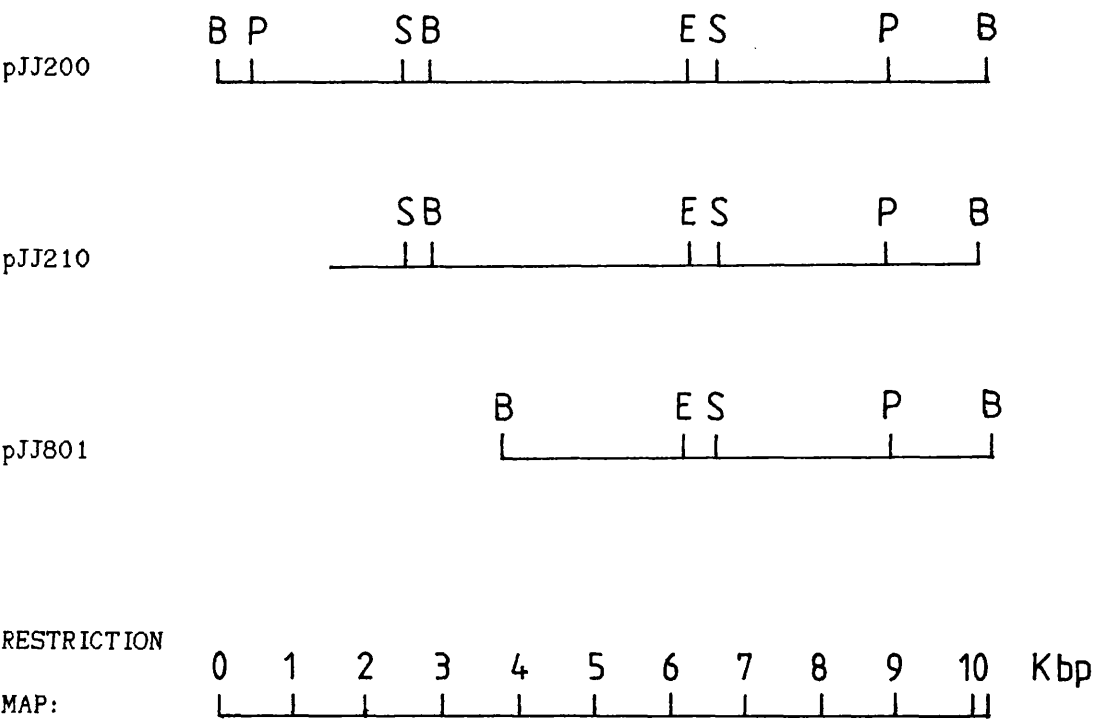
Group 1: 3-HPA and 4-HPA are taken up and hydroxylated adjacent to an existing hydroxyl group to produce HPC. These enzymes are induced by HPA and their respective genes have been designated *hpaA* (permease), *hpaB* (hydroxylase) and *hpaR* (regulatory protein).

Group 2: HPC is converted through a series of reactions to give the intermediates SSA and pyruvate. This group of genes are inducible by both HPA and HPC: as HPC can be transported into the cell and induces the genes for its own degradation this group represents a complete pathway.

Group 3: The conversion of SSA to succinate is achieved by the induction of an NAD-linked SSA dehydrogenase (encoded by the *sad* gene): induction is by SSA (Donnelly and Cooper, 1981).

From a study of mutants of the HPA pathway Skinner (1981) proposed that the product of a single regulatory gene controlled both the hydroxyphenylacetate (*hpa*) and the homoprotocatechuate (*hpc*) genes. A study of the cloned *hpc* genes of *E. coli* C led Jenkins (1987) to propose that control of expression of these genes was by a system of negative regulation.

FIGURE 1.10: PRIMARY CLONES ISOLATED BY JENKINS CONTAINING *hpc* GENES.

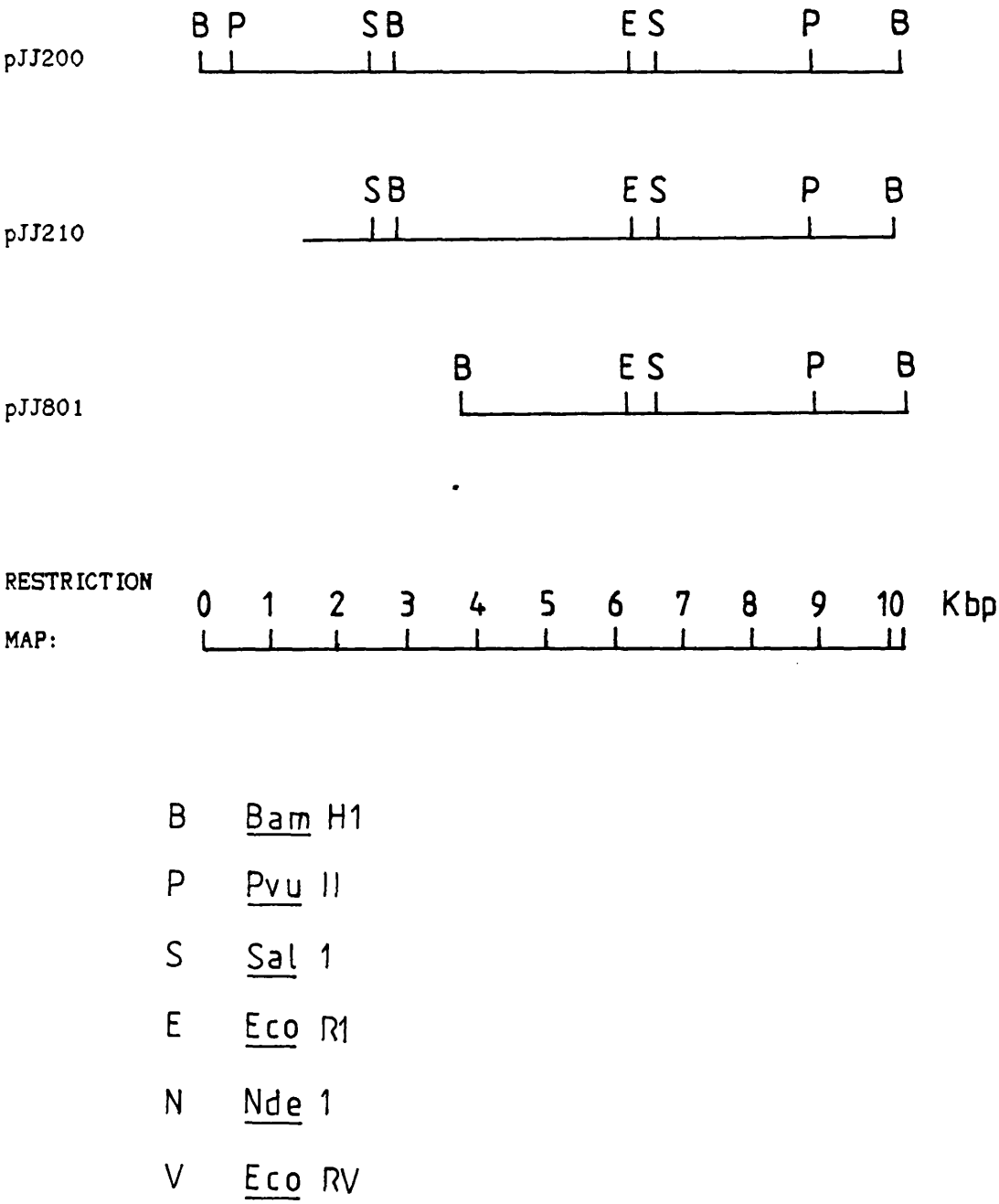


Analysis of Cloned *hpc* Genes.

The proteins encoded by the *hpc* genes are not detectable in strains of *E. coli* K-12 (Skinner, 1981); further hybridisation analysis of K-12 DNA with cloned *E. coli* C *hpc* genes (Jenkins and Cooper, 1988) or synthetic oligonucleotide probes to the *hpcG* gene (Ferrer and Cooper, 1988) and *hpcC* gene (Fawcett *et al.*, in press) show no significant hybridisation, supporting the idea that *E. coli* K-12 strains do not contain any *hpc* genes. This makes K-12 strains ideal as a genetic negative background for the analysis of cloned sequences containing these genes: the chosen K-12 strain was *E. coli* 5K, which is a *hsdR* derivative, allowing efficient transformation of DNA from the naturally *hsdM* *E. coli* C.

During analysis of the cloned *hpc* genes, Jenkins utilised both *E. coli* C and *E. coli* K-12 strains to produce a model for the organisation and expression of those genes (Jenkins, 1987). As the results of his experiments have direct relevance to this work the following section will describe them in some detail. Jenkins isolated three independent genomic clones containing genes of the *hpc* pathway by complementation of *E. coli* C strains which contained mutations in that sequence of genes. The chromosomal DNA in each case was partially digested using *Sau*3A and ligated into the *Bam*H1 site of either pBR328 or pBR322. This procedure resulted in one of the clones, pJJ801, having a terminal *Bam*H1 site which is not present in the chromosomal copy and a second clone, pJJ210, having one insert-vector junction as a *Sau*3A site only. The third clone, pJJ200, contains two *Bam*H1 fragments and is thought to represent the chromosomal distribution of these sites (fig. 1.10). The three clones have very similar restriction maps and from Southern blot analysis were shown to be from the same region of chromosomal DNA. The recombinant plasmids pJJ200 and pJJ210 contain all of the genes necessary for the conversion of HPC to succinic semialdehyde and pyruvate. These clones also contain a regulatory gene which means that the expression of their genes is controlled in *E. coli* 5K. The

FIGURE 1.10: PRIMARY CLONES ISOLATED BY JENKINS CONTAINING *hpc* GENES.

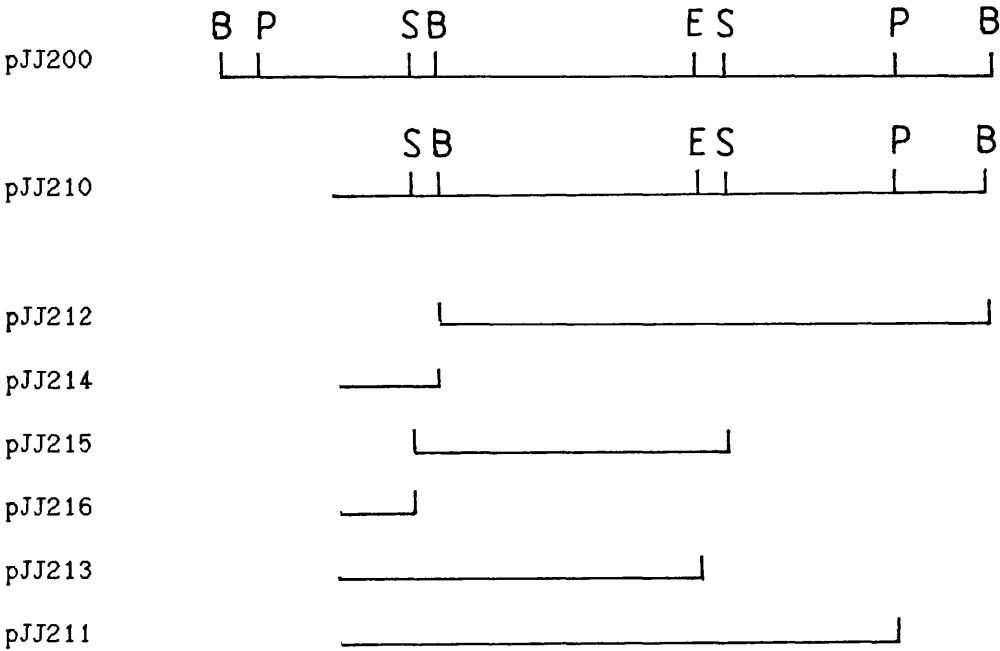


expression of the five genes present on pJJ801 is constitutive in *E. coli* 5K, but regulated in *E. coli* C strains, suggesting that pJJ801 contains an operator region to which a host-encoded regulatory factor can bind in trans. These experiments led Jenkins to propose a negative regulatory model and further evidence was obtained to implicate the role of cAMP-CRP as a positive element in the regulatory system. The HPC catabolic enzymes of both *E. coli* C and *E. coli* 5K (pJJ210) were shown to be subject to glucose repression which could be partially relieved by the addition of 5mM cAMP to the growth medium.

Oxidation experiments showed that whole cells of 5K (pJJ200/pJJ210) when grown on glycerol in the presence of HPA were not able to oxidise HPA or HPC even though cell extracts from these strains contained all of the necessary enzyme activities for the conversion of HPC to succinic semialdehyde and pyruvate. This data and the report that *E. coli* K-12 has the NAD-linked (SSA inducible) succinic semialdehyde dehydrogenase (Skinner and Cooper, 1982) led to the conclusion that the non-growth of 5K (pJJ200/pJJ210) on HPC was due to the lack of a permease. However, some HPA or HPC must enter 5K because the presence of these compounds in the medium significantly increased the levels of HPC catabolic enzymes.

The restriction sites present in the primary clones were used to construct subclones so that the sequence of genes could be distinguished and a summary of these is presented in figure 1.11. Using these subclones *hpcRGHBC* and *hpcF* were assigned positions but due to the lack of restriction sites between the *Sal*I and *Pvu*II sites of pJJ801 the positions of *hpcDE* could not be assigned. To confirm the gene order and to distinguish the sequence of *hpcDE*, Jenkins recovered the *Bam*HI insert of pJJ801 from a low melting point gel and partially restricted it with *Sau*3A. These fragments were ligated into *Bam*HI cut pBR328 and the mixture transformed into an *hpcC* mutant of *E. coli* C (JJ221). The transformants were tested for growth on HPA and those which were positive subjected to restriction mapping and full enzyme analysis. A summary of the

FIGURE 1.11a: SUMMARY OF THE SUBCLONES OF pJJ200/pJJ210 USED TO DETERMINE THE ORDER OF *hpc* GENES.



PLASMID	<i>hpc</i> GENES EXPRESSED							INDUCED BY HPA IN 5K.
	<i>hpcB</i>	<i>hpcC</i>	<i>hpcD</i>	<i>hpcE</i>	<i>hpcF</i>	<i>hpcG</i>	<i>hpcH</i>	
pJJ200	+	+	+	+	+	+	+	I
pJJ210	+	+	+	+	+	+	+	I
pJJ212	+	+	+	+	+	+	-	C
pJJ214	-	-	-	-	-	-	-	NT
pJJ215	+	-	-	-	-	+	-	C
pJJ216	-	-	-	-	-	-	-	NT
pJJ213	+	-	-	-	-	+	+	NT
pJJ211	+	+	+	+	-	+	+	I

FIGURE 1.11b: SUMMARY OF THE SUBCLONES OF pJJ801 USED TO DETERMINE THE ORDER OF *hpc* GENES.

		<i>hpc</i> genes expressed					
		B	C	D	E	F	I/C

	B E S P B						
pJJ801	-----	+	+	+	+	+	I
pJJ002	-----	+	-	-	-	-	NT
pJJ003	-----	-	-	-	-	-	NT
pJJ004	-----	+	-	-	-	-	NT
pJJ005	-----	-	-	-	-	-	NT
pJJ006	-----	+	+	+	+	-	I
pJRJ002	-----	+	+	+	+	-	C
pJRJ004	-----	-	+	+	+	-	C
pJRJ005	-----	+	+	-	-	-	I
pJRJ006	-----	+	+	-	-	-	C
pJRJ007	-----	+	+	-	-	-	I
pJRJ008	-----	+	+	+	-	-	I
pJRJ009	-----	-	+	+	-	-	C
pJRJ010	-----	+	+	+	-	-	I
pJRJ011	-----	-	+	-	-	-	C
pJRJ012	-----	-	+	+	+	-	C

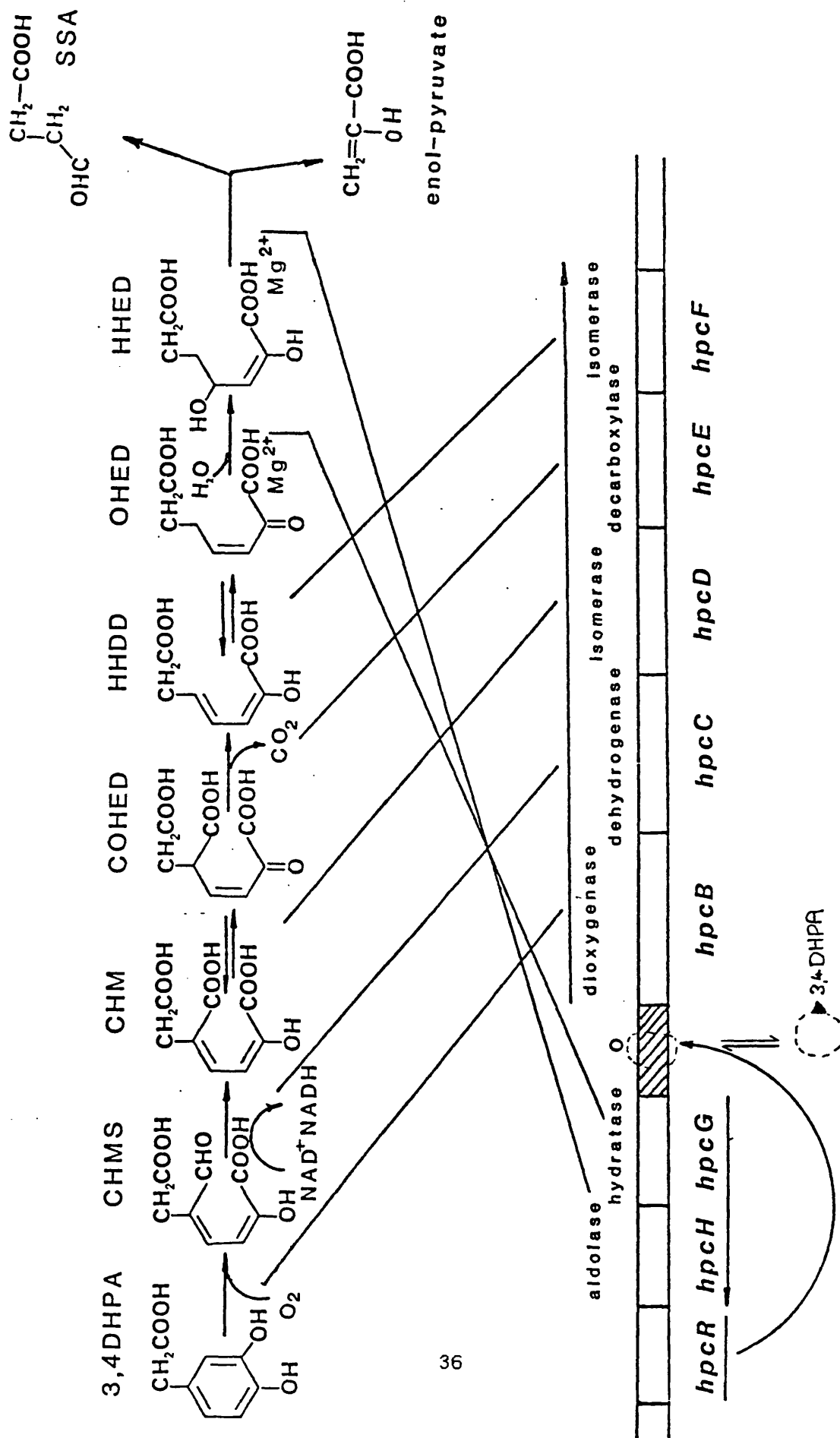
I/C refers to whether the plasmid is inducible or constitutively expressed in the *E. coli* C strain JJ221.

results obtained is presented in figure 1.11.

The expression of *hpc* genes on the various subclones in *E. coli* C strains was used to position an operator-promoter region. As removal of the 2.5kbp *Bam*H1-*Eco*R1 fragment of pJJ801 resulted in loss of expression of *hpc* genes (ie: pJJ002, but not pJJ003, expressed *hpc* genes) the operator-promoter site was thought to be in this area. Further, Jenkins indicated that from analysis of the *Sau*3A subclones of pJJ801 a site directly adjacent to the *hpcB* gene was indicated as the operator site. He also stated that the position of the operator region for the *hpcGH* gene block was 5' to the *hpcG* gene as only the phenotypes: hydratase+/aldolase+; hydratase+/aldolase- or hydratase-/aldolase- were observed. The phenotype hydratase-/aldolase+ was not seen indicating that the *hpcGH* gene block must be transcribed *hpcG* first. The model for the organisation and regulated expression of the *hpc* genes of *E. coli* C is shown in figure 1.12.

FIGURE 1.12: MODEL OF THE ORGANISATION OF *hpc* GENES AS PROPOSED BY JENKINS

(1987).



AIMS.

The aims of this work were three-fold:

- 1) To test the model for the organisation and expression of the *hpc* genes of *Escherichia coli* C, produce a finer map of the cloned genes and where necessary adjust the model.
- 2) Investigate the relatedness of *hpc* pathway genes in different organisms by a) comparison of a distinct enzyme of the HPC pathway from two organisms by N-terminal amino acid sequencing of purified proteins and b) Southern blot analysis of DNA from those organisms using nucleic acid probes-derived from either cloned genes or oligonucleotides synthesized using N-terminal amino acid sequence information.
- 3) To clone the *hpc* genes from another organism and compare them with the genes of *E. coli* C.

2: MATERIALS AND METHODS.

2.1: BIOCHEMICAL METHODS.

The bacterial strains and plasmids used in this study are listed in Table 2.1 and a diagram of the plasmids pUC18 and pUC19 is given in figure 2.1.

Growth Media and Conditions.

Bacteria were grown in either complex (Luria broth) medium, as described in Miller (1972), at 37°C or the minimal medium of Hareland *et al.* (1975) at 30°C. Liquid cultures were incubated in an orbital shaker at 200 r.p.m.; liquid media was solidified where appropriate by the addition of 1.6 % (w/v) Bacto agar (Difco).

Carbon sources and amino acids were sterilised separately and added to give final concentrations of: HPA (5 mM); HPC (5 mM); glycerol (20 mM); glucose (10 mM) and amino acids (80 $\mu\text{g ml}^{-1}$). Sodium dithionite (0.05% w/v) was added to medium containing HPC to slow the rate of spontaneous oxidation. Antibiotics were added as required to LB medium to final concentrations of: ampicillin (100 $\mu\text{g ml}^{-1}$); chloramphenicol (50 $\mu\text{g ml}^{-1}$) or tetracycline (12.5 $\mu\text{g ml}^{-1}$) and to minimal media at a quarter of these concentrations.

When selecting for the presence of inserts in the multiple cloning site of the vectors pUC18 and pUC19 by inactivation of the β -galactosidase gene, ampicillin (100 $\mu\text{g ml}^{-1}$) and isopropyl- β -D thiogalactopyranoside (IPTG, 300 μM) were incorporated into the medium. The agar plates were dried and then 50 μl of a solution of 2% (w/v) 5-bromo-4-chloro-3-indolyl- β -D galactoside (X-gal) in dimethyl formamide spread onto the surface prior to the addition of the bacteria.

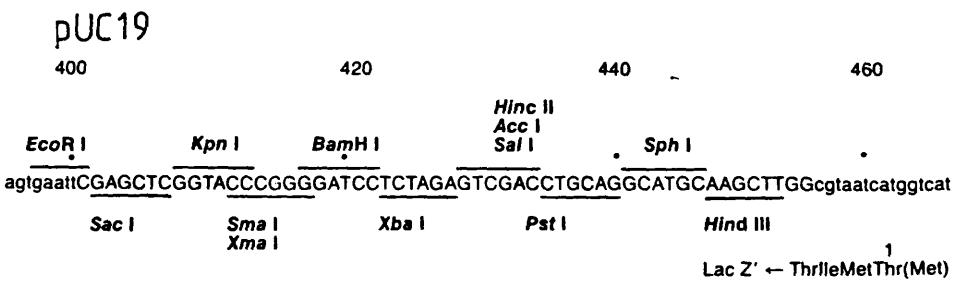
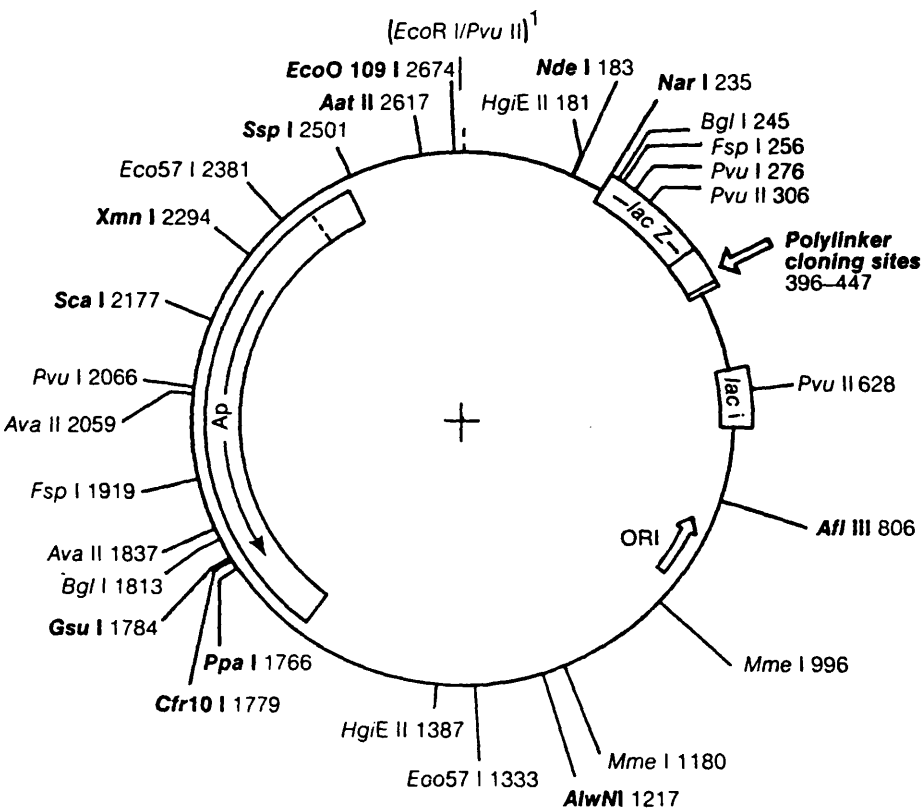
TABLE 2.1: BACTERIAL STRAINS AND PLASMIDS.

STRAIN	GENOTYPE	SOURCE/ REFERENCE
<i>Escherichia coli</i> C	Prototroph	Lab. stock
<i>E. coli</i> C JJ221	<i>hpcC</i> , <i>recA</i>	Jenkins, 1987
<i>E. coli</i> C JJ210	HPA ⁻ (even with pJJ801 present)	Jenkins, 1987
<i>E. coli</i> C D03	HPA ⁻ /HPC ⁺	Sudjadi, 1988.
<i>E. coli</i> 5K	F ⁻ , <i>supE</i> ⁴⁴ , <i>tonA</i> ²¹ , <i>hsdR</i> , <i>rpsL</i> , <i>thr-1</i> , <i>leu-B6</i> , <i>thi-1</i> , λ^-	Lab stock (Huback and Glover, 1970.)
<i>E. coli</i> K10	<i>thiA</i>	Lab. stock
<i>E. coli</i> B	Prototroph	Lab. stock
<i>E. coli</i> W	Prototroph	Lab. stock
<i>Pseudomonas putida</i> U	Prototroph	Lab. stock
<i>Klebsiella pneumoniae</i> M5a1	Prototroph	Lab. stock

PLASMID	VECTOR	GENES ENCODED	SOURCE/ REFERENCE
a) Plasmid vectors.			
	pBR322	Ap ^R , Tc ^R .	Soberon <i>et al.</i> , 1980
	pBR328	Ap ^R , Tc ^R , Cm ^R .	Soberon <i>et al.</i> , 1980
	pUC18/pUC19	Ap ^R .	Yanisch-Perron <i>et al.</i> , 1985
b) Recombinant plasmids containing cloned <i>E. coli</i> C genes.			
pJJ801	pBR328	<i>hpcBCDEF</i>	Jenkins, 1987.
pJJ200	pBR322	<i>hpcBCDEFGHR</i>	Jenkins, 1987.
pJJ210	pBR322	<i>hpcBCDEFGHR</i>	Jenkins, 1987.
c) Recombinant plasmids containing cloned <i>K. pneumoniae</i> M5a1 genes.			
pTF100	pBR322	<i>hpcBCDEFGHR</i>	This Study
pTF102	pBR322	<i>hpcBCDEFGHR</i>	This Study

FIGURE 2.1: THE PLASMID VECTORS pUC18 AND pUC19.

(Yanisch-Perron *et al.*, 1985)



Cell Free Extracts.

Bacteria from liquid media were harvested by centrifugation at 10,000 g for 5 min, washed in 0.2 volumes of 0.1 M sodium phosphate buffer pH 7.5 and resuspended in the same buffer in the ratio of 4 ml per 100 ml of the original volume. The cells were disrupted by ultrasonic oscillations in an MSE 100 W ultrasonic disintegrator in 4 ml samples at 7 μ m peak to peak amplitude, for 30 s, at 0°C. Remaining intact cells and cell wall material were removed by centrifugation at 20,000 g for 20 min. Where necessary the NADH oxidase activity associated with the cell membrane was removed by further centrifugation at 120,000 g for 90 min. All centrifugation steps were carried out at 4°C.

Protein Estimations.

The protein concentrations of crude cell free extracts were estimated by the method of Lowry *et al.* (1951), using crystalline bovine serum albumin as the standard protein. The concentration of protein was typically in the range of 2-4 mg ml⁻¹ from cells grown on minimal medium and 5-8 mg ml⁻¹ from cells grown on LB medium. The protein concentrations of purified protein solutions was estimated (spectrophotometrically) by the method of Warburg and Christian, 1941.

Washed Cell Suspensions and the Measurement of Oxygen Consumption.

Cells from logarithmic-phase cultures were harvested by centrifugation at 10,000 g for 5 min, washed twice in 0.1 M sodium phosphate buffer (pH 7.5) and resuspended in the same buffer to give a concentration of approximately 15 mg dry weight ml⁻¹. Consumption of oxygen was measured at 30°C in a Rank oxygen electrode, with a substrate concentration of 10 mM at 30°C. The solubility of oxygen was taken as 446 ng-atom ml⁻¹ at this temperature.

Preparation of Intermediates of the HPC Pathway.

E. coli 5K strains harbouring appropriate cloned *hpc* genes of *E. coli* C were utilised to obtain particular pathway intermediates. This method of preparing substrates is quantitative and produces pure compounds (Jenkins, 1987).

The general method used is as follows: 100 ml of Luria broth Ap cultures of strain *E. coli* 5K containing the appropriate plasmid was grown overnight and a cell free extract made as previously described. The activity of individual extracts was estimated by determining how much extract was required to completely convert a known quantity of HPC to CHMS. This assay only measures the activity of HPC dioxygenase but this is known to be the labile enzyme of the sequence, therefore, twice this amount was allowed in the incubation. The reaction mixture consisted in each case of 25 ml of 0.1 M sodium phosphate buffer pH 7.5 and 50 μ mol HPC. For the production of CHMS the reaction also contained extract from 5K(pJJ002); for the production of CHM, 67 μ mol NAD and extract from 5K(pJRJ003); for the production of HHDD, 67 μ mol NAD, and extract from 5K(pJJ801); and for the production of HHED, 67 μ mol NAD, 125 μ mol Mg^{2+} and extract 5K(pJJ212). The reaction mixture was gently shaken in a flask at 30°C until the reaction was complete as determined by the absence of HPC. This stage of the reaction was usually over within 10 minutes, but the incubation was allowed to continue for a further 10 minutes to ensure the reaction had gone to completion.

Concentrated HCl (1 ml) was added to the reaction which was then centrifuged to remove precipitated protein. The resulting supernatant was extracted three times with an equal volume of ethyl acetate; the pooled organic phases were dried over anhydrous sodium sulphate and evaporated to dryness under vacuum at 30°C in a rotary evaporator. The residue was dissolved in 1 ml of water and could be stored at -20°C or freeze dried and kept as a solid.

Enzyme Assays.

Enzyme assays were carried out in disposable plastic cuvettes ($\lambda > 320$ nm) or matched quartz cuvettes ($\lambda < 320$ nm) with a path length of 1 cm, at 30°C in a total volume of 1 ml, using a Pye-Unicam SP1800 recording spectrophotometer.

HPA HYDROXYLASE (EC 1.14.13.3)

(Skinner and Cooper, 1981)

The hydroxylation of HPA to HPC was measured by following the stoichiometric conversion of NADH to NAD at 340 nm. The reaction mixture contained 0.1 M sodium phosphate buffer pH 7.5, 0.13 μmol NADH and 10–100 μl of an ultracentrifuged cell extract. After measurement of the blank rate of NADH oxidation at 340 nm the reaction was started by the addition of 2.5 μmol HPA. A molar extinction coefficient of $6,200 \text{ dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$ was assumed for NADH.

HPC 2,3 DIOXYGENASE (EC 1.13.11.15)

(Skinner and Cooper, 1981)

Cleavage of the aromatic ring by HPC dioxygenase to form CHMS was monitored by the increase in absorbance at 380 nm. The reaction mixture contained 0.1 M sodium phosphate buffer pH 7.5, 0.2 μmol HPC and was started by the addition of 5–50 μl of extract. A molar extinction coefficient of $31,800 \text{ dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$ was assumed for CHMS at pH 7.5.

CHMS DEHYDROGENASE.

(Skinner and Cooper, 1981)

CHMS dehydrogenase was assayed by measuring the decrease in absorbance at 380 nm as CHMS was oxidised to CHM. The reaction mixture consisted of 0.1 M sodium phosphate buffer pH 7.5, 0.035 μmol CHMS and 5–50 μl of extract. The reaction was started by the addition of 0.2 μmol NAD. The molar extinction coefficient for CHMS was taken as $31,800 \text{ dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$.

CHM ISOMERASE

(Garrido-Pertierra and Cooper, 1981)

CHM isomerase activity was assayed by following the decrease in absorbance at 300 nm. The assay consisted of 0.1 M sodium phosphate buffer pH 7.5 and 0.05 μmol CHM; after measuring the blank rate the reaction was started by the addition of 5-100 μl of extract. The molar extinction coefficient for CHM was taken as $20,000 \text{ dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$.

COHED DECARBOXYLASE.

(Garrido-Pertierra and Cooper, 1981)

To produce the substrate for this assay, CHM was allowed to isomerise spontaneously in the cuvette to form a mixture of CHM and COHED and the absorbance at 300 nm monitored until an equilibrium mixture had formed (no further reduction in absorbance occurred). The reaction mixture consisted of 0.1 M sodium phosphate buffer pH 7.5 and 0.1 μmol spontaneously isomerised CHM; the reaction was started by the addition of 5-100 μl extract. The initial rate of reaction was obtained from the decrease in absorbance at 300 nm due to the spontaneous isomerisation of CHM to COHED to restore this equilibrium as the COHED is decarboxylated to form HHDD. The molar extinction coefficient of CHM was taken as $20,000 \text{ dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$.

HHDD ISOMERASE.

(Jenkins, 1987)

HHDD had a very rapid spontaneous isomerisation rate in phosphate buffer. However, the isomerisation rate can be slowed by carrying out the assay in deuterium oxide (D_2O). The reaction mixture consisted of D_2O , 20 μl 0.1 M sodium phosphate buffer pH 7.5 and 5-50 μl of extract. The molar extinction coefficient for HHDD was taken as $18,600 \text{ dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$.

OHED HYDRATASE.

(Jenkins,1987)

To assay the OHED hydratase enzyme an equilibrium mixture of HHDD and OHED was allowed to form in the cuvette, utilising the rapid spontaneous isomerisation rate in phosphate buffer. The reaction mixture consisted of 0.1 M sodium phosphate buffer pH 7.5 and isomerised HHED. When the decrease in absorbance at 276 nm had stopped 5 μ mol of magnesium chloride was added and the reaction was started by the addition of 5-50 μ l of extract. The initial rate of reaction was obtained from the decrease in absorbance at 276 nm as OHED is converted to HHED with the concomitant isomerisation of HHDD to OHED to maintain the equilibrium.

HHED ALDOLASE

(Jenkins, 1987)

The presence of the HHED aldolase was monitored by following the decrease in absorbance at 340nm as NADH was oxidised in the presence of lactate dehydrogenase (LDH) by the pyruvate formed from the HHED. The reaction was assayed in 0.1M sodium phosphate buffer pH8.0 containing 5 μ mol magnesium chloride, 0.15 μ mol NADH, 4 units of LDH and an excess of HHED.

SDS PAGE

Purification of protein was followed utilising 7.5 % to 20 % (w/v) gradient SDS PAGE. Samples of protein were mixed with 0.33 vol. sample buffer[0.18 M Tris.Cl pH 6.8; 5.7 % (w/v) SDS; 29 % glycerol; 0.005 % (w/v) bromophenol blue] and heated to 100°C for 2 min. The gel was developed at 4mA per track for 2.5h in an LKB vertical slab cell. The proteins were stained with Coomassie blue R 250 [50 % (v/v) methanol; 10 % (v/v) acetic acid 0.5 % (w/v) Coomassie blue R250] and destained in 7.5 % (v/v) acetic acid; 5.0 % (v/v) methanol. The following M_r standards were used as controls: bovine lactalbumin (14,200),

trypsin inhibitor (20,100), trypsinogen (24,000), carbonic anhydrase (29,000), glyceraldehyde-3-phosphate dehydrogenase (36,000), ovalbumin (45,000) and BSA (66,000).

Fast Protein Liquid Chromatography (FPLC).

The Pharmacia FPLC machine and columns were used according to the conditions recommended by the manufacturer. Columns used included the Mono Q anion exchange column; Phenyl superose hydrophobic interaction column and the Superose 12 gel filtration column. Full details of running conditions are given in the appropriate sections.

Amino Acid Sequencing.

The *E. coli* CHMS dehydrogenase was sequenced on an Applied Biosystems 470A gas-phase sequencer using a polybrene coated glass fibre disc and standard sequence cycles by Dr M.D. Davison.

Oligonucleotide Synthesis.

Oligonucleotides were synthesised by Mr. J. Kyte with an Applied Biosystems 380B DNA synthesizer using cyanoethyl phosphoramidate chemistry. The mixture of oligonucleotides was generated by mixed base addition.

GENETIC PROCEDURES.

2.2: PREPARATION AND MANIPULATION OF DNA AND RNA.

Preparation of Chromosomal DNA.

(Chow *et al.*, 1977)

High molecular weight chromosomal DNA was prepared from 200 ml overnight cultures of bacteria grown on LB medium. The cells were harvested by centrifugation at 7,000 g for 5 min at 4°C washed in 0.2 volumes of 10 mM Tris.Cl pH 7.9; 1mM EDTA; 1.0 M NaCl and resuspended in 30 ml of 10 mM Tris.Cl pH 7.9; 1 mM EDTA; 0.1 M NaCl containing 50 mg lysozyme. Incubation at 37°C for 10 min produced spheroplasts which were disrupted by the addition of 30 ml of the buffer containing 2% Sarcosyl NL97. After 1 h incubation at 42°C, self digested Proteinase K was added to a final concentration of 0.1 mg ml⁻¹ and the incubation continued overnight.

The cell lysate was extracted three times with an equal volume of neutral phenol (equilibrated with 0.5 M Tris.Cl pH 8.0), by gentle rotation on an orbital shaker for 30 min, and then once with an equal volume of chloroform/isoamyl alcohol (24:1 v/v). Chromosomal DNA was precipitated from the aqueous phase by two volumes of ethanol in the presence of 0.1 volumes of 3 M sodium acetate pH 5.2. The DNA was recovered by spooling onto a bent glass pipette and allowed to resuspend slowly in 10 mM Tris.Cl pH 7.4; 1 mM EDTA. This procedure typically yielded 4 mg of DNA with an average molecular weight of over 50 kbp.

Preparation of Plasmid DNA.

Small scale and large scale preparartions were made by the alkaline lysis method of Ish-Horowitz as described in Maniatis *et al.*, 1981. 1.5 ml cultures

were used to screen plasmid constructs of interest and larger, 50-400 ml, cultures were used to prepare stocks of plasmids.

An appropriate volume of an overnight LB (+ antibiotic) culture was harvested by centrifugation. The supernatant was removed and the pellet resuspended in 0.06 volumes of cold 25 mM Tris.Cl pH 8.0; 50 mM glucose; 10 mM EDTA and allowed to stand at room temperature for 5 minutes. This was followed by the addition of 0.12 volumes of alkaline SDS solution (1 % SDS in 0.2 M NaOH) and the contents of the tube mixed by three quick inversions. The mixture was incubated on ice for 5 minutes followed by the addition of 0.09 volumes of an ice cold solution of potassium acetate (3 M with respect to potassium and 5 M with respect to acetate; pH 5.0). The inverted tube was vortexed for 1 s and incubated on ice for a further 5 minutes. Precipitated chromosomal DNA and protein were removed by centrifugation for 5 minutes in a microfuge or at 15,000 g for 20 minutes depending upon the size of the sample. The supernatant was transferred to a fresh tube and extracted once with neutral phenol/chloroform, to remove remaining protein, and once with chloroform to remove residual phenol from the aqueous phase. Nucleic acids were precipitated by the addition of two volumes of ethanol and incubation at room temperature for two minutes. The precipitate was collected by centrifugation for 5 minutes in a microfuge or at 10,000 g for 15 minutes depending on the size of the sample.

Small scale preparations (from 1.5 - 50 ml of culture) were resuspended in 0.033 of the original volume of 10 mM Tris.Cl (pH 8.0), 1 mM EDTA (TE), containing 20 $\mu\text{g ml}^{-1}$ Ribonuclease A. The DNA obtained was suitable for transformation or restriction endonuclease analysis.

DNA from 400 ml of culture was prepared for dye-bouyant density equilibrium centrifugation by resuspending in 4.1 ml of TE, adding 4.3 g of CsCl and 218 μl of 10 mg ml^{-1} ethidium bromide. The refractive index of the solution was measured and if necessary adjusted to 1.3995-1.4000. The material was loaded

into a Beckman quick-seal tube and centrifuged overnight in a VT1 65.2 rotor of a Beckman L5-65 ultracentrifuge at 45,000 rpm. After centrifugation the covalently closed circular (ccc) plasmid DNA was located under long wave UV light, removed from the gradient with a syringe and large gauge needle, and the ethidium bromide removed by exhaustive extraction with sodium chloride saturated isopropanol. DNA was then precipitated with 2 volumes isopropanol in the presence of 0.1 volumes of 3M sodium acetate.

Agarose Gel Electrophoresis.

The molecular weights of DNA restriction fragments were analysed by electrophoresis through 0.5-1.0 % agarose gels. The agarose was made up with and run in TAE buffer (40 mM Tris-acetate pH 8.0; 1 mM EDTA). Before loading the DNA sample it was mixed with 0.1 volumes of loading solution [0.25 % (w/v) bromophenol blue; 0.25 % xylene cyanol (w/v); 30 % glycerol (v/v) in water] and centrifuged in a microfuge briefly to collect the sample. Gels were generally run at 5-10 V cm⁻¹ until the bromophenol blue dye had reached the end of the gel. Ethidium bromide (0.5 µg ml⁻¹) was included in the gel and this facilitated visualisation of DNA bands when illuminated with UV light. Permanent records of gels were made by photographing the fluorescence of DNA-ethidium bromide complexes under UV light.

Lambda DNA digested with the restriction enzyme *Hind*III was used to determine the size of unknown fragments: on some occasions *Xho* I and *Eco* RI were also used.

Restriction Enzyme Digestion.

Restriction digests were carried out according to the manufacturers recommendations (Bethesda Research Laboratories), using the buffers supplied. For digestions containing plasmid DNA, typically, 200-500 ng of DNA was digested in a 15 µl volume. Digestions with chromosomal DNA were in a 25 µl

volume and contained 5 µg of DNA. Following a restriction digest on any piece of DNA the fragments produced were electrophoresed through an agarose gel to ensure that the desired products had been achieved cleanly.

Phosphatase Treatment.

To minimise the number of recircularised vector molecules during any cloning experiment, terminal phosphate groups were removed by the action of calf intestinal alkaline phosphatase (CIP). CIP was purchased from Boehringer and added at 0.1 unit per microgram of DNA to the restriction digest after it had proceeded for one hour. The incubation was continued for a further 30 minutes; then the restriction enzyme and CIP removed by phenol extraction and the DNA precipitated with two volumes of ethanol in the presence of 0.1 volumes of 3M sodium acetate.

Isolation of DNA Fragments from Agarose Gels.

Approximately 1 µg of plasmid DNA was digested and electrophoresed through a 1 % ultrapure agarose gel as described previously. The required fragment(s) were located under long wave UV light and cut from the gel with a sterile scalpel blade. The gel slice was put into a 1.5 ml microfuge tube and centrifuged briefly to take it to the bottom of the tube. The tube was placed in a 65°C waterbath for 30 min to melt the agarose before the addition of one volume of butanol-saturated water and one volume of water saturated butanol. The mixture was vortexed vigorously for 2 min and the phases separated by centrifugation. The upper phase, but none of the interphase, was transferred to a fresh tube and the original mixture re-extracted with 0.5 volumes of water saturated butanol. Again the top layer was removed after vigorous vortexing and centrifugation and this was pooled with the first upper phase. This pooled material was extracted with 0.25 volumes of 0.3 M sodium acetate pH 7.0 and

after centrifugation the bottom phase was carefully collected. The solution was extracted once with an equal volume of chloroform/isoamyl alcohol [24:1 (v/v)] and the DNA precipitated with 2.5 volumes of ethanol in the presence of 4% (v/v) 3 M sodium acetate by incubation at -70°C for thirty min. The precipitate was collected by centrifugation for 10 min in a microfuge, dried briefly *in vacuo* and resuspended in water.

The solutions used were made as follows: An equal volume of butanol and water were vigorously shaken together and the phases allowed to separate. One gram of cetyl trimethyl ammonium bromide (CETAB) was dissolved in 100 ml of the upper (water saturated butanol) phase to which 100 ml of the lower (butanol saturated water) phase was added and the two mixed by shaking. The mixture was allowed to separate overnight and the phases put into separate bottles. Typically, 80% of any DNA fragment could be recovered by this method.

Ligation of DNA.

Plasmid and chromosomal DNA fragments were ligated using BRL ligation buffer and 1 unit of BRL T4 DNA ligase. The fragments were mixed in the ratio of 2:1, insert to vector, by the number of moles of ends. The reaction was allowed to proceed for 2-16 hours at room temperature before the mixture was used to transform the appropriate bacterial strain.

Transformation of *E. coli* with Plasmid DNA.

(Kushner, 1978)

The method of Kushner was used to make all bacterial strains, used in this study, competent to take up DNA from solution.

The recipient cells were grown in LB medium to an OD₆₈₀ of approximately 0.5 for *recA* strains and 0.2 for *recA*⁺ strains. The cells were harvested by centrifugation and washed in 0.33 volumes of 10 mM MOPS pH 7.0; 10 mM RbCl. The cells were resuspended in 0.33 volumes of 100 mM MOPS pH 6.5; 10 mM RbCl;

50 mM CaCl₂ and incubated on ice for 30-90 minutes. The cells were collected by a short spin and resuspended in 0.1 of the original volume of the same solution. DMSO was added to 0.2% (v/v) and 1-50 ng of DNA added. The mixture was incubated on ice for 1 hour and then heat shocked at 55°C for exactly 30s, placed in an ice/water bath for 2 minutes and 0.66 of the original volume of sterile LB added. The cells were incubated at 37°C for one hour to allow expression of the antibiotic resistance genes and then an aliquot spread onto an appropriate prewarmed LB-antibiotic agar plate.

Preparation of Total RNA.

Total cellular RNA was prepared by a modification of the method of Aiba *et al.*, (1981). Bacteria were grown to an OD₅₅₀ of 0.3 in minimal medium plus the required carbon source with any growth factors/antibiotics as required. The cells were harvested at 4°C by centrifugation, momentarily at 10,000g and quickly resuspended in 3 ml of 20 mM sodium acetate, pH 5.5; 0.5% (w/v) SDS; 1 mM EDTA. To the resuspended cells, 3 ml of redistilled phenol (equilibrated with 20 mM sodium acetate, pH 5.5) was added and the resulting mixture vortexed for 10 s. The sample was then incubated at 60°C for 5 minutes with gentle agitation. The phases were separated by centrifugation and the aqueous portion extracted a further twice with phenol and once with chloroform/isoamyl alcohol. The nucleic acids were precipitated with three volumes of ethanol and incubation at -70°C for 30 minutes. After centrifugation the pellet was dried *in vacuo* and then resuspended in 2 ml of sterile water. RNA was selectively precipitated by the addition of 3 volumes of 4 M sodium acetate, pH 6.0 and incubation overnight at 4°C (Kirby 1965). The RNA was collected by centrifugation and the pellet drained well before being resuspended in 1.0 ml of sterile water.

This method usually yielded 300-400 µg of RNA with a 260/280 ratio of 1.95-2.05 which was then stored as an isopropanol suspension at -20°C. When

vortexed the suspension was essentially homogeneous and the appropriate volume could be taken, precipitated and used in an experiment.

Formaldehyde/Agarose Gel Electrophoresis.

Total RNA was fractionated through 1% agarose gels containing 2.2 M formaldehyde essentially as described in Maniatis *et al.*, (1982).

Gels were made by adding 50 ml of a 2% agarose solution, maintained at 60°C, to 50 ml of a prewarmed solution containing: 20 ml 5x gel running buffer (MSE: 0.2 M MOPS, pH 7.0; 50 mM sodium acetate; 5 mM EDTA); 18 ml of 38% (w/v) formaldehyde and 12 ml water. The gels were poured, and all subsequent manipulations involving formaldehyde carried out in a fume hood.

RNA samples were prepared by collecting 50 µg of total RNA from isopropanol suspensions by centrifugation; the pellet was washed in 70% ethanol and dried briefly *in vacuo* then resuspended in 4.5 µl of water. To this 3.5 µl of formaldehyde, 10 µl of deionised formamide and 2 µl of 5x MSE were added. The sample was mixed and incubated at 55°C for 15 minutes followed by the addition of 2 µl of gel loading solution [50% glycerol; 1 mM EDTA; 0.4% (w/v) bromophenol blue; 0.4% (w/v) xylene cyanol]. The samples were centrifuged momentarily to collect them in the bottom of the tube before loading onto the gel. The running buffer was 1x MSE and gels were electrophoresed at 5V cm⁻¹ until the bromophenol blue dye had reached the edge of the gel.

The 23s and 16s rRNA species were used as size markers. In a separate lane 5 µg of RNA, prepared in the same way, was run; the RNA was blotted onto a nitrocellulose filter and baked and then the track containing the marker RNA was cut from the filter and stained. RNA was visualised by treating the filter with 5% acetic acid for 15 min, then staining with 0.04% (w/v) methylene blue in 0.5 M sodium acetate pH 5.2 for 15 min and destaining in water. The positions of the two bands were marked with pencil as they fade with time.

2.3: BLOTTING AND RADIOLABELLING TECHNIQUES.

Transfer of Nucleic Acids to Nitrocellulose.

DNA and RNA were transferred from agarose gels to nitrocellulose by capillary action essentially as described by Southern (1980).

DNA was denatured before transfer by illuminating with UV light for 30s and then immersing the gel in a solution of 0.5 M NaOH; 1.5 M NaCl for two 15 min periods with gentle shaking. The gel was neutralised by two 15 min incubations in a solution of 1.0 M Tris.Cl, pH 8.0; 1.5 M NaCl and the DNA was transferred using an apparatus shown in figure 2.2. The wick was pre-soaked in reservoir buffer; the membrane and other pieces of 3MM paper were pre-treated in 2xSSC (1xSSC is 0.15 M NaCl; 0.015 M Na citrate)

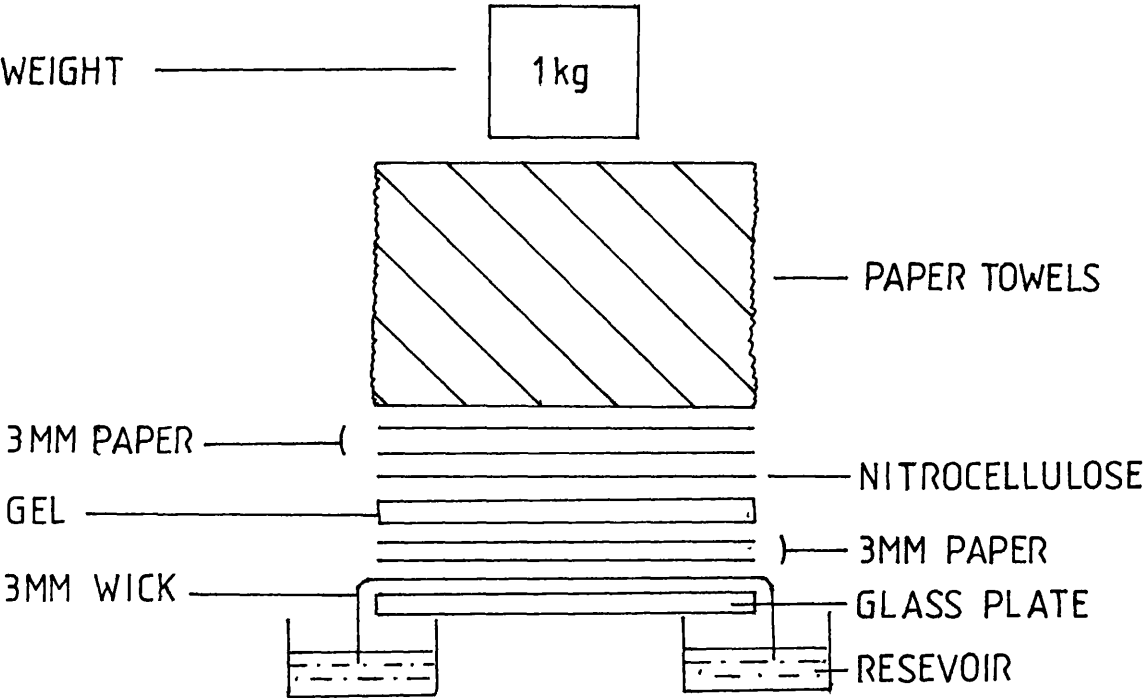
RNA was transferred to a nitrocellulose membrane that had been equilibrated with 20xSSC, without any treatment of the gel using the same apparatus as before. Transfers were allowed to proceed overnight after which the origin and orientation of the gel was marked on the filter in pencil and the filters baked at 80°C for two hours between two pieces of 3MM paper.

Colony Hybridisation.

(Grunstein and Hogness, 1975)

Colony hybridisation permits rapid screening of bacterial colonies to determine which contain specific DNA sequences of interest. Bacterial colonies containing recombinant plasmids were replica plated or patched onto nitrocellulose filters which were resting on LB-ampicillin containing agar plates. Colonies present after overnight incubation at 37°C were lysed by resting the filter, colony side up, on a stack of 3MM chromatography paper soaked in 0.5 M NaOH for 7 min. The nitrocellulose filter was placed on 3MM paper soaked in 1.0 M Tris.Cl

FIGURE 2.2: APPARATUS FOR THE TRANSFER OF NUCLEIC ACIDS TO NITROCELLULOSE.



(pH 7.4) for 4 min and then to 3MM paper soaked in 1.5 M NaCl; 0.5 M Tris.Cl (pH 7.4) for a further 4 min.

Radiolabelling of Oligonucleotide Probes.

Single stranded oligodeoxynucleotides, lacking a 5' phosphate group, were radiolabelled using [γ - 32 P] dATP and polynucleotide kinase (BRL) in the following reaction;

50 pmol 5' ends of DNA	x μ l
10x kinase buffer I	10 μ l
33 pmol [γ - 32 P] dATP	10 μ l
(Specific activity 3000 Ci mmol $^{-1}$)	
T4 polynucleotide kinase (10 units)	1 μ l
water (to 50 μ l)	

(Kinase buffer I is 500 mM Tris.Cl pH 7.6; 100 mM MgCl $_2$; 50 mM dithiothreitol; 1mM spermidine and 1mM EDTA.)

The mixture was incubated at 37°C for 30 min in a lead container, followed by a 15 min period at 65°C to inactivate the kinase enzyme.

Radiolabelling of DNA Restriction Fragments.

Plasmid DNA was cleaved with restriction enzyme(s) to give the required fragment for labelling. The digest was electrophoresed through a 1% low melting point agarose gel, containing 0.2 μ g ml $^{-1}$ ethidium bromide, and the required band excised with a minimum amount of extraneous agarose. The band was placed in a preweighed 1.5 ml microfuge tube and water added at a ratio of 1.5 ml of water per gram of agarose. A visual estimate of the quantity of DNA obtained was made from the gel so that the approximate concentration of DNA could be made. The agarose was melted and DNA denatured by placing the tube in a boiling water bath for 7 min.; the sample was maintained at 37°C prior to labelling.

The labelling reaction was carried out at room temperature by addition of the following reagents in the stated order:

- x μ l water (to a total of 25 μ l)
- 5 μ l OLB buffer
- 1 μ l BSA (BRL 5561, 10 mg ml⁻¹)
- x μ l DNA fragment (25 ng) up to 16.25 μ l
- 2.5 μ l [α -³²P] dCTP (10 μ Ci μ l⁻¹)
- 0.5 μ l Klenow fragment of DNA polymerase I (3 units)

The reaction was allowed to proceed for 5-16 hours then terminated by the addition of 100 μ l of stop buffer. The solutions used in this procedure are listed below.

TE BUFFER:	3 mM Tris.Cl pH 7.0; 0.2 mM EDTA.
SOLUTION O:	125 M Tris.Cl pH 8.0; 125 mM MgCl ₂ .
dNTP's:	100 mM solutions of dATP, dGTP, TTP in TE.
SOLUTION A:	1.0 ml solution O; 18 μ l 2-mercaptoethanol; 5 μ l each of dNTP solutions.
SOLUTION B:	2 M HEPES pH 6.6.
SOLUTION C:	Hexadeoxynucleotides, dissolved in 550 μ l TE to give a concentration of 90 OD ₂₆₀ ml ⁻¹
OLB BUFFER:	A mixture of solutions A, B and C in the ratio 10:25:15.
STOP BUFFER:	20 mM NaCl; 20 mM Tris.Cl, pH 7.5; 2 mM EDTA; 0.25% (w/v) SDS; 1 μ M dCTP.

Hybridisation of Radiolabelled Probes to Nucleic Acids.

The conditions of hybridisation of radiolabelled DNA to both DNA and RNA

immobilised on nitrocellulose membranes were essentially those of Maniatis *et al.*, (1981).

After baking, membranes were pre-hybridised in a solution of 6x SSC; 5x Denhardt's solution; 0.5% (w/v) SDS and 100 $\mu\text{g ml}^{-1}$ salmon sperm DNA for 1-4 hours. (Salmon sperm DNA was prepared by passing through a syringe needle and then boiling for 10 min; Denhardt's solution is 0.1% (w/v) ficoll; 0.1% (w/v) polyvinylpyrrolidone; 0.1% (w/v) BSA and 1x SSC is 0.15 M NaCl; 0.015 M Na citrate). The pre-hybridisation solution was removed and replaced by a hybridisation solution of the above containing 10 mM EDTA and radiolabelled probe. Hybridisations were allowed to continue overnight, usually 16 hours, after which the hybridisation solution was removed and the membrane washed to remove non-specific binding. The washing procedure employed was: 2x SSC; 0.5% (w/v) SDS for 5 min at room temperature, 2x SSC; 0.1% (w/v) SDS for 15 min at room temperature followed by two washes in 0.1x SSC; 0.5% (w/v) SDS for two hours at the temperature of hybridisation. The membrane was allowed to air dry, or alternately blotted dry, wrapped in Saran Wrap and exposed to X-ray film (Fuji RX 100). Intensifying screens were used where required and prolonged exposure was carried out at -70°C to reduce the effects of background radiation.

Hybridisation and final wash temperatures were determined by the method of Suggs *et al.*, (1981) for oligonucleotide probes and (Wetmar and Davidson, 1968) for labelled restriction fragments.

2.4: DNA SEQUENCING.

DNA sequence information was obtained by a modification of the denatured plasmid method of Hattori and Sakaki (1985).

Preparation of plasmid DNA: Overnight cultures of *E. coli* NM522 harbouring pUC recombinant plasmids were grown in LB-ampicillin: 1.5 ml of culture was taken for centrifugation at 12,000 rpm for 2 min in a 1.5 ml microfuge tube. The cell pellet was suspended and lysed gently in 100 μ l of lysozyme solution (0.4% [w/v] egg white lysozyme, 50 mM glucose, 25 mM Tris.Cl and 10 mM EDTA, pH 8.0) at room temperature for 5 min. Then 0.2 M NaOH-1% SDS solution (200 μ l) was added and mixed gently. After 5 min incubation on ice, 150 μ l of potassium acetate solution (3 M with respect to potassium and 5 M with respect to acetate) was added and mixed gently. After 5 min on ice, the sample was centrifuged at 12,000 rpm for 10 min. The supernatant was removed with a pipette and extracted once with phenol/chloroform/isoamyl alcohol (25:24:1 v/v) and once with chloroform. Two volumes of isopropanol were added to the aqueous phase and precipitates collected by centrifugation for 5 min at 12,000 rpm. The pellet was washed with ethanol, dried under vacuum and dissolved in 100 μ l TE. Ribonuclease A was added to a final concentration of 10 μ g ml⁻¹ and the solution incubated at 37°C for 30 min and then extracted with phenol and then with chloroform. Precipitates formed by the addition of 2 volumes of isopropanol and 0.1 volumes of 3 M sodium acetate (pH 5.2) after 5 min at room temperature were collected by centrifugation, washed in ethanol and vacuum dried.

Alkaline denaturation of plasmid DNA: The plasmid DNA was dissolved in 40 μ l TE and 15 μ l of this mixed with 40 μ l sterile water: 4 μ l of 2 M NaOH was added, mixed and kept at room temperature for 5 min. Then, 4 μ l of 2 M ammonium acetate (pH 4.5) was added and the denatured DNA precipitated by addition of 90 μ l of ethanol and cooling to -70°C for 5 min. The precipitates were

harvested by centrifugation at 12,000 rpm for 5 min, rinsed in 70% ethanol and dried under vacuum

Sequencing: The mixture of denatured plasmid (approximately 1 pmol) dissolved in 9.5 μ l water, primer (0.5 pmol in 1 μ l), 1.5 μ l of 10X Klenow buffer (100 mM Tris.Cl, pH 7.5; 660mM NaCl; 66mM MgCl₂) and 3 μ l of ³⁵S-dATP (1000 Ci mmol⁻¹) was heated at 37°C for 15 minutes in a 0.5 ml microfuge tube. Klenow fragment of *E. coli* DNA polymerase I (5 units, 1 μ l) was added and mixed well. The sample was immediately aliquoted into four 3 μ l parts and each part was mixed with 2 μ l of G-, A-, T- and C- specific dideoxy-deoxynucleotide mixture (see Table 2.3). The reactions were carried out at 50°C for 20 min, then a chase solution (2 μ l of a 1.25 mM dNTP solution) was added, mixed and further incubated at the same temperature for 20 min. Finally, 4 μ l of loading solution (90% formamide, 0.3% bromophenol blue, 0.3% xylene cyanol, 10 mM EDTA pH 8.3) was added and the samples kept on ice. Aliquots of the four reaction mixtures (2 μ l) were loaded onto the sequencing gel immediately following heating at 95°C for 3 min.

Sequencing gel and autoradiography: A 6% acrylamide-7M urea gel in 1x Tris-Borate-EDTA (TBE) electrophoresis buffer was employed and electrophoresis was carried out at 1700V with 1x TBE in the bottom reservoir and 0.5x TBE in the top reservoir. Following electrophoresis, the gel on one glass plate was immersed in 10% methanol-10% acetic acid for 30 min, transferred to Whatmann 3MM paper and dried on the paper at 80°C under vacuum. Autoradiography was done for 16-24 h at room temperature without intensifying screens.

Analysis of DNA sequence Data.

The DNA sequence data produced was analysed with the aid of packages available on the University mainframe computer.

TABLE 2.2: DIDEOXY-DEOXYNUCLEOTIDE SEQUENCING MIXTURES.

ddNTP-dNTP Mixture for Sequencing.

	dGTP	dTTP	dCTP	ddATP	ddCTP	ddGTP	ddTTP
ddATP-dNTP	62.5	62.5	62.5	31	-	-	-
ddCTP-dNTP	160	160	16	-	91	-	-
ddGTP-dNTP	16	160	160	-	-	91	-
ddTTP-dNTP	88	9	88	-	-	-	1000

The units shown are in μM . The mixture is dissolved in 25 mM Tris.Cl (pH 7.9), 16.5 mM MgCl_2 , 165 mM NaCl.

3: ORGANISATION AND EXPRESSION OF *Escherichia coli* hpc GENES.

3.1 FURTHER INVESTIGATIONS INTO THE ORGANISATION OF THE *hpc* GENES.

Southern Blot Analysis to Identify the Position of the *hpcG* Gene.

The product of the *hpcG* gene, 2-oxo-hept-3-ene-1,7-dioate (OHED) hydratase, had previously been purified and the amino-terminal amino acid sequence determined by automated techniques (Ferrer and Cooper, 1988). This analysis identified 34 residues (fig 3.1a), and examination of the sequence indicated a region suitable for the construction of a degenerate oligonucleotide probe. Of the first five residues Met is encoded by a unique triplet and the codons for Phe-Asp-Lys-His show only two degeneracies each. The sixth residue, Thr, has four degeneracies. Since the codon preference of *E. coli* C is not known a mixture of oligonucleotides that took into account the degeneracies in the coding of the first five residues but with only the first two bases corresponding to the sixth residue was constructed. This means that the probe consisted of a mixture of 16 different 17mer oligonucleotides as depicted in figure 3.1b. The mixed oligonucleotide probe, probe G, was synthesized without a 5'-phosphate group and could therefore be labelled directly with [γ - 32 P]-dATP and T4 polynucleotide kinase and used to detect the 5' end of the *hpcG* gene.

The *Bam*H1-*Bam*H1 subclone of pJJ200 that contains the 7.3 kbp fragment that extends from 2.9 kbp to 10.2 kbp on the restriction map, (pJJ212, fig.1.11a) expresses *hpcBCDEFG*, whereas, the genomic clone pJJ801 expresses *hpcBCDEF* (fig. 1.11b). This suggested that either the *hpcG* gene is not expressed or is missing totally or in part from pJJ801. To determine the location of the 5' end of the *hpcG* gene the plasmids pJJ200, pJJ212, pJJ801 and pJJ002 were analysed by a DNA-DNA dot blot. A sample of each of these plasmids was denatured and applied to a nitrocellulose membrane filter along with the plasmid controls pBR322 and pBR328. After baking the filter was pre-hybridised for 4 h at 30°C and then hybridised for 16 h at 30°C with radiolabelled probe G. As the actual DNA sequence was not known the temperature of dissociation of the probe from the

FIGURE 3.1a: SEQUENCE FROM THE FIRST 34 RESIDUES OF THE AMINO TERMINUS
OF OHED HYDRATASE.

Residue Number:	1	2	3	4	5	6	7	8	9	10
Residue:	Met	Phe	Asp	Lys	His	Thr	His	Thr	Leu	Ile
Residue Number:	11	12	13	14	15	16	17	18	19	20
Residue:	Ala	Gln	Arg	Leu	Asp	Gln	Ala	Glu	Lys	Gln
Residue Number:	21	22	23	24	25	26	27	28	29	30
Residue:	Arg	Glu	Gln	Ile	Arg	Ala	Ile	Ser	Leu	Asp
Residue Number:	31	32	33	34						
Residue:	Tyr	Pro	Glu	Ile						

FIGURE 3.1b: SEQUENCE OF THE SYNTHETIC OLIGONUCLEOTIDES CORRESPONDING TO
THE FIRST SIX AMINO ACID RESIDUES OF OHED HYDRATASE.

Residue Number:	1	2	3	4	5	6
Residue:	Met	Phe	Asp	Lys	His	Thr
Probe G:	ATG	TTT	GAT	AAA	CAT	AC
		C	C	G	C	

target sequence was calculated for that of the most weakly-hybridising species in the probe mixture by the empirical formula of Suggs *et al.*, (1981). The 17bp probe contained a maximum of thirteen A+T residues, thus from the calculation:

$$\begin{aligned} T_m(^{\circ}\text{C}) &= 2(A+T) + 4(G+C) && (\text{Suggs } et al., 1981) \\ &= 2(13) + 4(4) \\ &= 42^{\circ}\text{C} \end{aligned}$$

the temperature of dissociation (T_m) was calculated to be 42°C. Hybridisations were carried out at $T_m - 12^{\circ}\text{C}$ as in Maniatis *et al.*, 1982.

The recombinant plasmids pJJ200, pJJ212, pJJ801 and pJJ002, but neither of the plasmid controls pBR322 or pBR328 (fig. 3.2a) hybridised to probe G under the conditions tested indicating that at least the 5' end of the *hpcG* gene is present on pJJ801. This result was confirmed by Southern blot analysis of restriction fragments of the recombinant plasmids digested with either *Bam*H1 or a combination of *Bam*H1 and *Eco*R1 and hybridised to probe G (fig. 3.2b), using the same conditions of hybridisation and washing as for the dot blot experiment. The smallest fragment which hybridised to the probe was the insert of pJJ002; in each case the hybridising band corresponded to the insert between 3.9 kbp and 6.3 kbp or 2.9kbp and 6.3kbp on the restriction map (fig. 1.10). Neither of these experiments gave any indication of the position of the *hpcG* gene in relation to the other genes or whether the gene was intact on pJJ801 or pJJ002.

To facilitate the localisation of the 5' end of the *hpcG* gene, further restriction enzyme sites were sought in the region of the insert which hybridised to the probe. Previously, *Bcl*1, *Cla*1, *Hind*III, *Nco*1, *Pst*1, *Fvu*1, and *Sph*1 were found not to cut in the insert of pJJ801 (Jenkins, 1987). In addition *Bgl*III, *Sst*1, *Stu*1 and *Xba*1 also failed to restrict pJJ200. However, *Nde*1 and *Eco*RV digested the insert DNA of both pJJ200 and pJJ002 and the sites at which

FIGURE 3.2a: DNA-DNA dot blot hybridisation of various plasmids with probe G: The plasmids used were A, pJJ200; B, pJJ801; C, pJJ002; D, pJJ212; E, pBR322 and F, pBR328.

FIGURE 3.2b: Southern blot analysis of various plasmids with probe G: Panel A shows a photograph of agarose gel electrophoresis of restriction enzyme digested DNA and panel B shows the resulting blot after transfer of the DNA to nitrocellulose and probing with radiolabelled probe G. Lane 1 contained *Hind*III restricted λ -DNA; lane 2, *Bam*H1/*Eco*R1 digested pJJ212; lane 3, *Bam*H1 digested pJJ002; lane 4, *Bam*H1 digested pBR328; lane 5, *Bam*H1 digested pJJ212; lane 6, *Bam*H1/*Eco*R1 digested pJJ002; lane 7, *Bam*H1/*Eco*R1 digested pJJ200 and lane 8 *Bam*H1 digested pJJ200. The film was exposed to the blot for 3 hours at room temperature.

FIGURE 3.2a: DNA-DNA DOT BLOT HYBRIDISATION OF VARIOUS PLASMIDS WITH PROBE G.

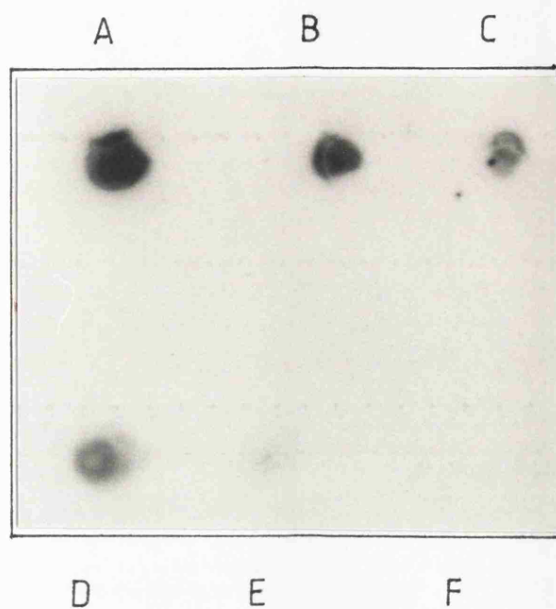
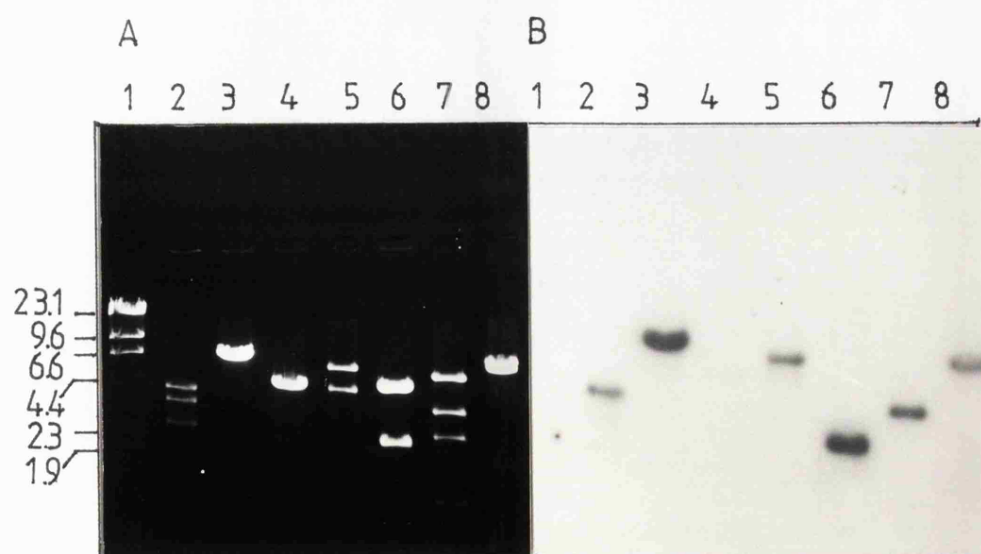


FIGURE 3.2b: SOUTHERN BLOT ANALYSIS OF RESTRICTION FRAGMENTS OF VARIOUS PLASMIDS WITH PROBE G



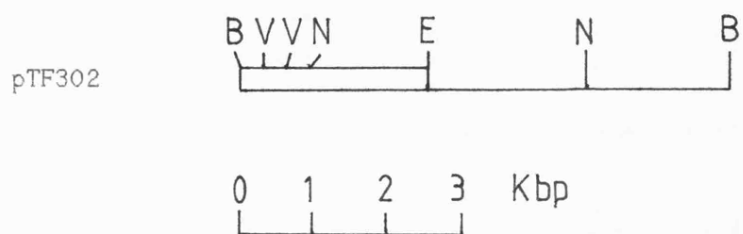
they cut were subsequently mapped in pTF302. (The subclone pTF302 contains the 2.4 kbp *Bam*H1-*Eco*R1 fragment of pJJ801, which extends from 3.9 kbp to 6.3 kbp on the restriction map, and was constructed to standardise the vector as pBR322.) One site for *Nde*I and two sites for *Eco*RV were mapped on this DNA, all of which were clustered towards the *Bam*H1 site as depicted in figure 3.3. Restriction digests of pTF302 were carried out either singly or in combination, the DNA electrophoresed through a 0.8% agarose gel and then blotted onto a nitrocellulose filter. After baking, the DNA was hybridised to radiolabelled probe G as before. The 5' end of the *hpcG* gene was localised to the 300bp *Eco*RV-*Eco*RV fragment (fig. 3.3) of pTF302 which lies between 4.2 kbp and 4.5 kbp on the restriction map (fig. 3.3); this hybridising fragment lies between 300 and 600bp from the *Bam*H1 site of pTF302. The subunit molecular weight of OHED hydratase estimated by SDS-PAGE is 32,500 (Ferrer and Cooper, 1988); when converted into DNA terms a protein of that size would be encoded by a gene of approximately 900 bp. As no OHED hydratase activity can be detected in cell extracts of *E. coli* 5K (pTF302) these results suggest that the *Bam*H1 site lies within the *hpcG* gene and that it is transcribed from right to left as drawn in figure 1.10.

DNA Sequencing of the 5' End of the *hpcG* Gene.

The direction of transcription of the *hpcG* gene was confirmed by DNA sequencing of the region within about 700 bp of the *Bam*H1 site of pTF302. The *Eco*RV-*Eco*RV fragment, to which probe G hybridised, was extracted from a 1% low melting point (LMP) agarose gel and ligated into the *Sma*I site of pUC18. The construct produced, pTF830, was sequenced from each end using universal and reverse oligonucleotide primers. This procedure gave single stranded DNA sequence over the whole length of the fragment and double stranded sequence in the central two-thirds of the region (fig. 3.4). As *Eco*RV is a restriction

FIGURE 3.3: Localisation of the 5' end of the *hpcG* gene: Panel A shows a photograph of agarose gel electrophoresis of restriction enzyme digested DNA and panel B shows the resulting blot after transfer of the DNA to nitrocellulose and hybridisation with radiolabelled probe G. Lane 1 contained *HindIII* restricted λ -DNA; lane 2-6 contained pTF302 DNA digested with: lane 2, *EcoRI*/*NdeI*; lane 3, *BamHI*/*NdeI*; lane 4, *EcoRV*; lane 5, *BamHI*/*EcoRI* and lane 6, *BamHI*

FIGURE 3.3: LOCALISATION OF THE 5' END OF THE *hpc* G GENE.



SOUTHERN BLOT ANALYSIS OF RESTRICTION FRAGMENTS OF pTF302.

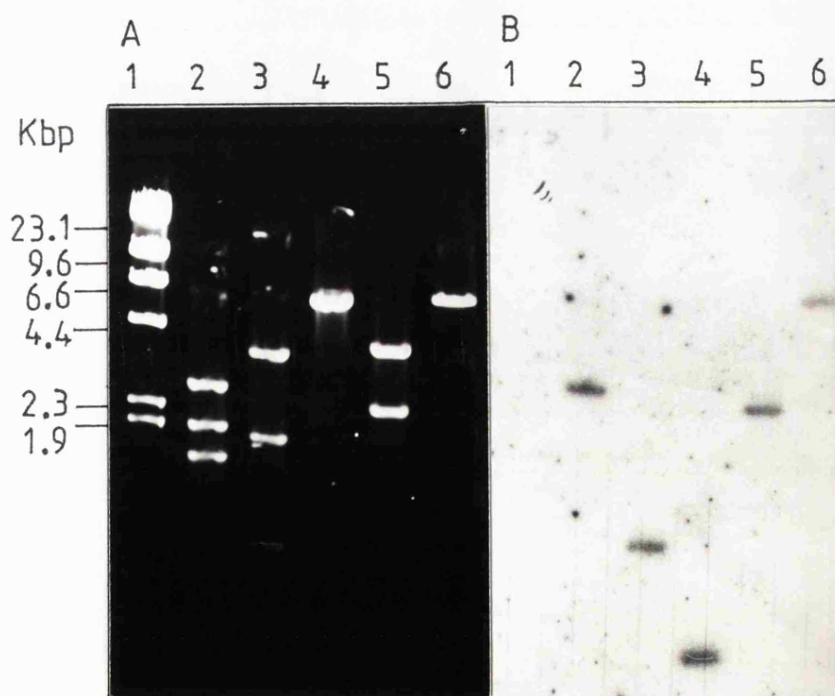
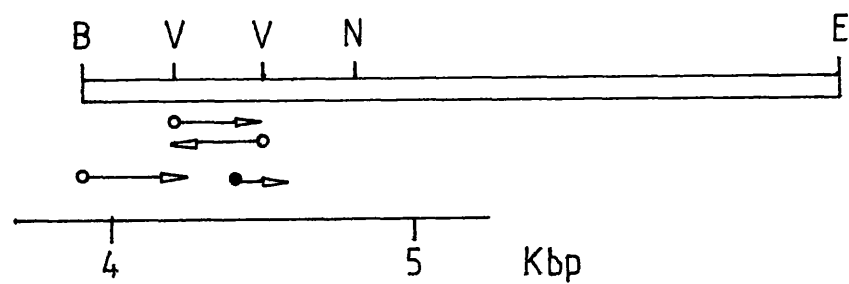


FIGURE 3.4: Sequence strategy used to identify the *hpcG* gene: The *EcoRV*-*EcoRV* (V-V) fragment was cloned into pUC18 and sequenced with universal and reverse primers; the *Bam*H1-*Eco*R1 (B-E) fragment was also cloned into pUC18 and sequenced with universal primer. In addition a synthetic oligonucleotide taken from a portion of the sequence was used to extend the information towards the *Eco*R1 site (closed circle). The scale shown refers to the position on the restriction map.

FIGURE 3.4: SEQUENCE STRATEGY USED TO IDENTIFY THE *hpcG* GENE.



DNA SEQUENCE INFORMATION.

	pUC DNA	<u>Bam</u> H1	INSERT DNA
1	GTCGACTCTA	GAGGATCCAG	TTGTCTGGTT TGATCGGACG ACCTCCGAGG
51	ATCGCCCCGG	CGTTGGCGGC	GTTATANTGG TGTCGAACAC TTTGCGCGGG
101	CGTGGGTTTC	CGGATCGATG	TTGTTGTGGC AGCGAGCGTC GATCAGCTCC
151	AGCGCCGGGA	TCAAGTAGTC	CGTGGCCGTT GTAGACGTCG AACAGCGTGC
201	AGTTTGGCCC	ACGCAGCGGT	TTTGCCAGCA CGAAAGCCAG CTCCACTTCA
251	TGCGCGGCAC	CATAAGCGAT	CGGTCTGGGAT ATCGGTTCAT CGAGGAAGAA
301	CATGTCATCC	AGCAGCGGCC	GTAATCCGGT TCGCCGATCT GCGAGCTTGC
351	CTGCATCGCT	TTCGAGGTCA	GGCCGATTTT GTGGCCTTTC AGCGTGCGAC
401	CTTCGGCGAT	TTTCAGGCCA	ACCCATTAC GCTGCACCGC GTAAGCGTCT
451	TCGGATGGGG	ATTTCTGGGGT	AATCCAGGCT GGAGGCGCGA ATCTGTTTCGC
501	GCTGTTTTTC	TGCCTGATCG	TCACGCTGGG CGATCAGGGT GTGGGTGTGT
551	TTATCGAACA	TAGCGATATC	CTGTGGCCAT TGCCGGATGC GGTTCGCGCT
601	TTATCCGCTA	CGGTGACGGT	TTGGGGTTTT GTAGGGCCCG GTAAGCGAGC
651	CACCGG		

endonuclease which produces blunt ended fragments, the orientation of the fragment could not be determined from this experiment. To this end the whole of the 2.4 kbp fragment from pTF302 was cloned into pUC18; the resultant plasmid, pTF802, was sequenced using universal primer so that DNA sequence information from the *Bam*H1 end of this clone, overlapping the information obtained from pTF830, could be obtained. Finally a synthetic oligonucleotide primer, complementary to a 17 bp stretch of the sequence, was used to extend the sequence information to a point beyond the 5' end of the gene. A diagram of the sequencing strategy used and the alignment of DNA and protein sequence can be seen in figure 3.4.

The DNA sequence information was analysed by the WIMP package which identified a single open reading frame that corresponded exactly with the known amino-terminus amino acid sequence of the purified OHED hydratase protein (fig. 3.5). The identified residues were encoded between the two *Eco*RV sites with the ATG-start codon close to the most internal of these as drawn in figure 3.5. The direction of transcription of the *hpcG* gene was confirmed as being towards the *Bam*H1 site which lies approximately two-thirds of the way along the gene (fig. 3.5). Upstream of the coding sequence of the *hpcG* gene a possible ribosome binding site (RBS) was identified: the polypurine consensus sequence for a RBS is shown in figure 3.5. For efficient interaction with the ribosome, 3-9 bases of the consensus sequence should be present and these should precede the ATG-start by 3-12 bases (Steitz, 1979). In this case the possible RBS contains four bases of the consensus and is 7bp from the start codon.

Subclones of pJJ801 and pJJ200 to identify *hpc* Gene Order.

The position of *hpcG* was identified very accurately by a combination of Southern blot analysis and DNA sequencing and could therefore be used as a reference point for other genes. For example the *hpcH* gene is expressed on the

FIGURE 3.5: Interpretation of DNA sequence information: The region of DNA sequence which corresponds to the 5' region of the *hpcG* gene and the amino acid sequence of the N-terminal amino acids of OHED hydratase are shown.

FIGURE 3.5: INTERPRETATION OF THE DNA SEQUENCE INFORMATION.

EcoRV

251TGCGCGGCACCATAAGCGATCGGTCGGGATATCGGTTTCATCGAGGAAGAA
ACGCGCCGTGGTATTTCGCTAGCCAGCCCTATAGCCTTGTAGCTCCTTCTT

301CATGTCATCCAGCAGCGGCCGTAATCCGGTTCGCCGATCTGCGAGCTTGC
GTACAGTAGGTCGTCGCCGCGATTAGGCCAAGCGGCTAGACGCTCGAACG

351CTGCATCGCTTTTCGAGGTCAGGCCGATTTTGTGGCCTTTCAGCGTGCGAC
GACCTAGCGAAAGCTCCAGTCCGGCTAAAACAGGCCAAAGTCGGACGCTG

401CTTCGGCGATTTTCAGGCGAACCATTACGCTGCACCGCGTAAGCGTCT
GAAGCCGCTAAAAGTCCGCTTGGGTAAGTGCACGCTGGCGCATTTCGAGA

451TCGGATGGGGATTTTCGGGGTAATCCAGGCTGGAGGCGGAATCTGTTTCGC
AGCCTACCCCTAAAGCCCCATTAGGTCCGACTACCGCGCTTAGACAAGCG
IleGluProTyrAspLeuSerIleAlaArgIleGlnGluArg

501GCTGTTTTTCTGCCTGATCGTCACGCTGGGCGATCAGGGTGTGGGTGTGT
CGACAAAAGACGGACTAGCAGTGCACCCGCTAGTCCACACCCACACA
GlnLysGluAlaGlnAspLeuArgGlnAlaIleLeuThrHisThrHis

EcoRV

551TTATCGAACATAGCGATATCCTGTGGCCATTGCCGGATGCGGTTGCGCCT
AATAGCTTGTATCGCTATAGGACACCGGTAACGGCCTACACCAACGCGGA
LysAspPheMetRBS

601TTATCCGCTACGGTGACGGTTTGGGGTTTTGTAGGGCCCGGTAAGCGAGC
AATAGGCGATGCCACTGCCAAACCCCAAACATCCCGGGCCATTGCTCG

651CACCGG
GTGGCC

POTENTIAL RIBOSOME BINDING SITE.

mRNA CONSENSUS:5' U A A G G A G G U 3'

DNA CODING STRAND:5' A C A G G A T A T C G C T A T G 3'
fMet

recombinant plasmid pJJ200 but not pJJ212 and must be encoded to the left of the *hpcG* gene as drawn on the restriction map (fig 1.11a). Subclones of the genomic clones pJJ801 and pJJ200 were made in order to distinguish the sequence of the other genes. As the main aim of these experiments was to identify gene order, rather than gain information about the regulation of those genes, some of the subclones were constructed using pUC vectors so that the expression of all genes present on the subclone could be achieved from the *lac* promoter in the absence of their normal promoter-operator elements. These experiments were carried out using *E. coli* 5K as the host strain.

The *Bam*H1-*Eco*R1 fragment which extends from 2.9kbp to 6.3kbp on the restriction map was isolated from pJJ200 by appropriate endonuclease digestion and recovery from a low melting point agarose gel. The fragment was ligated into appropriately cut and phosphatase treated pUC18 and pUC19 and recombinants selected by inactivation of β -galactosidase activity; the pUC18 derivative was named pTF820 and the pUC19 clone, pTF920. These recombinants were transformed into *E. coli* 5K and the genes expressed were analysed after growth in minimal media utilising glycerol as the carbon source. The data shown in Table 3.1 indicates that *hpcBD* were expressed from this 3.4kbp fragment of DNA. The presence of active CHM isomerase activity, encoded by *hpcD*, but no CHMS dehydrogenase activity, encoded by *hpcC*, was unexpected as Jenkins (1987) describes the gene sequence as *hpcHBCDEF* and this fragment of DNA as encoding active *hpcGB* genes only. The *Sal*I-*Bam*H1 fragment, taken from pJJ801, which extends from 6.7kbp to 10.2kbp on the restriction map (fig 1.10) was subcloned into pUC18 and pUC19 in a similar way and the enzymes encoded tested in 5K (Table 3.1). The *hpcEF* gene products were detected, but neither the *hpcC* or *hpcD* genes were present intact on this DNA (D. Roper, personal communication). Failure to detect the CHMS dehydrogenase activity on this subclone suggests

TABLE 3.1: ACTIVITIES OF *E. coli* 5K HARBOURING VARIOUS SUBCLONES OF pJJ801.

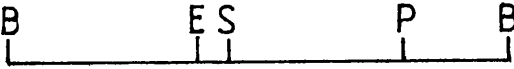







STRAIN	GROWTH	HPC	Specific Activity ($\mu\text{mol min}^{-1} \text{mg}^{-1} \text{protein}$)			
			CHMS	CHM	COHED	HHDD
	SUBSTRATE	DIOXYGENASE	DEHYDROGENASE	ISOMERASE	DECARBOXYLASE	ISOMERASE
5K (pTF820)	GLYCEROL	2.57	n.d.	1.70	N.T.	N.T.
5K (pTF920)	GLYCEROL	0.012	n.d.	0.011	N.T.	N.T.
5K (pTF802)	GLYCEROL	0.960	n.d.	1.120	N.T.	N.T.
5K (pTF902)	GLYCEROL	0.004	n.d.	0.011	N.T.	N.T.
x 5K (pDR1835)	GLYCEROL	N.T.	N.T.	n.d.	5.00	5.91
x 5K (pDR1935)	GLYCEROL	N.T.	N.T.	n.d.	n.d.	n.d.
5K (pTF303)	GLYCEROL	0.140	n.d.	0.160	N.T.	N.T.

n.d.=not detected

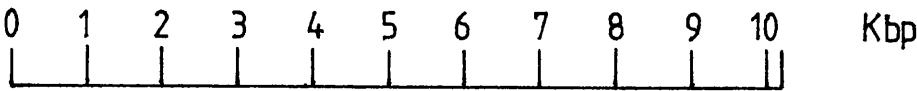
N.T.=not tested

x Information kindly supplied by D. Roper.

DIAGRAMS OF THE PLASMIDS REPRESENTED IN THIS TABLE 3.1.

PLASMID	VECTOR	
pJJ801	pBR328	
pTF820	pUC18	
pTF920	pUC19	
pTF802	pUC18	
pTF902	pUC19	
pDR1835*	pUC18	
pDR1935*	pUC19	
pTF303	pBR322	

The small arrows indicate the direction of transcription of the *lac* promoter in each case.



that the *hpcC* gene spans both the *EcoR*I and *Sa*I sites at 6.3kbp and 6.7kbp on the restriction map. The data presented here suggests an altered gene sequence [*hpcHG(BDX(EF))*] to that offered by Jenkins, but no distinction in the order of *hpc(BD)* or *hpc(EF)* was possible from these experiments. These were distinguished in a different way and will be described later.

The gene order *hpc(BDX)* was confirmed using an independent subclone from pJJ801; the *Bam*HI-*EcoR*I fragment from this genomic clone, which maps between 3.9kbp and 6.3kbp (fig. 1.10), was ligated into pUC18 and pUC19 to produce pTF802 and pTF902. Expression of the genes encoded by these constructs showed that *hpcBD* were present intact, but the gene product of *hpcC* was not detected (Table 3.1). Approximately 600bp of this subclone were known to be part of the *hpcG* gene; in an attempt to delineate the *hpc(BD)* genes further a subclone was constructed with the *Nde*I-*EcoR*I fragment which maps at 4.8kbp to 6.3kbp. This fragment was recovered from a low melting point agarose gel and ligated into appropriately cut and phosphatase treated pBR322 to produce pTF303. The products of both the *hpcB* and *hpcD* genes could be detected in cell extracts of *E. coli* 5K(pTF303) when grown on glycerol (Table 3.1) indicating that the *hpc(BD)* genes are present on approximately 1.5kbp of DNA which lie between 4.8kbp and 6.3kbp on the restriction map.

In each case where the same DNA fragment was ligated into the multiple cloning site of pUC18 and pUC19, greatly differing results were obtained for enzyme specific activities (Table 3.1). Cloning of a DNA fragment which has heterologous cohesive ends into the appropriate sites in these vectors places the genes encoded by the insert in both orientations relative to the *lac* operator-promoter region and thus when transcription from the *lac* promoter is in the same direction as the genes on the insert an increase in enzyme activity may result. Conversely, when transcription from the *lac* promoter is counter to the direction of transcription of the genes on the insert a reduction in associated enzyme activities may be expected. This theme was

investigated to determine the direction of transcription of the *hpc* genes and is described in more detail in section 3.2.

Localisation of the 5' end of the *hpcC* Gene.

During this work the product of the *hpcC* gene, CHMS dehydrogenase, was purified and the amino terminal amino acid sequence determined by automated techniques. A detailed description of these procedures is given in section 4.1 and is noted here since a degenerate oligonucleotide probe made from the protein sequence was used to identify the 5' end of the *hpcC* gene.

A total of 34 amino acid residues were assigned during sequencing of the amino-terminus of the CHMS dehydrogenase of *E. coli* C (fig. 3.6a). Residues 5-10 were chosen as a suitable region from which to make the corresponding degenerate oligonucleotide probe. The amino acids in this region are Asn-His-Trp-Ile-Asn-Gly, which have degeneracies of 2, 2, 1, 3, 2 and 4 respectively; utilising the first two bases only of the Gly, a mixture of 24 different 17 mer oligonucleotides was synthesised (fig. 3.6b). This mixed oligonucleotide probe, probe C, was radiolabelled using [γ - 32 P]-dATP and T4 polynucleotide kinase and used to detect the 5' end of the *hpcC* gene in a Southern blot experiment. The recombinant plasmid pJJ801 was digested with combinations of restriction enzymes, electrophoresed through a 0.8% agarose gel and then transferred to a nitrocellulose membrane; the hybridisation was carried out under standard conditions previously described. Autoradiography of the blot indicated that the 5' end of this gene was present on the 3.5kbp *SalI*-*Bam*HI fragment which lies between 6.7 and 10.2kbp on the restriction map (fig 3.7). As no CHMS dehydrogenase activity could be detected in extracts of 5K harbouring either the *SalI*-*Eco*R1 or *Bam*HI-*Eco*R1 fragments of pJJ801 (Table 3.1), the *hpcC* gene must lie across the *SalI* and *Eco*R1 sites with the 5' end to the right of the *SalI* site as drawn.

FIGURE 3.6a: SEQUENCE OF THE FIRST 34 RESIDUES FROM THE AMINO TERMINUS OF
CHMS DEHYDROGENASE.

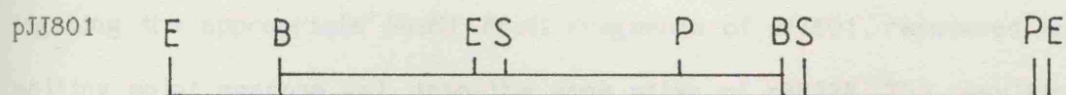
Residue Number:	1	2	3	4	5	6	7	8	9	10
Residue:	Met	Lys	Lys	Val	Asn	His	Trp	Ile	Leu	Gly
Residue Number:	11	12	13	14	15	16	17	18	19	20
Residue:	Lys	Asn	Val	Ala	Gly	Asn	Asp	Tyr	Phe	Gln
Residue Number:	21	22	23	24	25	26	27	28	29	30
Residue:	Thr	Thr	Asn	Pro	Ala	Thr	Gly	Glu	Val	Leu
Residue Number:	31	32	33	34						
Residue:	Ala	Asp	Val	Ala						

FIGURE 3.6b: SEQUENCE OF THE SYNTHETIC OLIGONUCLEOTIDES CORRESPONDING TO
RESIDUES 5-10 OF THE AMINO TERMINUS OF CHMS DEHYDROGENASE.

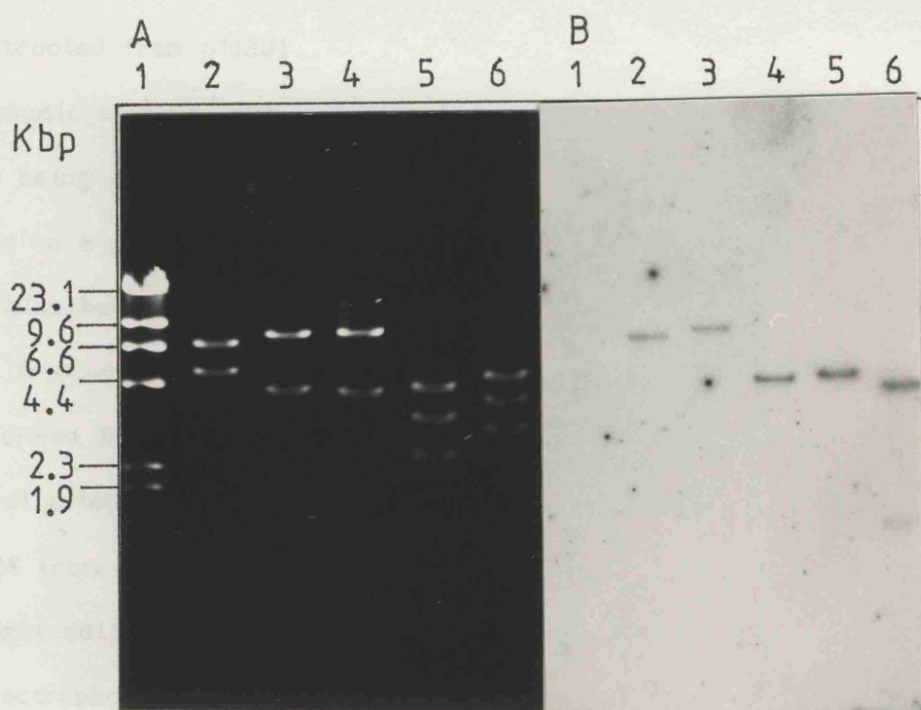
Residue Number:	5	6	7	8	9	10
Residue:	Asn	His	Trp	Ile	Leu	Gly
Probe C:	TTA	GTA	ACC	TAA	TTA	CC
	G	G		G	G	
				T		

FIGURE 3.7: Southern blot analysis of pJJ801 with probe C: Panel A shows gel electrophoresis of restriction fragments of pJJ801 and panel B shows the resulting Southern blot after transfer of the DNA to nitrocellulose and hybridisation with radiolabelled probe C. Lane 1 contained *Hind*III restricted λ -DNA and lanes 2-6 contained pJJ801 digested with: 2, *Bam*H1; 3, *Eco*R1, 4, *Sal*I; 5, *Bam*H1/*Eco*R1 and 6, *Bam*H1/*Sal*I. Exposure of the filter to X-ray film was for two hours at room temperature. Analysis of the blot shows that the hybridisations occur to the right of the insert *Sal*I site (as drawn).

FIGURE 3.7 : SOUTHERN BLOT ANALYSIS OF pJJ801 WITH PROBE C.



0 1 2 3 Kbp



Location of the *hpcF* Gene.

Jenkins (1987) reported that the gene encoding the previously unrecorded HHDD isomerase activity, *hpcF*, could be located on pJJ801 by analysis of the subclones pJJ005 and pJJ006 (fig 3.8). These derivatives were constructed by ligating the appropriate *Bam*H1-*Pvu*II fragments of pJJ801, recovered from a low melting point agarose gel, into the same sites of pBR328. The resulting clone pJJ005 was reported as containing no active genes while the clone pJJ006 was reported to express *hpcBCDE* but not *hpcF* (fig 1.11b). The inference made from these data was that the *Pvu*II site which formed the junction of the two subclones was in the coding sequence of the *hpcF* gene and thus neither subclone expressed active HHDD isomerase. Since the original subclones were no longer available verification of this data was not possible: the subclones were reconstructed from pJJ801.

The genomic clone pJJ801 was digested with a combination of *Bam*H1 and *Pvu*II before being religated (fig.3.8). In this way the insert DNA of pJJ801 would be excised on a *Bam*H1 fragment and cleaved at the single *Pvu*II site; the vector would also be cleaved at the *Pvu*II site, leaving two vector fragments and two insert fragments with *Bam*H1 and *Pvu*II cohesive ends. The ligation mixture was transformed into *E. coli* 5K and Ap-resistant, Tc-sensitive clones identified so that religated vector molecules could be eliminated from further investigation. The DNA from a selection of these transformants was prepared from 1.5 ml overnight cultures and analysed by restriction enzyme digestion and agarose gel electrophoresis. Those which contained a 1.3kbp *Pvu*II-*Bam*H1 fragment, equivalent to pJJ005, were named pTF305 and those which contained a 5.2kbp fragment, equivalent to pJJ006, were called pTF306. The enzyme activities associated with these subclones were analysed in strain 5K using glycerol as the carbon source and are presented in Table 3.2. As with pJJ005, pTF305 contained no detectable enzyme activities, however, pTF306 expressed *hpcF* as well as *hpcBCDE*.

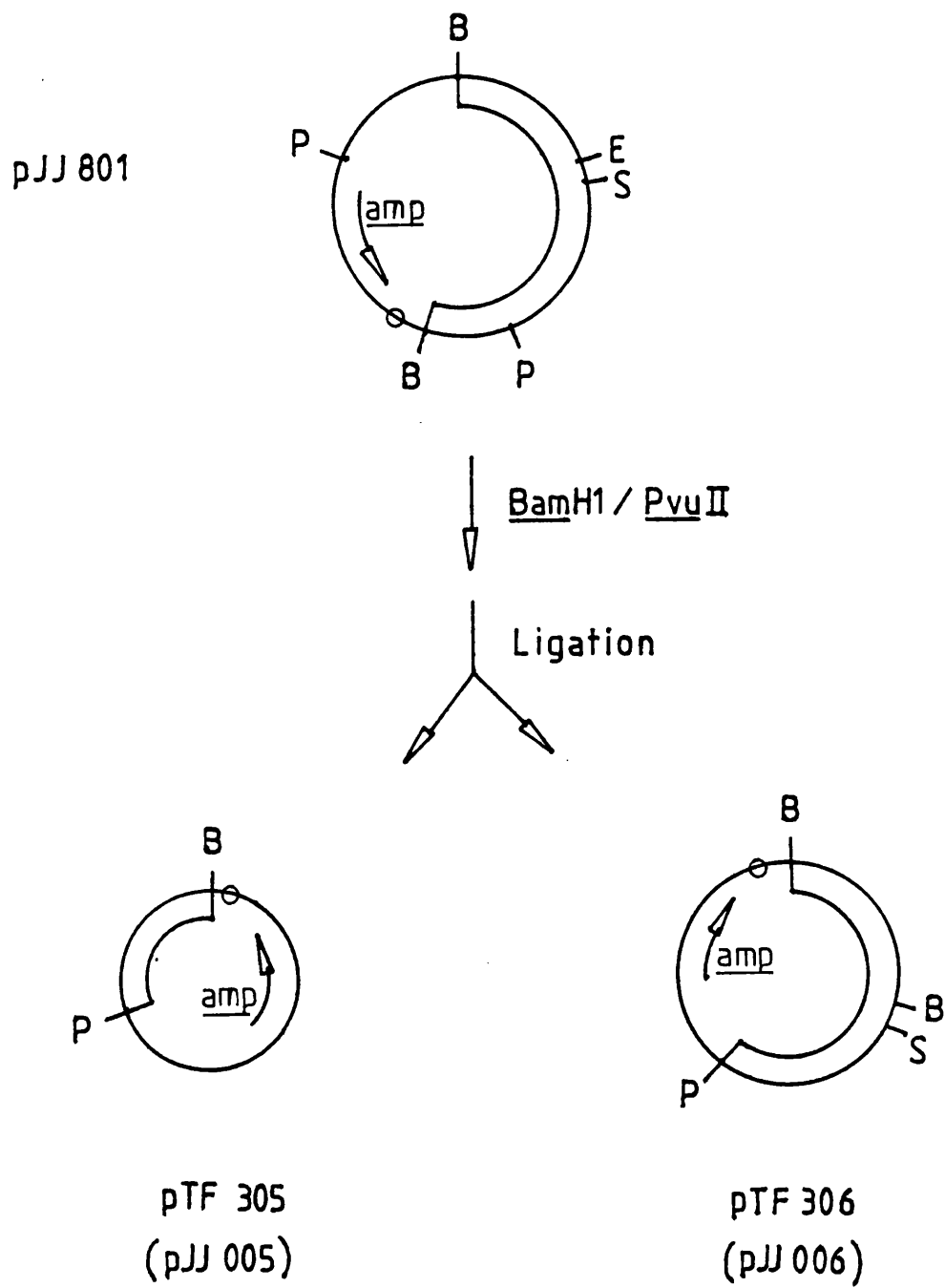
TABLE 3.2: ENZYME ACTIVITIES WITH SUBCLONES OF pJJ801 WHEN HARBOURED IN 5K.

STRAIN	GROWTH	HPC	Specific Activity ($\mu\text{mol min}^{-1} \text{mg}^{-1} \text{protein}$)				
			CHMS	CHM	COHD	HHDD	
	SUBSTRATE	DIOXYGENASE	DEHYDROGENASE	ISOMERASE	DECARBOXYLASE	ISOMERASE	
5K (pJJ801)	GLYCEROL	0.875	0.400	0.520	0.090	0.650	
5K (pTF306)	GLYCEROL	1.010	0.490	0.370	0.140	0.610	
5K (pTF302)	GLYCEROL	0.190	n.d.	0.160	N.T.	N.T.	

n.d.=not detected

N.T.=not tested

FIGURE 3.8: CONSTRUCTION OF THE SUBCLONES pTF305 AND pTF306.



○ = Origin of Replication

One possible explanation for the discrepancy in recorded enzyme activities between pJJ006 and pTF306 was that two *PvuII* sites lay close together, one being in the coding region of the *hpcF* gene and the other outside of that region. In constructing a subclone by cutting and re-ligation, as in pTF306, it was possible that any such *PvuII*-*PvuII* fragment would become ligated back into the correct position to restore an intact *hpcF* gene. This situation is much less likely to arise when fragments of DNA are recovered from gel slices for the construction of subclones as in pJJ006. To test this possibility pTF306 was restricted with an excess of *PvuII* to ensure complete digestion and electrophoresed through a 1% low melting point gel. A single band was seen on the gel; this was extracted, the DNA precipitated and re-ligated. In this way any small *PvuII*-*PvuII* fragment would be well separated from the 8.8kbp band. The ligation mixture was transformed into *E.coli* 5K; six transformants were grown overnight in LB-Ap and each was tested for the presence of HHDD isomerase activity. All of the transformants tested contained an active *hpcF* gene. DNA from a 1.5 ml portions of these cultures was analysed by restriction with *PvuII* and agarose gel electrophoresis and was indistinguishable from similarly treated pTF306. These data show that the *hpcBCDEF* genes are present on the 5.2kbp *BamHI*-*PvuII* fragment which lies between 3.9kbp and 8.9kbp on the restriction map (fig. 1.10)

Ordering the *hpc(EF)* Genes.

Jenkins (1987) ordered the *hpcEF* genes by analysis of the subclones pJJ005 and pJJ006 as previously discussed. However it is now known that these constructs do not give any information about the relative position of the *hpcEF* genes and the smallest region to which they can be localised is the 2.1 kbp *SalI*-*PvuII* fragment that lies between 6.8 and 8.9kbp. Due to the lack of known restriction enzyme sites in this region of DNA it was not possible to construct a subclone

which contained only one of these genes in the usual way, therefore, *Ba*31 exonuclease was used to construct a deletion in this region.

The primary clone pJJ801 contains a single *Sst*II restriction enzyme site approximately 100bp from the *Sal*I site, towards the *Pvu*II site (D. Roper, personal communication). As the vector DNA contains no such site, cleavage with *Sst*II produces a single linear molecular species such that one exposed end is close to the *hpc*(*EF*) gene. Four micrograms of *Sst*II linearised pJJ801 were resuspended in 50 μ l of 500 μ g ml⁻¹ BSA and 50 μ l 2x *Ba*1 31 buffer. After 30 min incubation at 30°C, 0.5 units of *Ba*31 were added and 25 μ l aliquots removed after of 1, 5, 15 and 30 min incubation. The reaction was stopped by immediate extraction with phenol/chloroform and the DNA precipitated with ethanol. After checking by agarose gel electrophoresis that some of the DNA had been deleted, the molecules were made blunt-ended by the action of Klenow fragment of DNA polymerase I in the presence of a mixture of four deoxynucleotide triphosphates (dNTP's) in the following reaction.

DNA (approximately 1 μ g)	x μ l
10x NT buffer	2 μ l
1.25 mM dNTP's	2 μ l
Klenow (3 units)	0.5 μ l
water (to 25 μ l)	

The reaction was allowed to proceed at room temperature for 30 min before being stopped by the addition of EDTA to a final concentration of 20 mM. The mixture was extracted with phenol/chloroform, the DNA precipitated with ethanol and resuspended in water. The DNA was ligated overnight at room temperature and then transformed into the strain *E. coli* C JJ221. Transformants were selected on LB-Ap containing agar plates and those which appeared after overnight incubation at 37°C tested for their ability to grow on minimal media using HPA as the sole carbon source. DNA was prepared from twelve

transformants which showed the phenotype of growth on LB-Ap, but no growth on HPA and this was transformed into the strain 5K for the analysis of enzyme activities still encoded on the DNA. Each of the freshly transformed derivatives of 5K were grown overnight in liquid LB-Ap cultures, cell free extracts prepared and enzyme assays carried out (Table 3.3). Of the 12 clones analysed the phenotypes, COHED decarboxylase+/HHDD isomerase+ or COHED decarboxylase-/HHDD isomerase- were seen, but none which distinguished the order of the genes encoding these proteins. This is surprising, especially as three clones only expressed these two genes, indicating that *hpcBCD* had been deleted and that the exonuclease activity had proceeded preferentially in one direction. This assumption is supported by the enzyme restriction digests of these clones: those which express only *hpcEF*, when digested with *Bam*H1, show only a single band indicating the removal of one *Bam*H1 site. Analysis of more deletion clones of this type should be able to distinguish the order of the *hpc(EF)* genes, but this has not proved possible to date.

Ordering the *hpc(BD)* Genes.

The ordering of the *hpc(BD)* genes is discussed in section 3.2.

TABLE 3.3: ENZYME ACTIVITIES DETECTED IN *Ba*31 DELETION SUBCLONES.

SUBCLONE	<i>hpc</i> GENES EXPRESSED				
	<i>hpcB</i>	<i>hpcC</i>	<i>hpcD</i>	<i>hpcE</i>	<i>hpcF</i>
1	-	-	-	+	+
2	-	-	-	+	+
3	+	-	+	+	+
4	+	-	+	+	+
5	-	-	-	+	+
6	+	-	+	+	+
7	N.T.	N.T.	N.T.	+	+
8	N.T.	N.T.	N.T.	+	+
9	N.T.	N.T.	N.T.	+	+
10	N.T.	N.T.	N.T.	+	+
11	N.T.	N.T.	N.T.	+	+
12	N.T.	N.T.	N.T.	+	+

N.T.=not tested.

3.2 EXPRESSION OF THE *hpc* GENES.

Direction of Transcription of the *hpc* Genes.

The direction of transcription of *hpcG* was determined by a combination of Southern blot analysis and DNA sequencing (Section 3.1) as being from right to left as drawn on the restriction map from the 4.5kbp coordinate. This provided a good reference point for the ordering of the other genes as well as the direction of transcription of those genes.

Transcription of *hpcG(BD)* Genes.

The model of transcription proposed by Jenkins (1987) suggests that a divergent promoter region for the transcription of two gene blocks, *hpcGH* and *hpcBCDEF*, was situated between the *hpcB* and *hpcG* genes. In order to test this model a DNA fragment which spanned this region was inserted into the multiple cloning site of pUC18 and pUC19. The fragment chosen was the 3.4kbp *Bam*H1-*Eco*R1 piece of pJJ200 which maps between 2.9 and 6.3kbp; heterologous cohesive ends allow this DNA to be inserted into the pair of pUC vectors in opposite orientations compared to the *lac* promoter of the plasmid. If the model of Jenkins was correct then a gene with transcription concurrent with that of the *lac* promoter may be expected to have an increased number of transcripts which could be translated into active proteins and thus the end result would be an increase in that enzymes' specific activity. Similarly a gene transcribed counter to that of the *lac* promoter would suffer a reduction in the number of viable transcripts through either collision of RNA polymerases travelling in opposite directions (Brewer, 1988) or the binding of sense and anti-sense mRNA's (Coleman *et al.*, 1984). The proposal of Jenkins would in this case increase the level of OHED hydratase and reduce the level of HPC dioxygenase (and CHM isomerase) in the pUC18 construct; the reverse being true in pUC19. However, when present in *E. coli* 5K, the construct pTF820 expressed the *hpcGBD*

genes at levels approximately 200 times greater than the pTF920 construct (Table 3.1), indicating that all three genes are transcribed in the same direction. The orientation of the *lac* promoter in the pTF820 construct is concurrent with the known orientation of the *hpcG* gene. Thus the *hpcGBD* genes are transcribed towards the *Bam*H1 end of this fragment. Addition of an inducer of the *lac* promoter, IPTG, did not affect the recorded enzyme specific activities when added to a final concentration of 1mM (Table 3.4). As *E.coli* 5K contains only a single copy of the *lac* repressor gene, *lacI*, it appears that the number of repressor protein molecules produced was insufficient to significantly interfere with transcription, allowing near-constitutive levels of transcription in the absence of IPTG.

The direction of transcription of *hpc(BD)* was confirmed by an independent construction in pUC18 and pUC19 using the 2.4kbp *Bam*H1-*Eco*R1 fragment from pJJ801, which lies between 3.9kbp and 6.3kbp on the restriction map. In a similar experiment to the one just described the pUC18 construct, pTF802, expressed both *hpcB* and *hpcD* at levels about 100 times higher than the equivalent pUC19 construct, pTF902 (Table 3.1). This result confirms that the 5' ends of the *hpcB* and *D* genes are closer to the *Eco*R1 site of these constructs than they are to the *Bam*H1 site. This experimental concept was carried further by the insertion of the whole 6.5kbp *Bam*H1 fragment of pJJ801 into pUC18 in opposite orientations. The resulting recombinant plasmids pTF882 and pTF883 were assayed for enzyme specific activities after growth on glycerol containing media in the presence or absence of 1mM IPTG (Table 3.4). The activities of *hpcBCDEF* were either very low or not detectable in pTF882, but were all present at a high specific activity in pTF883 indicating that all of these genes are expressed in the same direction. Again the enzyme activities in each construct were not significantly affected by the presence of IPTG.

These data contradict the model of Jenkins which proposes that transcription of

**TABLE 3.4: ACTIVITIES OF HPC CATABOLIC ENZYMES OF VARIOUS SUBCLONES IN pUC VECTORS:
THE EFFECT OF IPTG.**

STRAIN	GROWTH	HPC	Specific Activity ($\mu\text{mol min}^{-1} \text{mg}^{-1} \text{protein}$)			
			CHMS	CHM	COHED	HHDD
	SUBSTRATE	DIOXYGENASE	DEHYDROGENASE	ISOMERASE	DECARBOXYLASE	ISOMERASE
5K (pTF820)	GLY	2.57	n.d.	1.70	N.T.	2.51
5K (pTF820)	GLY/IPTG	2.79	n.d.	1.44	N.T.	2.46
5K (pTF920)	GLY	0.012	n.d.	0.011	N.T.	0.033
5K (pTF920)	GLY/IPTG	0.010	n.d.	0.009	N.T.	0.033
5K (pTF802)	GLY	0.960	n.d.	1.120	N.T.	N.T.
5K (pTF802)	GLY/IPTG	1.110	n.d.	1.050	N.T.	N.T.
5K (pTF902)	GLY	0.004	n.d.	0.011	N.T.	N.T.
5K (pTF902)	GLY/IPTG	0.003	n.d.	0.009	N.T.	N.T.
5K (pTF882)	GLY	n.d.	n.d.	n.d.	0.002	0.015
5K (pTF882)	GLY/IPTG	n.d.	0.001	0.003	0.003	0.017
5K (pTF883)	GLY	2.50	1.14	1.28	4.06	2.30
5K (pTF883)	GLY/IPTG	2.57	1.25	1.57	3.56	2.14

the genes present on pJJ801 is in the order *hpcBCDEF*. This evidence indicates that the order of transcription is *hpc(FE)C(DB)*.

Evidence for Two Transcriptional Units on pJJ801.

The primary clone pJJ801 when present in *E.coli* 5K expresses *hpcBCDEF* constitutively, but these genes are regulated in an *E. coli* C background (Jenkins, 1987; Table 3.5). From this it appears that the plasmid pJJ801 contains a promoter region(s) for the initiation of transcription of the *hpc* genes and also an operator region to which a chromosomally encoded repressor protein can act in trans, regulating that expression. Analysis of the behaviour of subclones of pJJ801 when present in *E. coli* C JJ221 and *E. coli* 5K suggest that two transcriptional units are present and this evidence is discussed here. The subclone pTF306, which contains the 5.2kbp *Bam*HI-*Pvu*II fragment of pJJ801, expresses *hpc(FE)C(DB)* in 5K. (Table 3.5). When present in JJ221 and grown on glycerol the subclone pTF306 expresses *hpc(FE)CD* constitutively; *hpcB* is also expressed but at a low rate (Table 3.5). In the presence of an inducer, HPA, the levels of enzymes encoded by *hpc(FE)CD* increase approximately 1.5-fold and the HPC dioxygenase(*hpcB*) increases by approximately nine-fold. The presence of HPA in the medium induces the chromosomally encoded *hpc* genes of JJ221 and the contribution of these genes is included in the enzyme specific activities. This experiment indicates the sites of two operator-promoter regions. All of the enzyme activities of pJJ801 are present on pTF306, but the latter expresses *hpc(FE)CD* constitutively in JJ221 showing that the *Pvu*II site falls either in a promoter region or between the promoter and the translational start signal of the first gene. HPC dioxygenase activity remains inducible when measured in JJ221(pTF306), suggesting that it is normally expressed on a different transcriptional unit to *hpc(FE)CD*. This observation was verified using

TABLE 3.5: ENZYME ACTIVITIES ASSOCIATED WITH SUBCLONES OF pJJ801 WHEN HARBOURED IN 5K AND JJ221.

STRAIN	GROWTH	HPC	Specific Activity ($\mu\text{mol min}^{-1} \text{mg}^{-1} \text{protein}$)			
			CHMS	CHM	COHD	HHDD
	SUBSTRATE	DIOXYGENASE	DEHYDROGENASE	ISOMERASE	DECARBOXYLASE	ISOMERASE
5K (pJJ801)	GLY	0.875	0.400	0.520	0.090	0.650
JJ221 (pJJ801)	GLY	n.d.	0.001	0.002	0.007	0.004
JJ221 (pJJ801)	GLY/HPA	0.190	0.170	0.110	0.060	N.T.
5K (pTF306)	GLY	1.010	0.490	0.370	0.140	0.610
JJ221 (pTF306)	GLY	0.049	0.230	0.280	0.055	0.110
JJ221 (pTF306)	GLY/HPA	0.445	0.117	0.425	0.090	0.190
5K (pTF302)	GLY	0.190	n.d.	0.160	N.T.	N.T.
JJ221 (pTF302)	GLY	0.008	n.d.	0.175	N.T.	N.T.
JJ221 (pTF302)	GLY/HPA	0.360	n.d.	0.285	N.T.	N.T.

a different subclone in a similar experiment. The 2.4kbp *Bam*H1-*Eco*R1 fragment of pJJ801 was subcloned into pBR322; the resulting construct, pTF302, expresses *hpcBD* in 5K (Table 3.5). When harboured by JJ221 the *hpcD* gene is expressed constitutively and the *hpcB* gene is expressed in a regulated manner (Table 3.5). Thus a second operator-promoter region must lie 5' to *hpcB*, between the *hpcB* and *hpcD* genes.

As *hpcD* has been shown to be part of the *hpc(FE)CD* transcriptional unit the complete gene order must be *hpcHGB hpcDC(EF)*.

Northern Blot Analysis of *hpc* Transcripts.

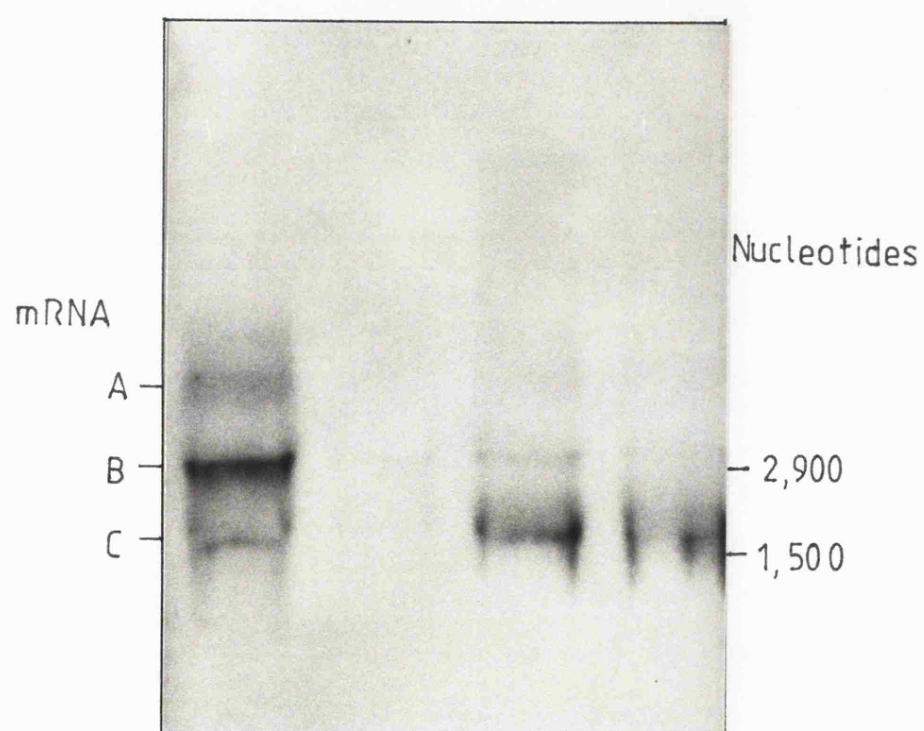
Two transcriptional units had been identified on pJJ801: one containing the *hpc(FE)CD* genes and the other starting 5' to the *hpcB* gene. It is not possible to determine from previous experiments if the structural genes *hpcGH* are a part of the same unit as *hpcB* or if they are transcribed together or singly on separate units. The regulatory gene *hpcR* may be transcribed along with some structural genes as in the negatively autoregulating *hut* operon or may be present on a completely separate unit.

To determine the number and approximate size of *hpc* transcripts the total RNA extracted from a number of exponentially growing cultures was electrophoresed through a 1% agarose gel containing formaldehyde. The RNA was transferred to a nitrocellulose membrane without prior treatment of the gel and the filter baked to immobilise the nucleic acid. Hybridisation with the whole *Bam*H1 insert fragment of pJJ200, radiolabelled along its length using [α - 32 P] dCTP and a random hexanucleotide primer, indicated the presence of three transcripts (fig. 3.9). Using the 23s and 16s rRNA bands as size markers three transcripts which hybridised to the probe in the *E. coli* C (HPA) lane were calculated as being approximately 4.5, 2.7 and 1.6 kb in size. In a sample of total RNA taken from

E.coli C cells grown on glucose and treated in the same way, the 2.7 and 1.6 kb bands were just visible but the 4.5 kb band was not (fig. 3.9).

FIGURE 3.9: Northern blot analysis: Autoradiograph of a Northern blot probed with the whole of pJJ200 insert DNA. The lanes contained RNA (50µg) from *E.coli* C grown on HPA (lane 1); glucose (lane 2) and from *E. coli* 5K (pJJ200) grown on glycerol in the presence of HPA (lane 3) and on glycerol (lane 4). The blot was exposed to film for 3 days at -70°C.

FIGURE 3.9: NORTHERN BLOT ANALYSIS OF TOTAL RNA FROM *E. coli* C AND *E. coli* 5K(pJJ200) GROWN IN THE PRESENCE AND ABSENCE OF HPA.



A sample of RNA from *E. coli* 5K harbouring the clone pJJ200 was also run on the same gel to compare the transcripts associated with the cloned material with those of the wild type. Unfortunately a direct comparison was not possible as *E. coli* 5K(pJJ200) is not able to utilise HPA or HPC as the sole carbon source. Therefore, these cells were grown using glycerol as the carbon source in the presence of 4-HPA: RNA extracted from 5K(pJJ200) cells grown on glycerol only was used as the control (fig. 3.9).

Laser densitometry of the autoradiograph which resulted from this experiment clearly shows the three peaks of hybridisation in the *E. coli* C sample when grown on 4-HPA (fig. 3.10), although the largest transcript (mRNA A) is diffuse. The scan of *E. coli* C when grown on glucose shows no obvious peaks. In the case of *E. coli* 5K(pJJ200), cells grown on glycerol contain predominantly mRNA C but those grown in the presence of 4-HPA also contain mRNA A and mRNA B. The amount of hybridisation in each track of the gel and the distribution between the three peaks is given in Table 3.6. The total intensity figures are a direct indication of the amounts of mRNA transcribed from the *hpc* genes and this varied in each case: *E. coli* C grown on 4-HPA contained approximately three times more hybridising material than *E. coli* 5K(pJJ200) grown on glycerol in the presence of 4-HPA. This discrepancy can be accounted for by a combination of some catabolite repression being exerted by the glycerol in the growth medium and the small number of copies of the recombinant plasmid at the time of harvesting the cells. This can be demonstrated by the increase in enzyme specific activities associated with an increase in the optical density of the bacterial culture of 5K(pJJ200) (Table. 3.7).

The smallest transcript, mRNA C, was produced in the absence of any inducer in 5K(pJJ200) grown on glycerol: this along with the absence of enzyme activities detected in this culture (Table 3.7) suggests that the 1.6kb transcript is that of the regulatory gene, *hpcR*. The larger two transcripts mRNA A and mRNA B are only present when an inducer was in the growth medium and these are probably

FIGURE 3.10: The Northern blot in Figure 3.9 was scanned at 633nm in an LKB densitometer. The lanes correspond to those on the gel: lane 1, *E. coli* C/HPA; lane 2, *E. coli* C/glucose; lane3, *E. coli* 5K (pJJ200)/glycerol+HPA and lane 4, *E. coli* 5K(pJJ200)/glycerol.

FIGURE 3.10: LASER DENSITOMETRIC SCANS OF NORTHERN HYBRIDISATIONS.

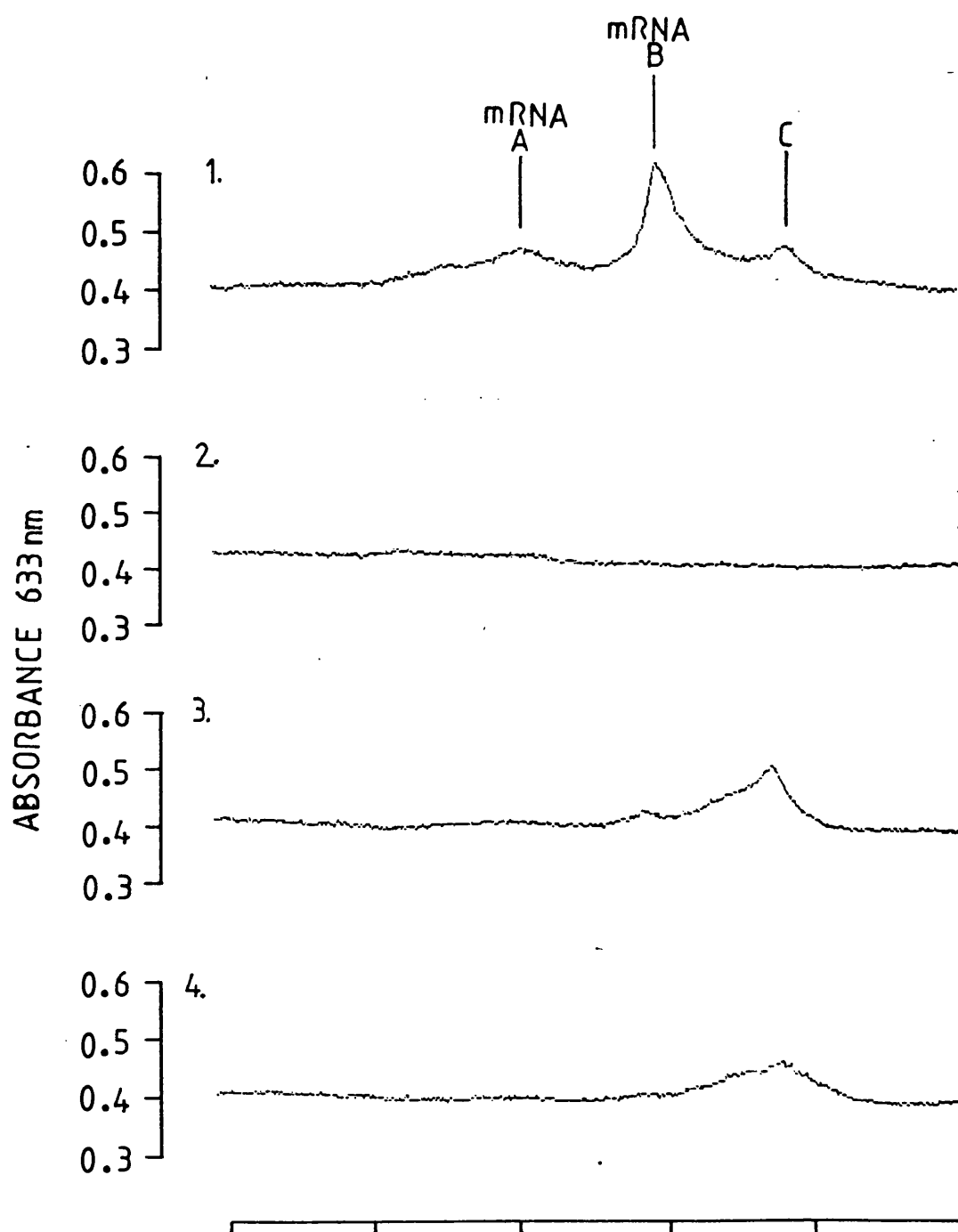


TABLE 3.6: TO SHOW THE PERCENT INTESITIES OF HYBRIDISATION OF pJJ200 INSERT
DNA TO TOTAL RNA FROM VARIOUS SOURCES.

TRANSCRIPT	SAMPLE (PERCENT INTENSITY)			
	<i>E. coli</i> C HPA	<i>E. coli</i> C GLUCOSE	5K (pJJ200) GLYCEROL	5K (pJJ200) GLYCEROL/HPA
mRNA A	40.8	0	1.6	10.8
mRNA B	45.8	0	0	20.3
mRNA C	13.3	0	98.4	67.8
TOTAL INTENSITY	2077	141	456	711

TABLE 3.7: TO SHOW THE INCREASE IN ENZYME SPECIFIC ACTIVITY WITH INCREASED OPTICAL DENSITY OF A CULTURE OF 5K (pJJ200).

		Specific Activity ($\mu\text{mol min}^{-1} \text{mg}^{-1} \text{protein}$)			
OPTICAL	HPC	CHMS	CHM	COHED	HHDD
DENSITY	DIOXYGENASE	DEHYDROGENASE	ISOMERASE	DECARBOXYLASE	ISOMERASE
(550nm)					HYDRATASE
0.3	0.003	0.003	0.020	0.019	0.014
0.5	0.009	0.036	0.040	0.024	0.017
0.9	0.033	0.042	0.085	0.036	0.030
					n.d.
					0.013
					0.042

n.d.=not detected.

the only transcripts which contain structural genes. A total of approximately 7.3 kb of mRNA is present in these two transcripts which is sufficient to encode the seven structural genes required for the conversion of HPC to succinic semialdehyde and pyruvate. A model for the organisation and expression of the *hpc* genes of *E. coli* C is given in section 3.3.

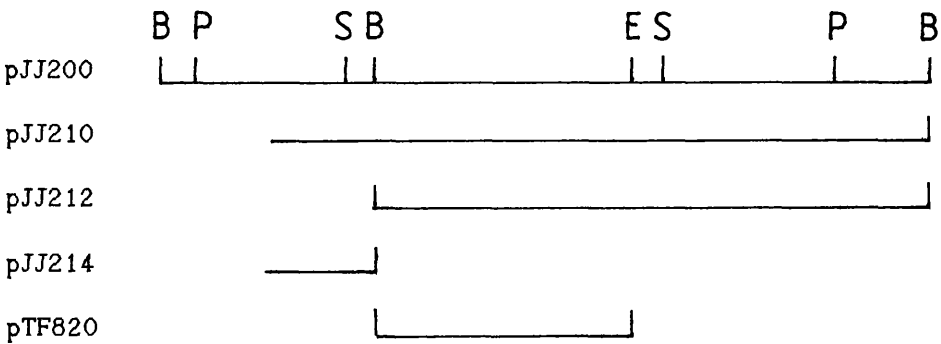
3.3: MODEL OF ORGANISATION AND EXPRESSION OF *hpc* GENES OF *E. coli* C.

The Gene Order.

A previous model for the organisation of *hpc* genes suggested that the gene order was *hpcRHGBCDEF* (Jenkins, 1987). The present study shows that the order is in fact *hpcRHGBDC(EF)*.

The primary clones pJJ200 and pJJ210 are phenotypically similar, expressing all of the genes required for the conversion of HPC to central metabolites in a regulated manner. Deletion of the region to the left of the internal *Bam*H1 site of pJJ210 produces the subclone pJJ212 which expresses *hpcBCDEFG* constitutively in *E. coli* 5K. As the fragment removed, pJJ214, does not express the *hpcH* gene it appears that the *Bam*H1 site is in the coding sequence of that gene (fig 3.11). The 5' end of the *hpcG* gene has been accurately placed 561bp from the left hand *Bam*H1 site of pJJ801 by Southern blot analysis using a degenerate oligonucleotide probe deduced from the primary amino acid sequence of the *hpcG* gene product. The *hpcG* gene product, OHED hydratase was not detectable in extracts of 5K(pJJ801) but was in 5K(pJJ212): this shows that the *hpcG* gene spans the left hand *Bam*H1 site of pJJ801 (fig.3.11). Analysis of the 2.4kbp *Bam*H1-*Eco*R1 fragment of pJJ801 in pUC18 (pTF802; Table 3.1) or pBR322 (pTF302; Table 3.2) detected the products of *hpcB* and *hpcD* but not *hpcC*. The *hpc(BD)* genes were ordered by their behaviour in the *E. coli* C strain, JJ221. The subclone pTF302 when present in JJ221 expresses *hpcD* constitutively, but *hpcB* is expressed in a regulated manner (Table 3.5). Similarly the subclone pTF306 expresses *hpcCDEF* constitutively in JJ221, but *hpcB* is regulated (Table 3.5). These data suggest that *hpcD* is expressed in the same transcriptional unit as *hpcCEF* and must therefore lie to the right of *hpcB* as drawn. The *hpcC* gene is expressed on pJJ801 but not on either the 2.4kbp *Bam*H1-*Eco*R1 fragment or the 3.5kbp *Sal*I-*Eco*R1 fragments derived from it.(Table 3.1; fig.3.11) indicating that the *hpcC* gene spans both the *Eco*R1 and *Sal*I restriction sites.

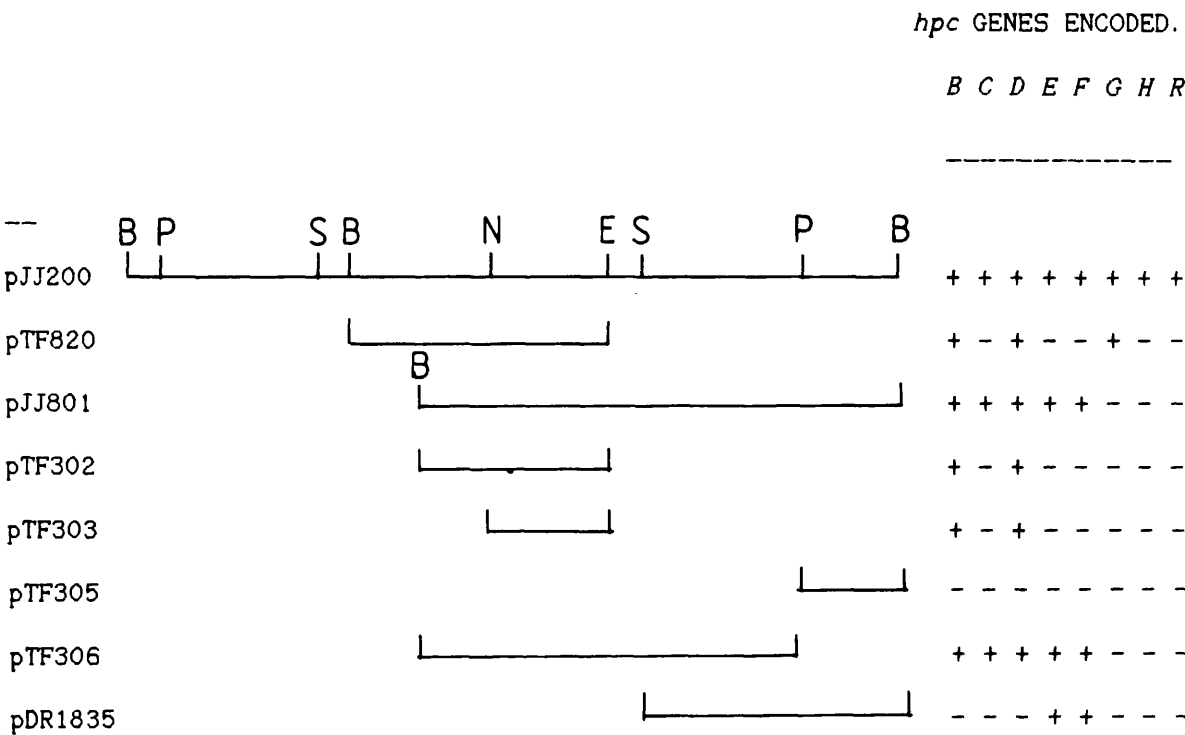
FIGURE 3.11a: SUMMARY OF CONSTRUCTS USED TO TO IDENTIFY THE *hpc* GENE ORDER.



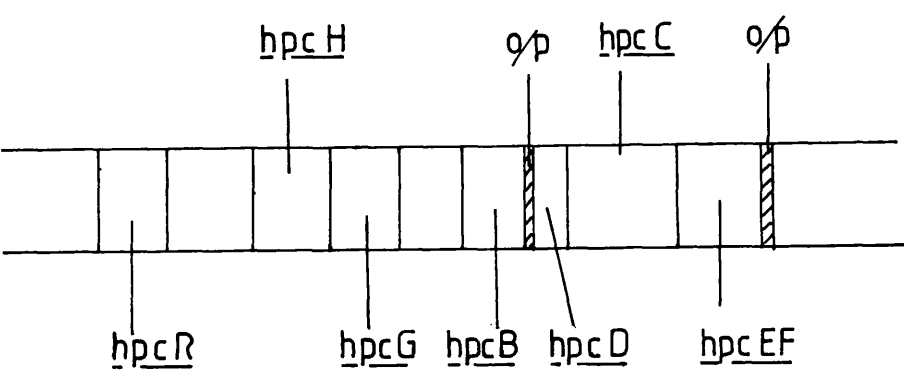
PLASMID	<i>hpc</i> GENES EXPRESSED.							INDUCED BY HPA
	<i>hpcB</i>	<i>hpcC</i>	<i>hpcD</i>	<i>hpcE</i>	<i>hpcF</i>	<i>hpcG</i>	<i>hpcH</i>	IN 5K.
pJJ200	+	+	+	+	+	+	+	I
pJJ200	+	+	+	+	+	+	+	I
pJJ212	+	+	+	+	+	+	-	C
pJJ214	-	-	-	-	-	-	-	NT
pTF820	+	-	+	NT	NT	+	NT	NT

I= inducible
C= constitutive
NT= not tested

FIGURE 3.11b: SUMMARY OF CONSTRUCTS USED TO IDENTIFY *hpc* GENE ORDER.



POSITIONS OF THE *hpc* GENES.



Expression of the *hpc* Genes.

Jenkins (1987) used the expression of *hpc* genes of various subclones in *E. coli* C strains to position an operator-promoter region (fig 1.11; fig1.12). Further, analysis of subclones produced by *Sau*3A partial digestion of the insert DNA of pJJ801 placed the operator region responsible for the control of the *hpcBCDEF* gene block within 300bp of the left hand *Bam*H1 site as drawn. This interpretation was supported by DNA sequence information obtained using M13 containing the 2.4kbp *Bam*H1-*Eco*R1 fragment of pJJ801 and specific oligonucleotide primers. Computer analysis of the DNA information identified potential -10 sites for both gene blocks and a GTG-start codon with an open reading frame for the *hpcB* gene which extended to the end of the sequence information (Jenkins, 1987). The author of this work stressed that no evidence in the form of S1 nuclease mapping or protein sequence data was available to corroborate the DNA sequence information.

Some 650bp of DNA sequence were obtained from the left hand *Bam*H1 site of pJJ801 in the present study, however, none of this corresponded with the DNA sequence information of Jenkins. The nucleotide sequence determined in this work corresponds exactly with the DNA sequence deduced from the primary amino acid data of purified OHED hydratase (figs. 3.1a; 3.5). Further no potential -10 or -35 consensus sequences were found, on either strand, within 100bp of the 5' end of the *hpcG* gene.

A subclone of pJJ801 made in this study, pTF303, which extends from 4.8-6.3kbp on the restriction map (fig 3.11) retains HPC dioxygenase activity which is the product of the *hpcB* gene (Table 3.1) Thus the coding region of the *hpcB* gene must lie to the right of the *Nde*I site as drawn in figure 3.11. Recombinant M13 clones are subject to instability and it is possible that a deletion occurred during the sequencing procedure of Jenkins resulting in a rearrangement of the insert DNA and subsequent analysis of the wrong portion of DNA.

Analysis of the clones pJJ801, pTF306 and pTF302 in *E. coli* C JJ221 (Table 3.5) indicated that two transcriptional units were present on pJJ801. Deletion of the small *PvuII*-*Bam*HI fragment of pJJ801, to produce pTF306, led to the constitutive expression of *hpcCDEF* in JJ221: as a negative regulatory model is proposed an operator region controlling the expression of *hpcCDEF* must have been removed in this construct. The expression of *hpcB* in JJ221(pTF306) was regulated, suggesting that it is present on a different transcriptional unit to *hpcCDEF* (Table 3.5). This situation was mirrored in the subclone pTF302. Total RNA from *E. coli* C grown on HPA and 5K(pJJ200) grown on glycerol in the presence of HPA, probed with the inset DNA of the primary clone pJJ200, indicated that three transcripts associated with *hpc* genes were present. The smallest of these was also expressed constitutively (in the absence of an inducer) and is, therefore, probably the transcript of the regulatory gene, *hpcR* (fig.3.9). The larger of the two structural gene transcripts is probably associated with the *hpcCDEF* genes and the smaller with *hpcBGH* genes.

Direction of Transcription.

The direction of transcription of *hpcG* was shown to be from right to left, as drawn, by DNA sequence analysis. When present in pUC18 and pUC19 the amount of detectable gene products of various subclones was always increased when the direction of transcription of the *hpcG* gene was concomitant with that of the *lac* promoter associated with the vector (Table 3.4). This showed that all of the genes of pJJ801 were expressed in the same direction and were not present on two diverging gene blocks as previously suggested (Jenkins, 1987) but on two gene blocks which are both transcribed right to left as drawn in figure 3.12.

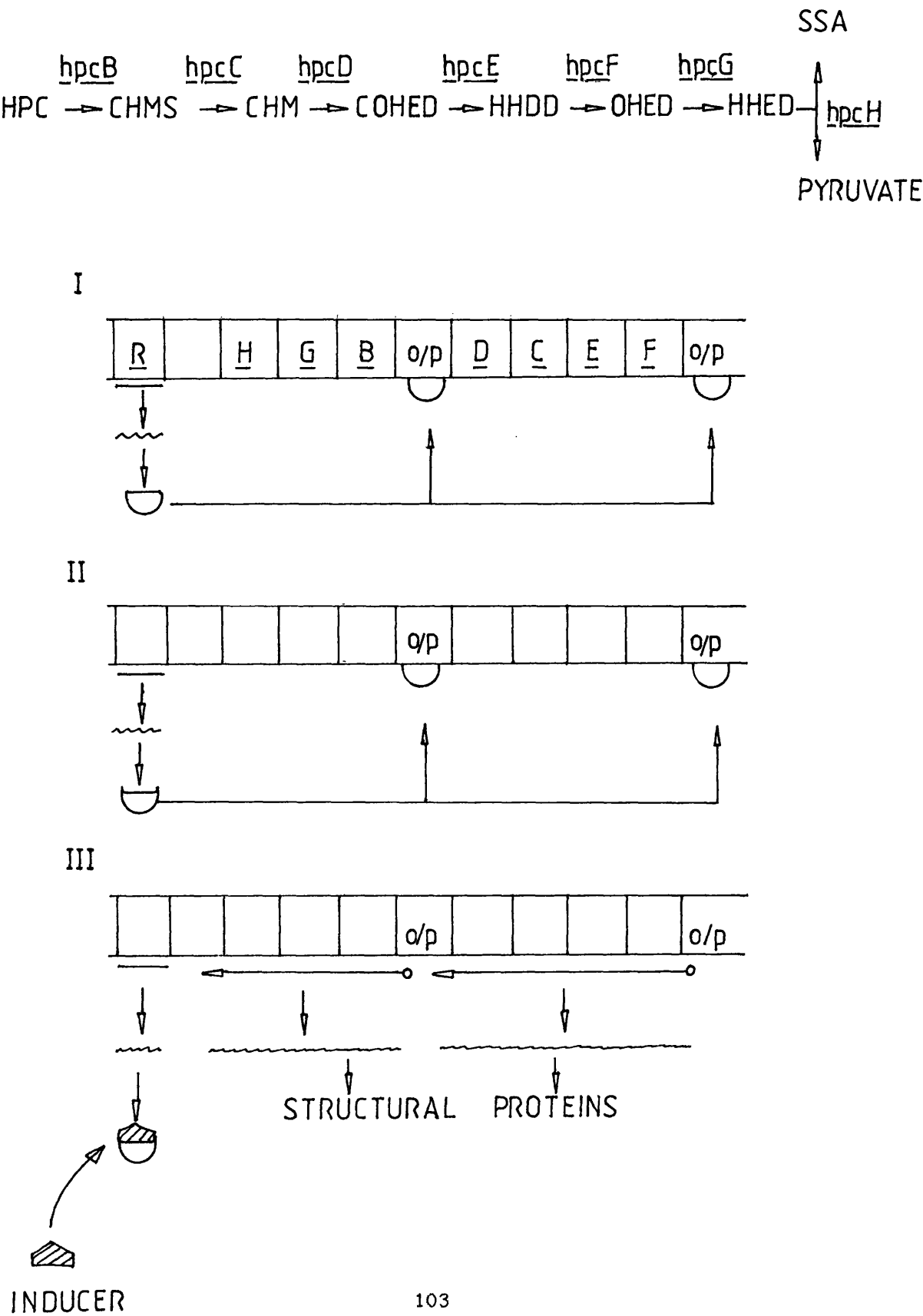
FIGURE 3.12: MODEL OF ORGANISATION:

I- In the presence of glucose in the medium: The regulatory gene, *hpcR*, is transcribed and the product of that gene binds to the two operator sites of the *hpc* operon and prevents transcription of the structural genes.

II-In the absence of glucose and the absence of an inducer of the *hpc* genes: Transcription of the *hpcR* gene and binding of the regulatory protein to the operator sites occurs as in I.

III-In the absence of glucose, but in the presence of a specific inducer of the *hpc* genes: The *hpcR* gene is transcribed and a mature regulatory protein produced. However, the inducer in some way interferes with the binding of the regulatory protein to the operator site; thus the two blocks of structural genes *hpcFECD* and *hpcBGH* are transcribed and mature proteins produced.

FIGURE 3.12: MODEL FOR THE ORGANISATION AND EXPRESSION OF *E. coli* C
hpc GENES.



4: INVESTIGATIONS INTO THE
RELATEDNESS OF *hpc* PATHWAY GENES
IN DIFFERENT BACTERIA.

4.1: Purification and Characterisation of the 5-Carboxymethyl-2-Hydroxymuconic Semialdehyde (CHMS) Dehydrogenase from *Escherichia coli* C.

Escherichia coli strains B, C and W, but not K-12 can catabolise HPA and HPC. Jenkins (1987) showed that the *Bam*HI insert DNA of the clone pJJ801 had very little homology with chromosomal DNA of *E. coli* K-12 or *P. putida* DNA by Southern hybridisation, even after washing the filter at low stringencies (0.1x SSC at 30°C) and extended exposures of the blot to X-ray film (8 days). An oligonucleotide mixture corresponding to the primary amino acid sequence of the purified OHED hydratase (Ferrer and Cooper, 1988) was used to probe chromosomal DNA from several bacteria. The probe hybridised to DNA from *E. coli* B as well as *E. coli* C, but not to DNA from *E. coli* W, K-12 or *Klebsiella pneumoniae*. The work presented here aims to investigate more systematically the evolutionary relatedness of the *hpc* genes of two organisms: *E. coli* C and *K. pneumoniae* M5a1 by examining a particular enzyme of HPC catabolism, CHMS dehydrogenase.

The source of the *E. coli* C CHMS dehydrogenase was *E. coli* 5K containing the cloned *hpcC* gene from *E. coli* C isolated by Jenkins (Jenkins, 1987; Jenkins and Cooper, 1988), carried on the recombinant plasmid pJJ212. *E. coli* 5K (pJJ212) cells from 400ml of Luria broth-Ap medium were resuspended in 16ml of 25mM Tris.Cl buffer pH 7.5 and 4ml portions of this ultrasonicated and centrifuged, as described in the Materials and Methods, to produce a crude cell extract. The crude extract was centrifuged further at 180,000g for 90 min at 4°C to remove any membraneous material. The supernatant was carefully poured off and treated with protamine sulphate (40 mg ml⁻¹ in 25mM Tris.Cl pH 7.5) in the proportion of 1mg of protamine sulphate to 20mg of bacterial protein. The extract was maintained at 0°C with gentle stirring for 15 min then centrifuged at 15,000g for 15 min at 4°C. The treatment was repeated once more, the filtered supernatant applied to an HR 10/10 Mono Q anion exchange column attached to a Pharmacia FPLC. A 100ml gradient of 0-0.6M NaCl in 25 mM Tris.Cl pH 7.5 and a

flow rate of 4ml min^{-1} was used to elute the enzyme. The CHMS dehydrogenase eluted at approximately 0.25M NaCl (fig. 4.1) and the fractions with the highest specific activity were pooled, diluted with an equal volume of 2M ammonium sulphate, filtered and applied to a HR 5/5 Phenyl Superose hydrophobic interaction column. A 15ml gradient of $1.7\text{--}0.0\text{M}$ ammonium sulphate in 0.1M sodium phosphate buffer pH 7.5 was applied at a flow rate of 0.5ml min^{-1} and the CHMS dehydrogenase was eluted at the end of this gradient (fig. 4.2). Analysis of $10\mu\text{l}$ samples of each fraction by SDS-PAGE showed the progression of the purification and can be seen in figure 4.3. To establish the purity of the final product, 1 , 5 and $10\mu\text{g}$ of protein (as estimated by the method of Warburg and Christian, 1945) were run on an SDS-PAGE gel (fig. 4.4). The lower loading indicates that a single band was present in the fraction and the higher loading shows the absence of minor bands of other molecular weights. From this gel it was estimated that the single band accounted for over 95% of the total protein detected. Details of the purification are given in Table 4.1.

Characterisation of Purified CHMS Dehydrogenase.

SUBUNIT MOLECULAR WEIGHT.

The 7.5% to 20% gradient SDS-PAGE gel pictured in figure 4.4 was used to estimate the subunit molecular weight of the purified protein. Seven standard proteins were used to calibrate the gel and the resulting graph (fig. 4.5) indicated a subunit molecular weight of approximately $52,000$.

FIGURE 4.1: MONO Q COLUMN: Protamine sulphate treated material was applied to a Mono Q anion exchange column and proteins eluted using a 100 ml gradient of 0-0.6 M NaCl. The CHMS dehydrogenase activity eluted at approximately 0.25 M NaCl- the activity of the peak fractions is given as the open circles. The scale is $\mu\text{mol min}^{-1} \text{mg}^{-1}$ protein.

FIGURE 4.1: ELUTION OF PROTEINS FROM A HR 10/10 MONO Q COLUMN.

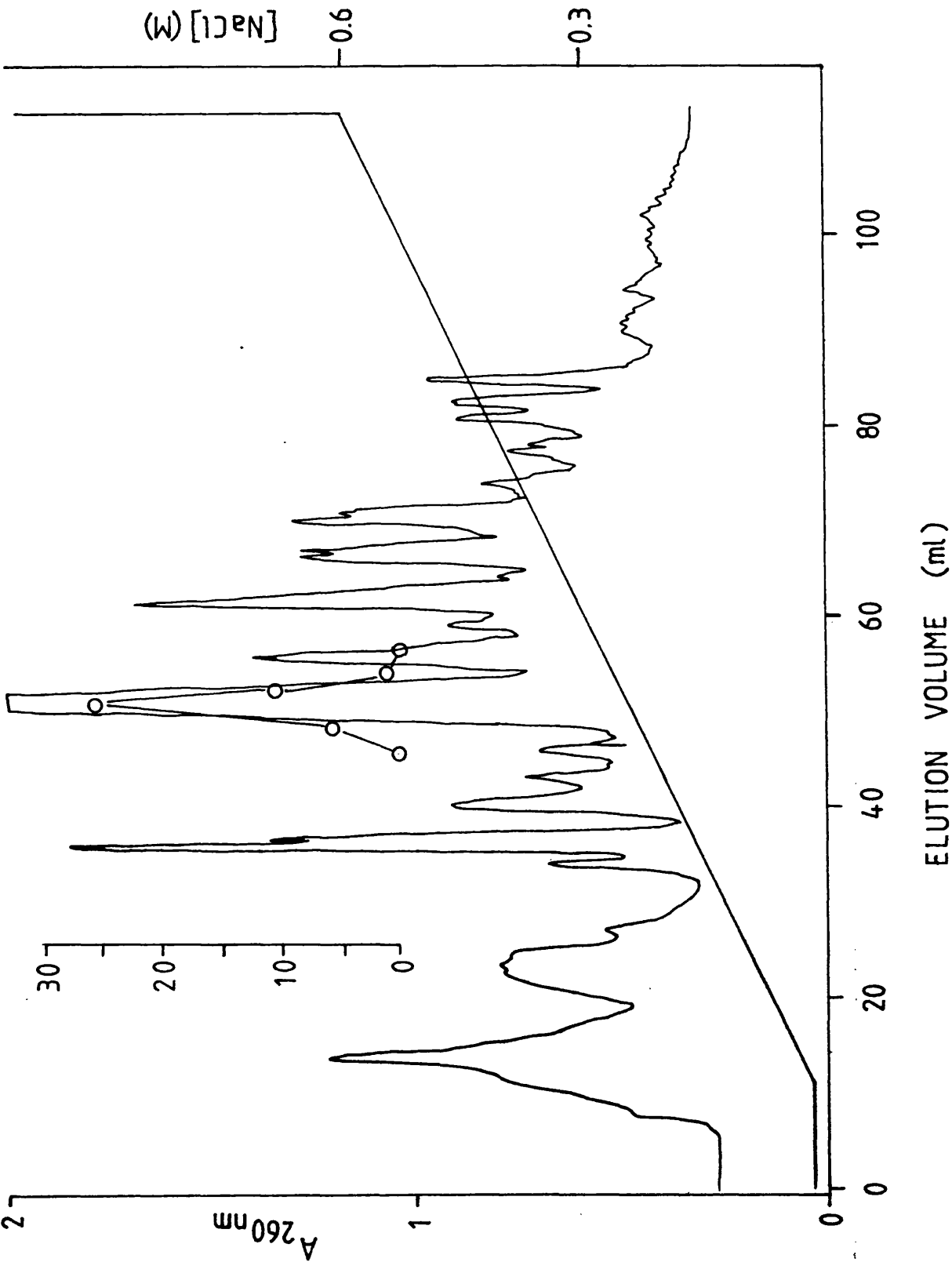


FIGURE 4.2: PHENYL SUPEROSE COLUMN: Peak activity fractions from the Mono Q column were run on a Phenyl Superose column and proteins eluted using a 1.7-0 M gradient of ammonium sulphate. The CHMS dehydrogenase activity was present in the large peak.

FIGURE 4.2: ELUTION OF CHMS DEHYDROGENASE FROM A HR 5/5 PHENYL SUPEROSE COLUMN.

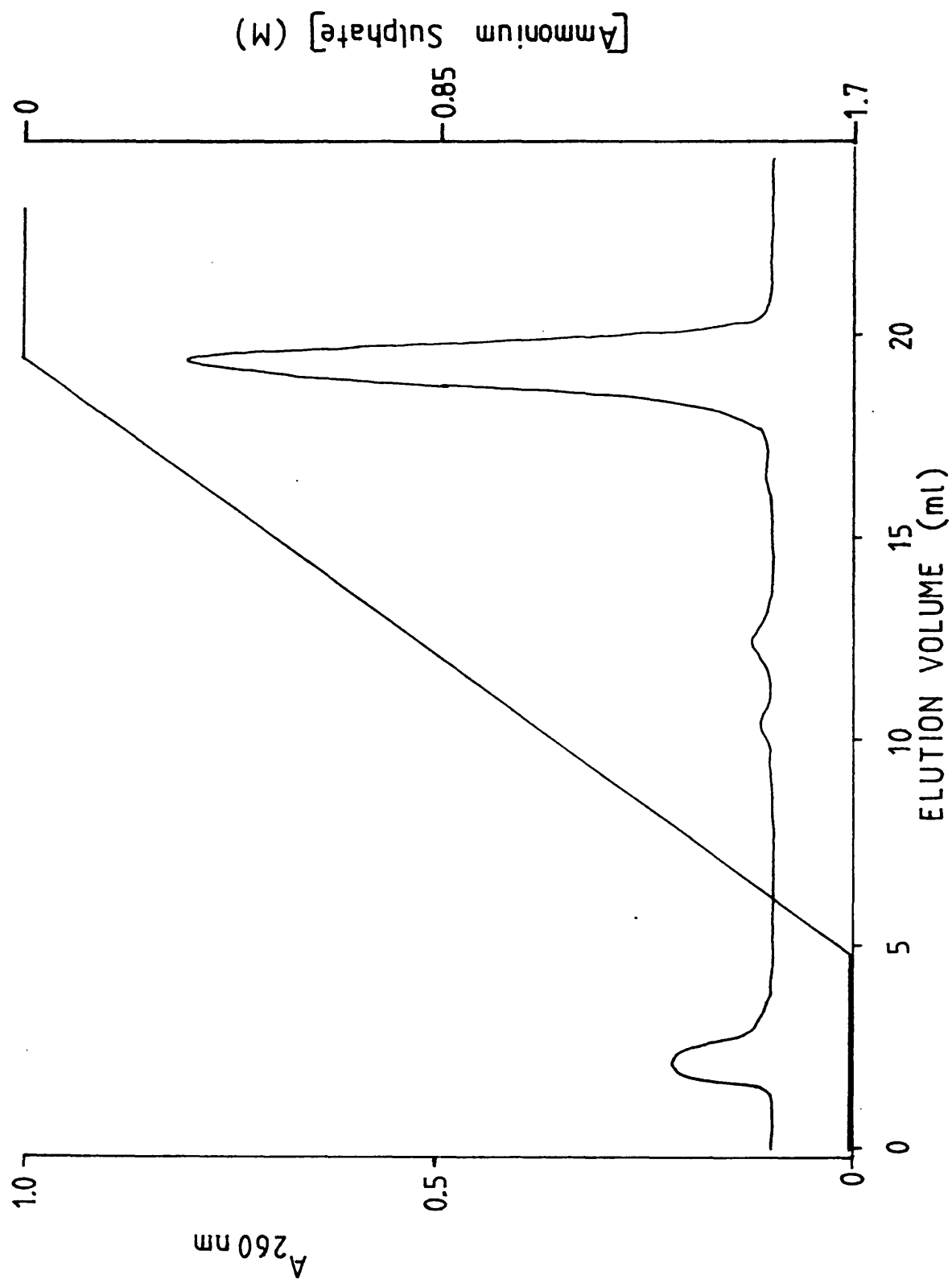


FIGURE 4.3: SDS-PAGE analysis of protein samples: 10 μ l of each fraction was run on a 7.5%-20% SDS-polyacrylamide gradient gel. The lanes contained: 1, crude extract; 2, first protamine sulphate fractionation supernatant; 3, second protamine sulphate fractionation supernatant; 4, second protamine sulphate supernatant filtered; 5, peak activity fraction from Mono Q column; 6, peak activity fraction from Mono Q column, filtered; 7 and 8, peak activity fraction from separate runs on a phenyl superose column. The proteins used to calibrate the gel were; bovine lactalbumin (14,200); trypsin inhibitor (20,100); trypsinogen (24,000); carbonic anhydrase (29,000); glyceraldehyde 3-phosphate dehydrogenase (36,000); ovalbumin (45,000) and bovine serum albumin (66,000) and these were present in the lanes marked M.

FIGURE 4.3: SDS-PAGE GEL SHOWING STEPS IN THE PURIFICATION OF CHMS
DEHYDROGENASE FROM *E. coli* C.

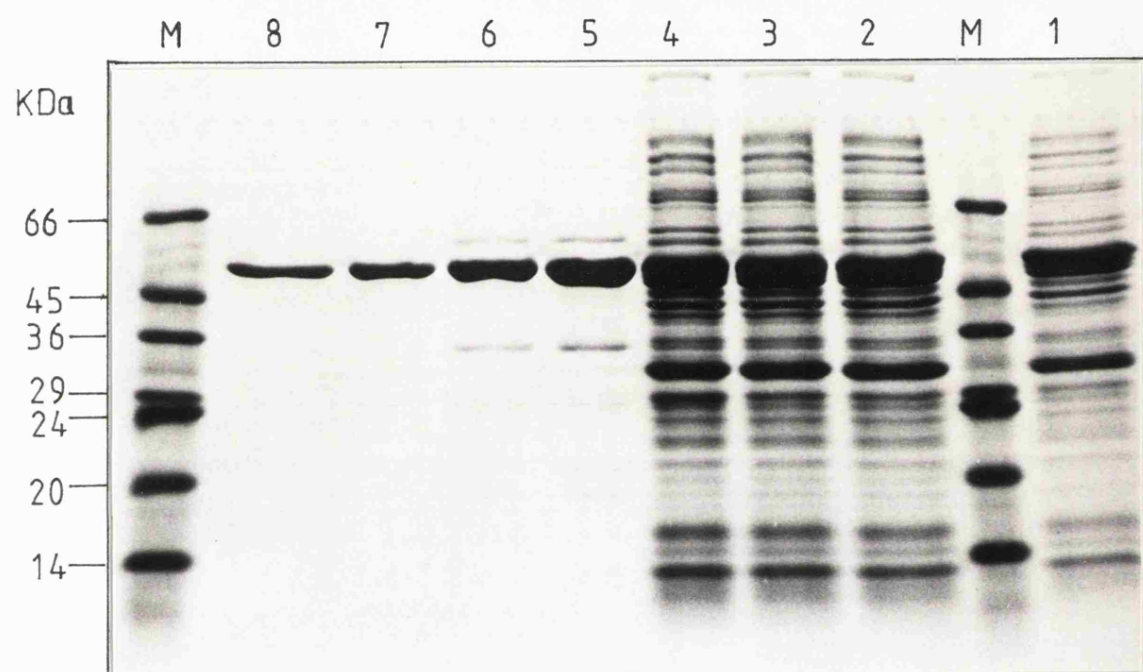


FIGURE 4.4: Purity of CHMS dehydrogenase: Samples equivalent to 1 μ g (lane 1), 5 μ g (lane 2) and 10 μ g (lane 3) of protein -measured by the method of Warburg and Christian- were electrophoresed on a 7.5%-20% SDS-polyacrylamide gradient gel. The proteins used to calibrate the gel were; bovine lactalbumin (14,200); trypsin inhibitor (20,100); trypsinogen (24,000); carbonic anhydrase (29,000); glyceraldehyde 3-phosphate dehydrogenase (36,000); ovalbumin (45,000) and bovine serum albumin (66,000) and these were present in the lanes marked M.

FIGURE 4.4: SDS-PAGE GEL SHOWING PURITY OF CHMS DEHYDROGENASE.

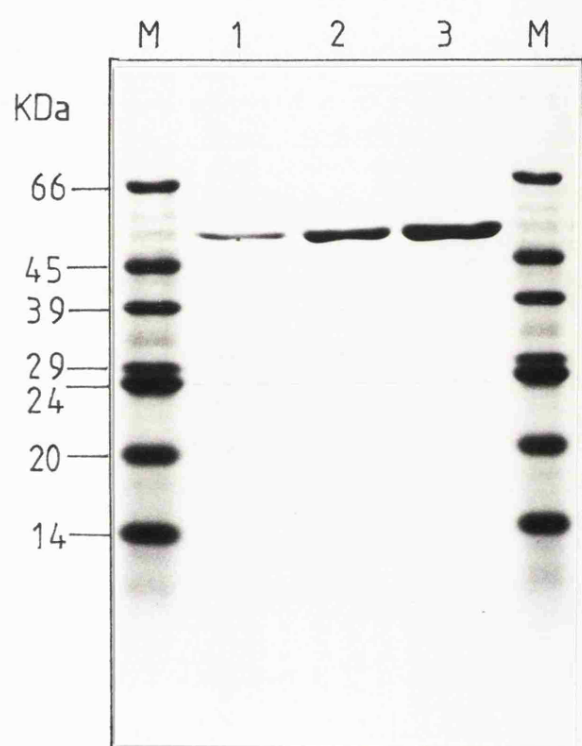
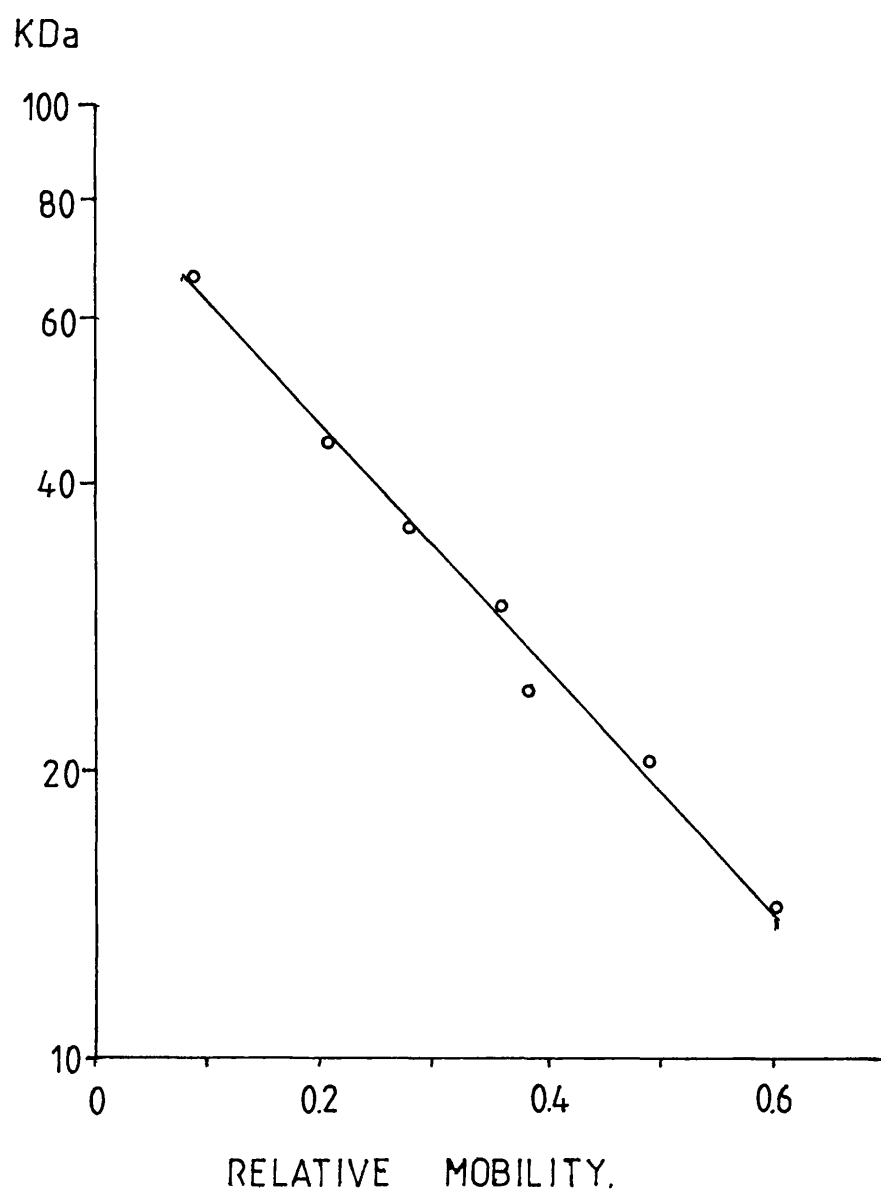


TABLE 4.1: A SUMMARY OF THE PURIFICATION OF CHMS DEHYDROGENASE FROM *E. coli* C.

STEP	VOLUME (ml)	TOTAL PROTEIN (mg)	TOTAL UNITS ($\mu\text{mol min}^{-1}\text{ml}^{-1}$)	SPECIFIC ACTIVITY (Units mg protein $^{-1}$)
1. Ultracentrifuged Extract	16	83	400	4.8
2. Protamine Sulphate	16	102	379	3.7
3. Mono Q	4	6.8	182	26.8
4. Phenyl Superose	4	2.2	77	35

FIGURE 4.5: CALIBRATION GRAPH: The standard proteins were: bovine lactalbumin (14,200); trypsin inhibitor (20,100); trypsinogen (24,000); carbonic anhydrase (29,000); glyceraldehyde 3-phosphate dehydrogenase (36,000); ovalbumin (45,000); and bovine serum albumin (66,000). The mobilities are relative to the bromophenol blue marker dye.

FIGURE 4.5: A GRAPH SHOWING THE CALIBRATION CURVE OF THE SDS-PAGE GEL IN
FIGURE 4.4.



GEL FILTRATION.

The native protein molecular weight was estimated by Suparose 12 gel filtration chromatography of a sample of purified CHMS dehydrogenase. The column was run with 50mM sodium phosphate buffer pH 7.0 containing 0.15M NaCl and a flow rate of 0.4ml min^{-1} . The enzyme activity eluted from the column in a single symmetrical peak and analysis of a calibration graph (fig. 4.6) constructed from the elution profile of standard proteins produced an estimate of $125,000(\pm 5,000)$ for the native molecular weight of the CHMS dehydrogenase. As a single band was evident on a denaturing polyacrylamide gel the evidence suggests that the native form of the enzyme was a dimer of identical subunits.

MOLAR EXTINCTION COEFFICIENT OF CHMS.

The molar extinction coefficient of CHMS at pH 7.8 is $35,500\text{ dm}^3\text{ mol}^{-1}\text{ cm}^{-1}$ (Y-L.T.Lee, M.S. thesis, University of Minnesota, St. Paul, 1977). However, addition of the same volume of a preparation of CHMS to buffers with different pH values resulted in different absorbance readings; ie: $\epsilon_{330}\text{ CHMS}$ is pH dependent. The molar extinction coefficient of CHMS at different pH values was calculated by comparing the absorbance at pH 7.8 of a fixed volume of CHMS with the absorbance at different pH values. $\epsilon_{330}\text{CHMS}$ varies between 8,700 at pH 6.0 and 45,600 at pH 8.0 when measured in phosphate buffer (Table 4.2) and these figures were used subsequently when determining the pH optimum of the purified CHMS dehydrogenase.

pH OPTIMUM.

A constant concentration of CHMS ($26\mu\text{M}$) was used throughout this experiment; the concentration of NAD was 0.2mM and each assay contained approximately 69ng of purified enzyme. The rate of reaction was highest in 0.1M sodium phosphate buffer at pH 7.5. The rate of reaction was at least two thirds of the maximum rate over the range pH 6.8–8.0 indicating a broad range of activity. Below

FIGURE 4.6: Gel filtration chromatography: Samples of purified CHMS dehydrogenase were run in 50mM sodium phosphate buffer-150mM NaCL pH 7.0 at a flow rate of 0.4ml min⁻¹. The column was calibrated using IgG (160,000), human serum albumin (67,000), β -lactoglobulin (35,000) and cytochrome *c* (12,400) as standards.

FIGURE 4.6: CALIBRATION GRAPH OF THE SUPEROSE 12 GEL FILTRATION COLUMN.

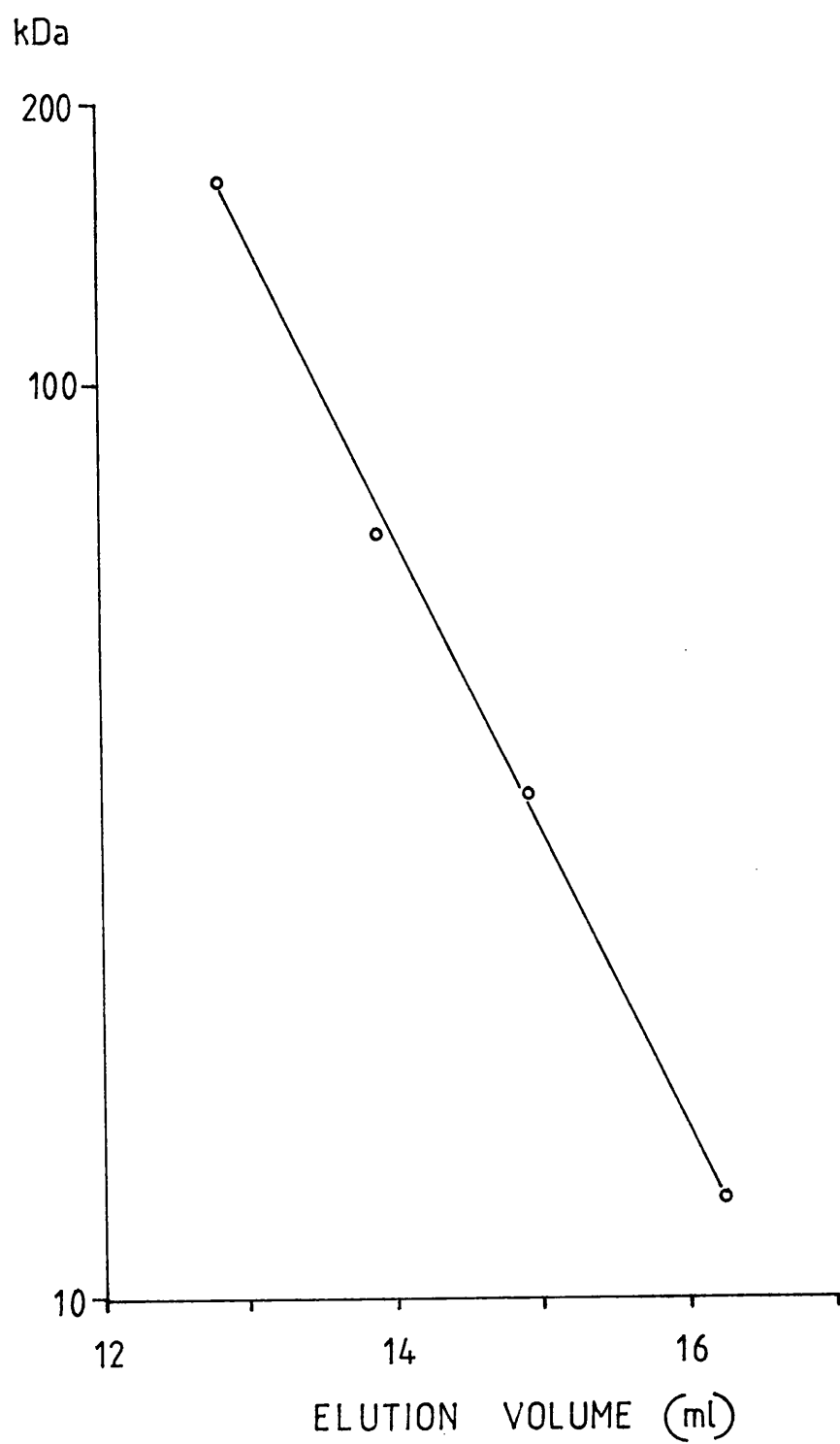


TABLE 4.2: MOLAR EXTINCTION COEFFICIENTS OF CHMS AT DIFFERENT pH VALUES.

pH	$\epsilon_{330, \text{CHMS}}$
6.0	8,700
6.6	14,800
6.8	21,900
7.0	23,100
7.2	23,800
7.4	26,400
7.5	31,800
7.6	32,800
7.8	35,500
8.0	45,600

pH 6.8 the activity dropped off quickly (fig. 4.7).

APPARENT K_m DETERMINATIONS.

The K_m values for each substrate were measured at pH 7.5 using a single saturating concentration of the other substrate. With $196\mu\text{M}$ NAD the K_m for CHMS was $9.2\pm 1.4\mu\text{M}$ and with $26\mu\text{M}$ CHMS the K_m for NAD was $11.7\pm 1.5\mu\text{M}$. The method used to determine these values and the resulting graphs are shown in section 4.2 (page 121) where they are compared to the corresponding data from the purified *Klebsiella pneumoniae* CHMS dehydrogenase.

AMINO-TERMINAL AMINO ACID SEQUENCING.

A 0.5 ml sample of the peak activity fraction from the Phenyl Superose column was dialysed overnight against 100 volumes of 25mM sodium phosphate buffer pH 7.5 before being submitted for amino-terminal amino acid sequencing. The sequencing was kindly carried out by Dr.M.D. Davison using an Applied Biosystems 470A gas phase sequencer. The first 34 residues of the enzyme were determined without any ambiguities and the sequence is shown in figure 4.8.

FIGURE 4.7: GRAPH SHOWING THE VARIATION IN ENZYME SPECIFIC ACTIVITY WITH pH.

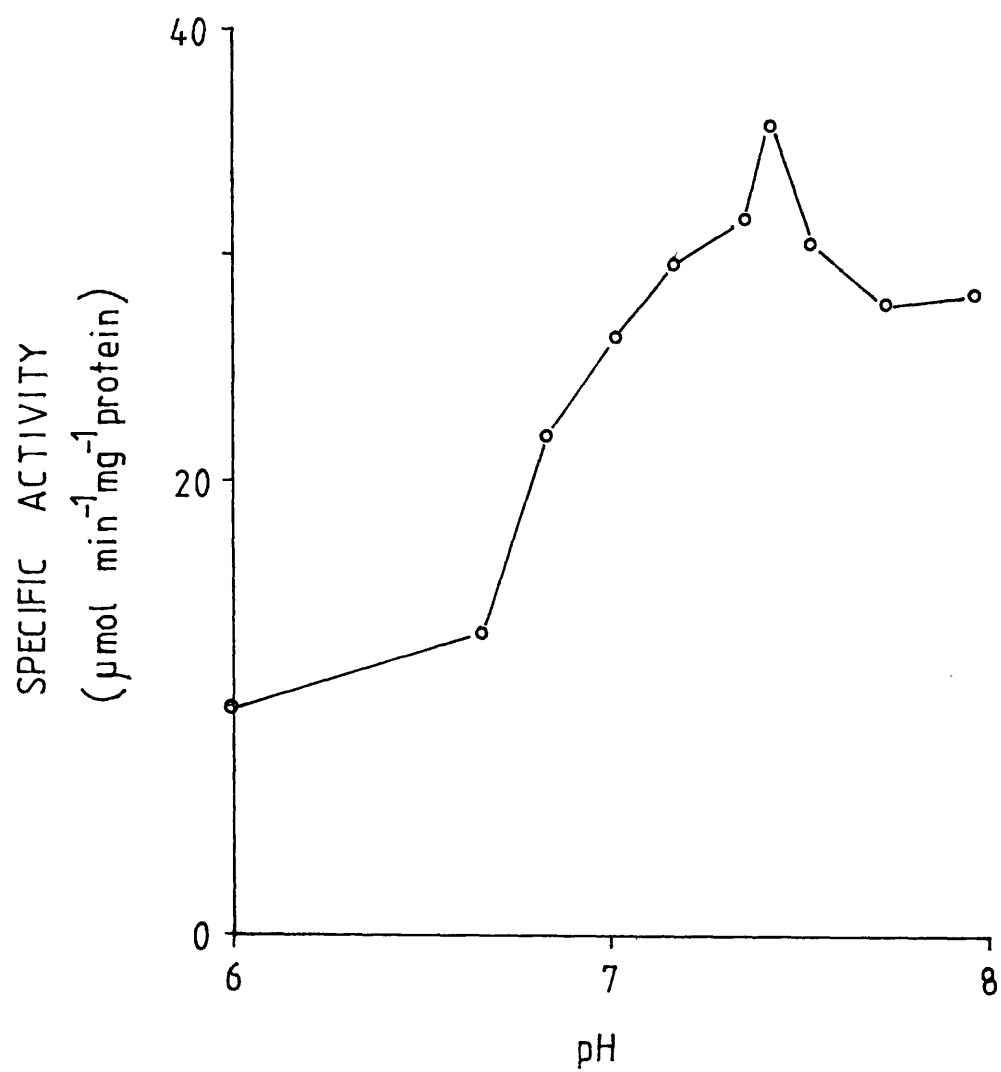


FIGURE 4.8: SEQUENCE FROM THE FIRST 34 RESIDUES OF THE AMINO TERMINUS OF *E. coli* CHMS DEHYDROGENASE.

Residue Number:	1	2	3	4	5	6	7	8	9	10
Residue:	Met	Lys	Lys	Val	Asn	His	Trp	Ile	Asn	Gly
Residue Number:	11	12	13	14	15	16	17	18	19	20
Residue:	Lys	Asn	Val	Ala	Gly	Asn	Asp	Tyr	Phe	Gln
Residue Number:	21	22	23	24	25	26	27	28	29	30
Residue:	Thr	Thr	Asn	Pro	Ala	Thr	Gly	Glu	Val	Leu
Residue Number:	31	32	33	34						
Residue:	Ala	Asp	Val	Ala						

4.2: HOMOLOGIES BETWEEN THE CHMS DEHYDROGENASE ENZYMES OF *E. coli* AND *K.*

pneumoniae.

The CHMS dehydrogenase from *Escherichia coli* B has been purified and characterised (Alonso and Garrido-Pertierra, 1982; 1986). The native molecular weight of the purified enzyme was found to be 210,000 and the subunit weight 58,000, suggesting a tetramer of identical subunits. The enzyme displayed a narrow pH optimum with maximum activity at pH 7.8 in phosphate buffer (Alonso and Garrido-Pertierra, 1982). When CHMS and NAD were used as variable substrates the K_m for CHMS was found to be $9.0 \pm 1.0 \mu\text{M}$ and the K_m for NAD $29.1 \pm 4.6 \mu\text{M}$ (Alonso and Garrido-Pertierra, 1986). No amino-terminal sequencing of this protein was undertaken.

The corresponding enzyme from *K. pneumoniae* M5a1 was purified by Professor A. Garrido-Pertierra and the amino-terminal amino acid sequence determined by Dr.M.D. Davison. The first 20 amino-terminal residues of the *Klebsiella* enzyme were determined without any ambiguities. Figure 4.9 compares the sequences of *E. coli* C and *K. pneumoniae* M5a1 enzymes. Of the 20 amino acids for which direct comparison is possible, 18 of the residues are identical and the other two are conservative substitutions. The subunit molecular weight and native molecular weight of the *Klebsiella* enzyme were also determined by Professor Garrido-Pertierra and were 52,000 and 102,000 ($\pm 5,000$) respectively.

pH OPTIMUM OF *Klebsiella* ENZYME.

A constant concentration of CHMS (26 μM), which produced an absorbance of 0.825 at pH 7.5, was used throughout this experiment; the concentration of NAD was 0.2mM and each assay contained approximately 100ng of purified enzyme. The rate of reaction was highest in 0.1M sodium phosphate buffer at pH 7.5.

**FIGURE 4.9: AMINO-TERMINAL AMINO ACID SEQUENCES OF THE CHMS DEHYDROGENASES OF
E. coli C AND *K. pneumoniae* M5a1.**

Residue Number:	1	2	3	4	5	6	7	8	9	10
<i>E. coli</i> :	Met	Lys	Lys	Val	Asn	His	Trp	Ile	Asn	Gly
<i>K.pneumoniae</i> :	Met	Lys	Lys	Ile	Asn	His	Trp	Ile	Asn	Gly
Residue Number:	11	12	13	14	15	16	17	18	19	20
<i>E. coli</i> :	Lys	Asn	Val	Ala	Gly	Asn	Asp	Tyr	Phe	Gln
<i>K. pneumoniae</i> :	Lys	Asn	Val	Ala	Gly	Ser	Asp	Tyr	Phe	Gln

APPARENT K_m DETERMINATIONS OF *Klebsiella* CHMS DEHYDROGENASE.

The K_m values for each substrate were measured at pH 7.5 using a single saturating concentration of the other substrate. With 196 μ M NAD the K_m for CHMS was 5.2 \pm 1.6 μ M and with 26 μ M CHMS the K_m for NAD was 11.7 \pm 1.5 μ M. The method used to determine these values and those of the *E. coli* enzyme was by a non-linear least squares regression fit of v against S by the computer programme of Walmsley and Lowe (1985). The resulting fit produced by this analysis is pictured in figure 4.10. Both of the enzymes show hyperbolic kinetics when CHMS or NAD was used as the variable substrate.

Comparison of purified CHMS dehydrogenases.

Three CHMS dehydrogenase enzymes have now been purified from various organisms which have the ability to catabolise HPC (Alonso and Garrido-Pertierra, 1982; Fawcett *et al.*, 1989). Comparison of the three enzymes is given in Table 4.3. The subunit molecular weights of the enzymes are all similar, but the *E. coli* B enzyme appears to be associated into a tetrameric form rather than a dimer (the most likely configuration of the other two enzymes). The three enzymes exhibit similar, but not identical, affinities for their two substrates and all operate optimally at a similar pH value. The most striking similarity, however, is the amino-terminal sequences. Unfortunately the data was not available for the *E. coli* B enzyme, but the other two CHMS dehydrogenases revealed a 90% identity in the first twenty amino acids at their amino-termini. The other two amino acids showed conservative substitutions.

FIGURE 4.10: Kinetic plots of v against S with a single saturating concentration of one of the substrates for two purified CHMS dehydrogenases. a) *K. pneumoniae* M5a1 enzyme with 26 μ M CHMS and various NAD concentrations, c) *K. pneumoniae* M5a1 enzyme with 196 μ M NAD and various CHMS concentrations, b) *E. coli* C enzyme with 26 μ M CHMS and various NAD concentrations and d) *E.coli* C enzyme with 196 μ M NAD and various CHMS concentrations.

FIGURE 4.10: KINETIC PLOTS OF v AGAINST S FOR CHMS DEHYDROGENASES OF *E.coli* C

AND *K. pneumoniae* M5a1.

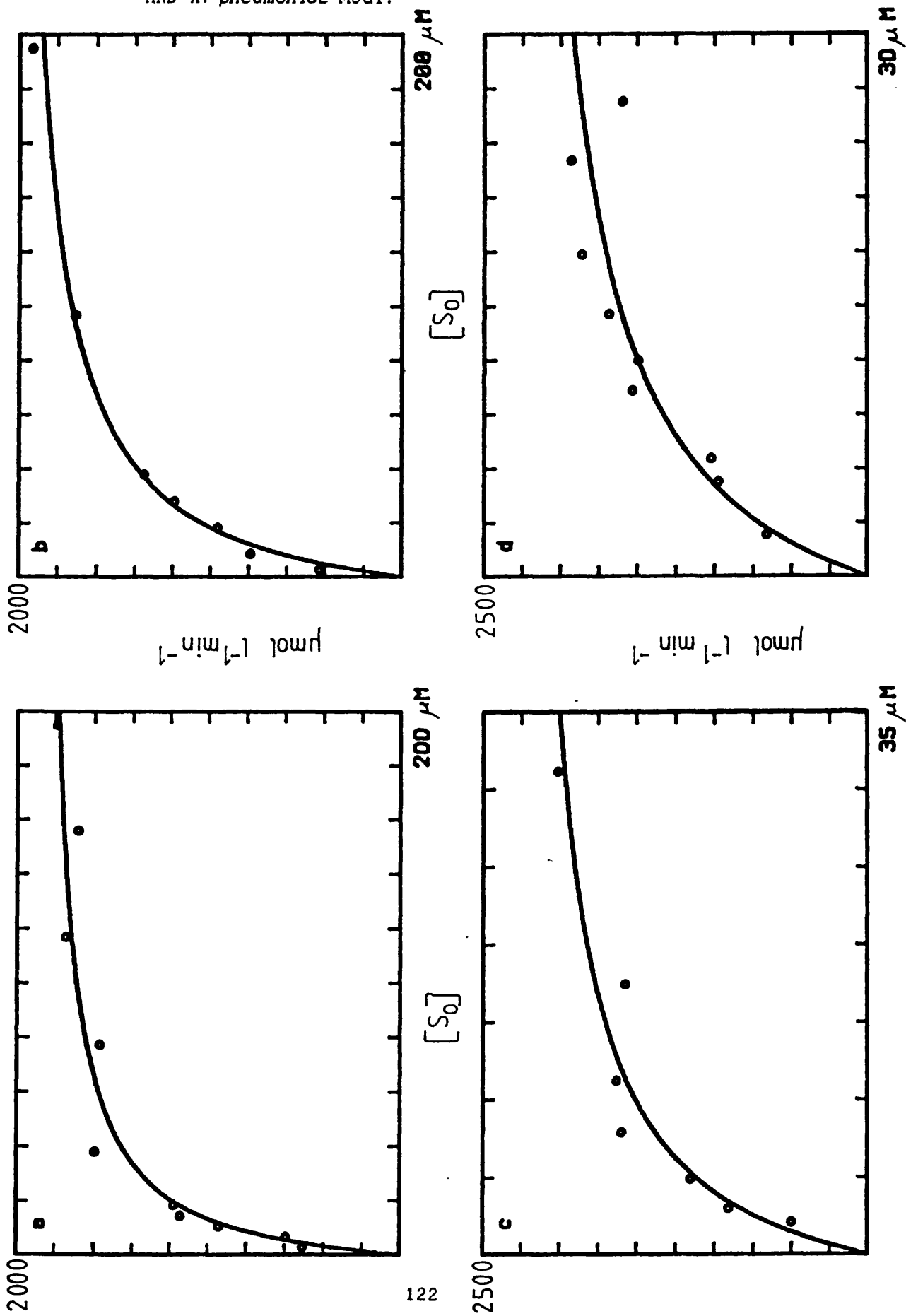


TABLE 4.3: SUMMARY OF KNOWN CHARACTERISTICS OF PURIFIED CHMS DEHYDROGENASES.

CHARACTERISTIC	<i>E. coli</i> C	<i>K. pneumoniae</i> M5a1	<i>E. coli</i> B
Subunit Molecular Weight	52,000	52,000	58,000
Native Molecular Weight	120,000 ($\pm 5,000$)	102,000 ($\pm 5,000$)	210,000 ($\pm 10,000$)
pH Optimum	7.5	7.5	7.8
K_m (CHMS)	9.2 ± 1.4	5.2 ± 1.6	9.0 ± 1.0
K_m (NAD)	19.4 ± 2.7	11.7 ± 1.5	29.1 ± 4.6
Specific Activity	35	54	54
Amino-Terminal Homology	90% homology in first 20 residues		N/T

N/T= not tested.

Specific Activity measured in $\mu\text{mol min}^{-1} \text{mg}^{-1}$ protein.

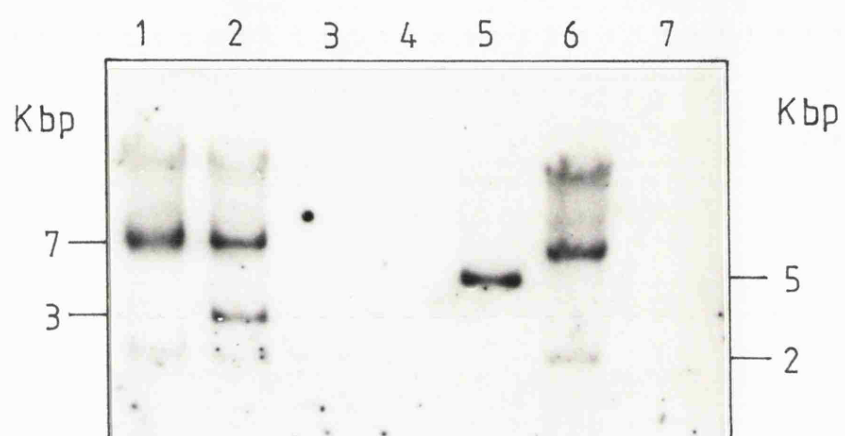
Southern Blot Analysis of CHMS Dehydrogenase Genes Using a Probe Derived from the Amino-Acid Sequence.

The physical characteristics of the CHMS dehydrogenases (Table 4.3) show some similarities. However, properties such as K_m values do not provide the kind of evidence on the relatedness of proteins as primary DNA or protein sequence data do. In the absence of purified proteins or sequenced genes it is possible to assess how well conserved a particular DNA sequence is in different species by DNA hybridisation.

A portion of amino acid sequence which was absolutely conserved between the CHMS dehydrogenases of *E. coli* C and *K. pneumoniae* M5a1 was chosen and a degenerate oligonucleotide probe constructed such that all possible DNA sequences for that peptide sequence were represented; the oligonucleotide mixture was called probe C (fig. 3.6b). Five micrograms of chromosomal DNA from *E. coli* C, B, W and K-10, *K. pneumoniae* M5a1 and *Pseudomonas putida* were each digested with *Bam*H1 and then electrophoresed through a 0.6% agarose gel, the DNA blotted onto nitrocellulose and immobilised by baking for 2h at 80°C. The nitrocellulose filter was treated with radiolabelled probe C and washed. The probe hybridised strongly to an approximately 7kbp fragment of *E. coli* B, C and W DNA. *E. coli* W showed a second strongly hybridising band of approximately 3kbp. *K. pneumoniae* gave a single strongly hybridising band of approximately 5kbp. No strongly hybridising band was seen for *E. coli* K-12 (strain K-10) or *P. putida* DNA. However, all of the *E. coli* samples showed a very weakly hybridising band of approximately 2kbp (fig. 4.11). Since there is no *Bam*H1 site in any of the sequences probed for, the two strongly hybridising bands seen for *E. coli* W suggests that there may be a duplicate form of the CHMS dehydrogenase gene in that organism. The failure to detect any significant hybridisation with *P. putida* DNA indicates that its CHMS dehydrogenase may have diverged significantly from that of *E. coli* C and *K. pneumoniae*.

FIGURE 4.11: Southern blot of chromosomal DNA: Autoradiograph of a Southern blot of *Bam*H1 digested bacterial DNA hybridised with radiolabelled probe C. The lanes contained DNA from the following sources: 1, *E.coli* C; 2, *E. coli* W; 3, *E. coli* K10; 4, *P. putida*; 5, *K. pneumoniae* M5a1; 6, *E. coli* B and 7, *Hind*III digested λ -DNA.

FIGURE 4.11: SOUTHERN BLOT OF *Bam*H1 DIGESTED CHROMOSOMAL DNA'S HYBRIDISED WITH PROBE C.



4.3: CLONING AND PRIMARY ANALYSIS OF *hpc* GENES FROM *Klebsiella pneumoniae*

M5a1.

Isolation of Clones.

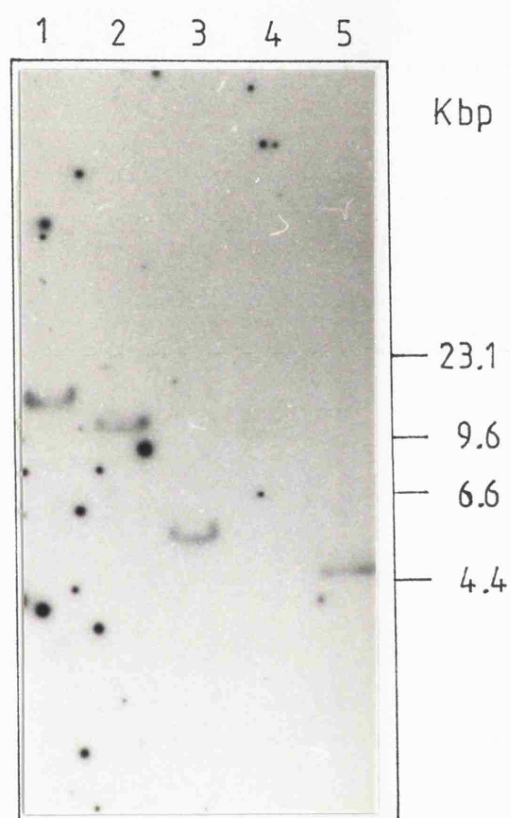
The amino terminal amino acid sequences of the *E. coli* C and *K. pneumoniae* CHMS dehydrogenase enzymes have been shown to be highly homologous. A degenerate oligonucleotide probe derived from a portion of those sequences (probe C) was able to hybridise to chromosomal DNA from *E. coli* C and *K. pneumoniae* M5a1 as well as to chromosomal DNA from *E. coli* strains B and W (Fawcett *et al.*, 1989). Although strong hybridisation was seen for *K. pneumoniae* DNA the size of the labelled fragment was different to that for *E. coli* C suggesting that the two organisms did not have identical restriction maps for the *hpc* genes. Detection of chromosomal copies of *hpc* genes with probe C indicated that a possible method of isolating the genes on a cloned fragment would be colony hybridisation.

Chromosomal DNA from M5a1 was digested with several restriction enzymes and electrophoresed through a 0.5% agarose gel. The DNA was blotted onto nitrocellulose and hybridised with radiolabelled probe C (fig. 4.12). Each track contained a single hybridising band; DNA digested with *Hind*III and *Sal*I hybridised to the probe at sizes which lay between the 23kbp and 9.6kbp λ -markers. *Bam*HI and *Pst*I digested DNA contained hybridising fragments of about 5kbp. In an attempt to obtain all of the *hpc* genes on a single fragment of DNA the restriction endonuclease *Hind*III was chosen to construct the gene library, as probe C hybridised to a large (approximately 15kbp) *Hind*III fragment. To reduce the number of recombinant plasmids which had to be screened in the library, completely digested chromosomal DNA was size fractionated and fragments of the required size selected prior to ligation into a plasmid vector.

K. pneumoniae M5a1 chromosomal DNA (330 μ g) was completely digested with 400

FIGURE 4.12: Southern blot of *Klebsilla* DNA: Autoradiograph of a Southern blot of *K. pneumoniae* M5a1 chromosomal DNA hybridised with probe C. The lanes contained DNA digested with the following enzymes; 1, *Hind*III; 2, *Sal*I; 3, *Bam*H1; 5, *Pst*I. Lane 4 contained *Hind*III digested λ -DNA.

FIGURE 4.12: SOUTHERN HYBRIDISATION OF M5a1 DNA WITH PROBE C.

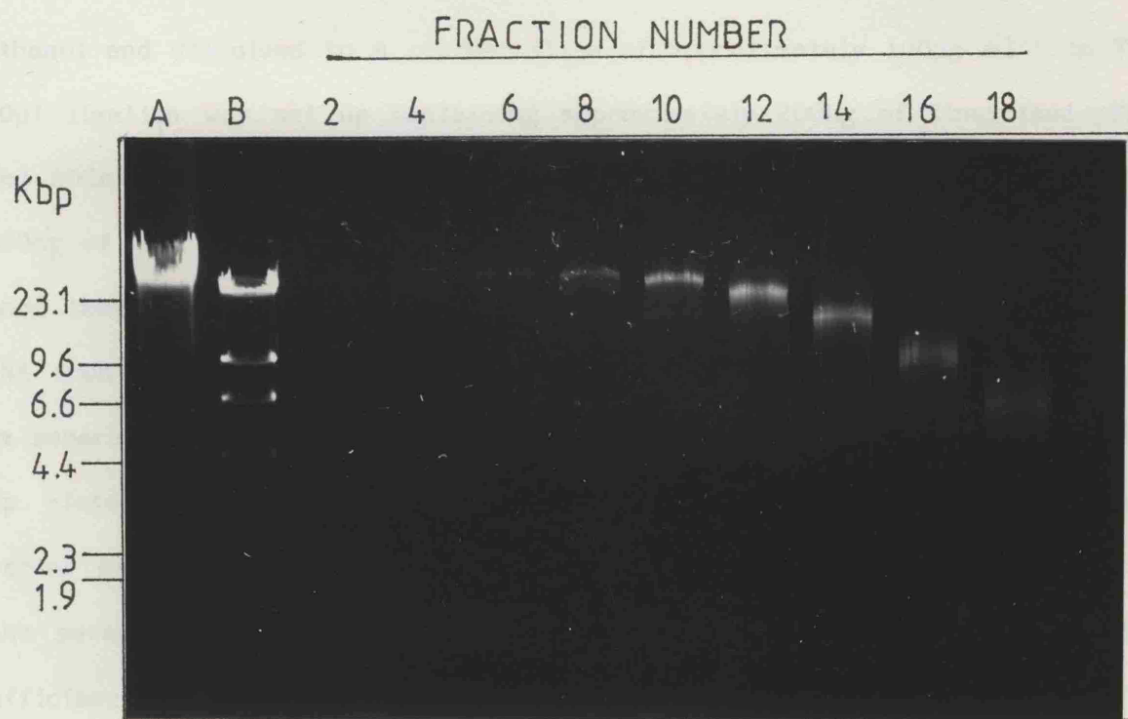


units of *Hind*III, the mixture extracted twice with phenol/chloroform and the DNA precipitated with two volumes of ethanol in the presence of 0.1 volumes of 3M sodium acetate. The precipitate was collected by centrifugation and washed with 70% ethanol, the pellet vacuum dried and dissolved in 500 μ l TE. A 5 μ l sample of the solution was electrophoresed through a 0.6% agarose gel to ensure that complete digestion had taken place and the remainder of the DNA solution was laid on top of a 10-40% linear sucrose gradient (in 1M NaCl, 20mM Tris.Cl pH 8.0, 5mM EDTA). The gradient was centrifuged for 24h at 25,000rpm in a Beckman SW27 rotor and then the gradient was collected in 1ml fractions using a peristaltic pump. A 20 μ l sample of each even numbered fraction was diluted with an equal volume of water, mixed and electrophoresed through a 0.5% agarose gel along with some non-digested chromosomal DNA and *Hind*III restricted λ -DNA. Figure 4.13 shows a photograph of the resulting gel and the successful size fractionation of the *Klebsiella* DNA.

Fractions which contained fragments of above 9.6kbp were tested for the presence of *hpc* coding sequences by a DNA-DNA dot blot assay. The DNA from 100 μ l of each even numbered fraction (up to number 18) was precipitated and redissolved in 20 μ l of water and denatured by the addition of 5 μ l of 4M NaOH and incubation at 37°C for 10 min. The samples were then maintained on ice for two min. before being spotted onto a nitrocellulose filter. The filter was air dried and baked at 80°C for two hours and then hybridised to radiolabelled probe C as previously described. The autoradiograph revealed that fractions 8, 10, 12, 14 and 16 hybridised to the probe to some extent; the strongest hybridisation being with fraction 14 and the second strongest in fraction 16. Therefore, 400 μ l of fractions 14 and 15 were pooled and the DNA precipitated by an equal volume of isopropanol in the presence of tRNA at a final concentration of 20 μ g ml⁻¹. The precipitate was collected by centrifugation and washed in ethanol, vacuum dried briefly and dissolved in 200 μ l of sterile water. Electrophoresis of a sample of the DNA indicated that the concentration

FIGURE 4.13: Size fractionation of chromosomal DNA: The DNA was fractionated on a 10-40% linear sucrose gradient (in 1M NaCl, 20mM Tris.Cl pH 8.0, 5mM EDTA) by centrifuging at 25,000rpm for 24h in a Beckman SW27 rotor and 1ml fractions collected using a peristaltic pump; 20 μ l samples of each even numbered fraction were electrophoresed through a 0.6% agarose gel. lane A contained undigested chromosomal DNA and lane B contained *Hind*III digested λ -DNA.

FIGURE 4.13: GEL SHOWING THE SIZE FRACTIONATION OF CHROMOSOMAL DNA.

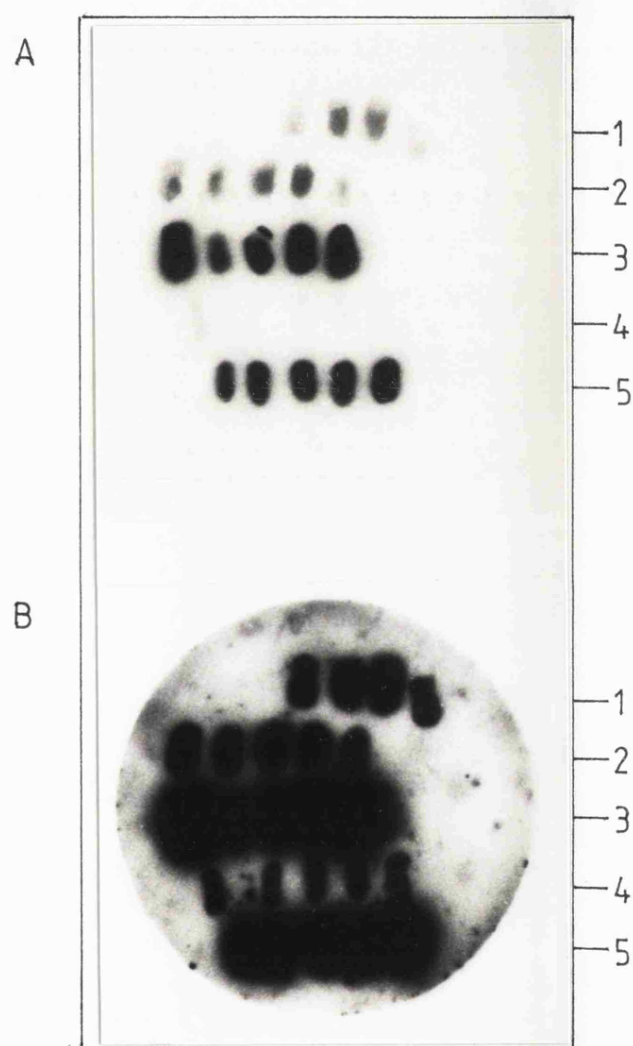


was approximately $50\mu\text{g ml}^{-1}$ and had a size distribution of between approximately 10 to 15kbp.

The chosen vector, pBR322, was prepared by restriction enzyme digestion with *Hind*III and phosphatase treatment to remove terminal phosphate groups. The solution was extracted with phenol/chloroform and the DNA precipitated with ethanol and dissolved to a concentration of approximately $100\mu\text{g ml}^{-1}$ in TE. A 20 μl ligation was set up containing approximately 200ng of linearised pBR322 and 200ng of size selected chromosomal DNA; a control ligation contained only 200ng of linearised pBR322. The ligations were allowed to proceed overnight at room temperature and then 100ng of DNA from the control ligation and 200ng of DNA from the test ligation were used to transform competent *E. coli* 5K cells in separate experiments. The transformed cells were spread onto pre-warmed LB-Ap plates and incubated overnight. The linearised and phosphatase treated vector gave 4,300 transformants per microgram of vector DNA whereas each of the seven test plates contained approximately 1,000 colonies (an approximate efficiency of 70,000 transformants per microgram of vector DNA). All of the test plates were replica-plated onto separate nitrocellulose filters resting on top of LB-Ap plates; these were incubated overnight such that small colonies were present on the filters. The filters were marked with a soft pencil, removed from the plates, the colonies lysed (as described in methods) and the DNA fixed to the nitrocellulose by baking for 2h at 80°C. Each filter was pre-hybridised and hybridised with radiolabelled probe C, air dried and exposed to X-ray film. Two hours exposure of the filters to the film revealed approximately 100 possible positive colonies. As the master plates were overcrowded a selection of positives colonies were taken from them and spread onto fresh LB-Ap plates to obtain single colonies. Several single colonies from each putative positive were patched onto nitrocellulose filters resting on top of LB-Ap plates and the screening procedure was repeated for a second time. Figure 4.14 shows an example of an autoradiogram of a single

FIGURE 4.14: Screening putative positive clones: Panel A shows an autoradiogram of a single nitrocellulose filter containing 5 putative positive clones from the initial round of screening after 30 min. exposure of the filter to X-ray film. Panel B shows an autoradiogram of the same filter after overnight exposure of the filter to X-ray film.

FIGURE 4.14: AUTORADIOGRAM OF A SECONDARY SCREEN WITH PROBE C.



nitrocellulose filter after 30 min and overnight exposure to X-ray film. The filter contains single colonies from 5 putative positive clones in the first round of screening; two of the groups were all positive after 30 min exposure and an additional two were regarded as positive after overnight exposure. The next to bottom row were regarded as negative. In all, 13 clones were positive after the second round of screening by colony hybridisation and these were further investigated by the restriction pattern of the DNA after digestion with *Bam*H1. The DNA from each of the positive clones was restricted completely with *Bam*H1 and electrophoresed through a 0.8% agarose gel (fig. 4.15); the clones fell into 5 different restriction pattern groups and were named pTF100-pTF104. The clones which gave the strongest hybridisation in the second round of screening fell into two groups by their restriction patterns (pTF100 and pTF102).

Complementation of *E. coli* C Mutants of the HPC Pathway with Putative Clones.

Several mutant strains of *E. coli* C were available which contained lesions in *hpa* or *hpc* genes, including an *hpcC* strain. As the clones had been selected using a DNA probe derived from the amino-terminal amino acid sequence of the *hpcC* gene product, complementation of this mutant would be likely if an active *hpcC* gene had been cloned. This method although quick and easy could give false-negative results if only a portion of the gene was cloned or the cloned *K. pneumoniae* genes were not expressed in the *E. coli* host. The cloned genes were also transformed into other mutant strains: JJ210 is a strain which remains HPA⁻ even with pJJ801 (*hpcBCDEF*) present and D03 is a strain which is HPA⁻/HPC⁺. *E. coli* 5K was also transformed with the clones and tested for growth on HPA and HPC in order to determine whether all of the genes required for growth on either of these two carbon sources had been cloned. Table 4.4 shows the growth characteristics of these strains containing the various clones after 48h incubation at 37°C. *E. coli* C and *K. pneumoniae* M5a1 were included on

FIGURE 4.15: Photograph of restriction digest analysis of putative positive clones: the DNA from each clone was restricted with *Bam*H1 and electrophoresed through a 0.8% agarose gel. The clone in lane 1 was called pTF100; those with the restriction pattern of lanes 2 and 11, pTF101; lanes 5, 12 and 13, pTF102; lanes 4, 7 and 9, pTF103 and lanes 6, 8 and 10, pTF104. The lanes marked M contained *Hind*III digested λ -DNA.

FIGURE 4.15: RESTRICTION ANALYSIS OF PUTATIVE POSITIVE CLONES.

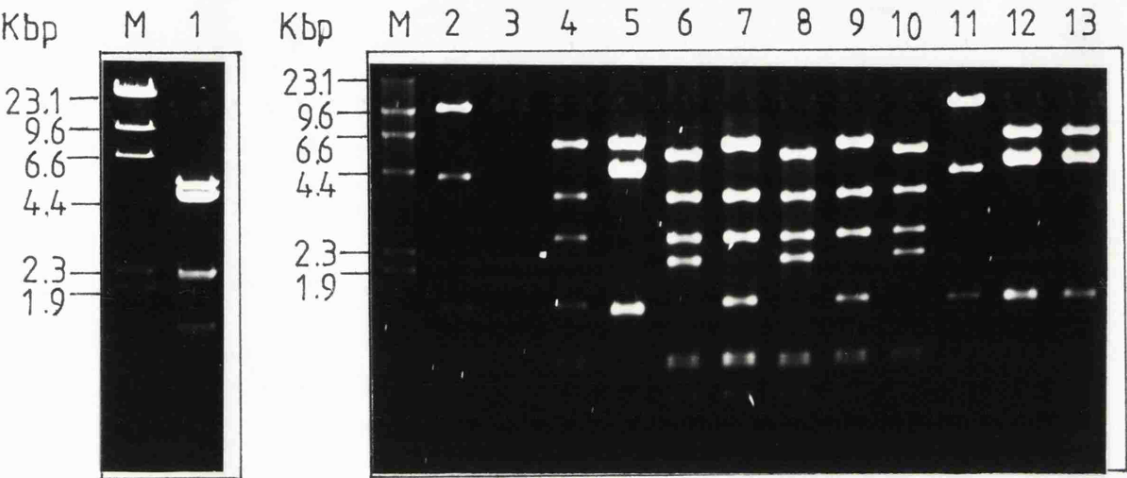


TABLE 4.4: COMPLEMENTATION OF VARIOUS STRAINS BY CLONED *Klebsiella* DNA.

STRAIN	GROWTH SUBSTRATE		
	HPC	HPA	GLUCOSE
<hr/>			
<i>E. coli</i> C	4+	5+	5+
M5a1	3+	5+	5+
JJ221	-	-	5+
JJ221 (pTF100)	2+	3+	5+
JJ221 (pTF101)	-	-	5+
JJ221 (pTF102)	2+	4+	5+
JJ221 (pTF103)	-	-	5+
JJ221 (pTF104)	-	-	5+
JJ210	-	-	5+
JJ210 (pTF100)	2+	2+	5+
JJ210 (pTF101)	-	-	5+
JJ210 (pTF102)	2+	2+	5+
JJ210 (pTF103)	-	-	5+
JJ210 (pTF104)	-	-	5+
D03	4+	-	5+
D03 (pTF100)	4+	-	5+
D03 (pTF101)	4+	-	5+
D03 (pTF102)	4+	-	5+
D03 (pTF103)	4+	-	5+
D03 (pTF104)	4+	-	5+
5K	-	-	5+
5K (pTF100)	-	-	5+
5K (pTF101)	-	-	5+
5K (pTF102)	-	-	5+
5K (pTF103)	-	-	5+
5K (pTF104)	-	-	5+

Growth was marked on a scale of 0-5+; the maximum being taken as the growth using glucose as the carbon source after two days incubation.

the plates as positive controls.

The clones pTF100 and pTF102 complemented the mutations in strains JJ221 and JJ210; these clones also gave the strongest hybridisation to radiolabelled probe C. No clones were able to repair the mutation in D03, which is able to grow on HPC but not HPA and is therefore probably deficient in an *hpa* gene, or allow 5K to grow on either HPA or HPC. This experiment suggests that the *hpcC* gene and at least one other gene of the *hpc* pathway were present on both pTF100 and pTF102, but they did not contain all of the genes required for growth on HPA or HPC as the sole carbon source. The more weakly hybridising clones (pTF101, pTF103 and pTF104) do not appear to contain an intact *hpcC* gene and are not able to repair the mutation in pJJ210.

Enzyme Activities Associated with the Clones.

The presence of enzymes which are responsible for the catalytic conversion of HPA to SSA and pyruvate were tested for by measuring their activities in cell free extracts of *E. coli* 5K containing the various clones after growth in LB-Ap liquid medium in the presence of HPA. Those clones which failed to complement any mutants were also tested as they may have contained genes of the HPC pathway other than those tested in the previous experiment. The plasmids pTF100 and pTF102 contain active *hpcBCDEFGH* genes whereas, pTF101, pTF103 and pTF104 contain no active *hpc* genes. Thus the latter three clones were eliminated from further investigation. If *hpc* genes have evolved by a process of retrograde evolution as proposed by Horowitz (1945; 1965) then the lower level of hybridisation exhibited by some clones may have been due to hybridisation of probe C to an ancestrally related sequence, in the absence of the *hpcC* gene.

The clones pTF100 and pTF102 contain the same *hpc* genes, but exhibit a different restriction pattern when treated with *Bam*H1. Each was grown in LB-Ap with and without the presence of HPA and cell free extracts made to determine

TABLE 4-5

ACTIVITIES ASSOCIATED WITH M5a1 AND 5K(pTF100/pTF102) UNDER NON-INDUCING AND INDUCING CONDITIONS

STRAIN	GROWTH	HPA	HPC	Specific Activity (μmol min ⁻¹ mg ⁻¹ protein)					
				CHMS	CHM	COHD	HHDD	OHED	HHED
M5a1	LB	0.010	0.002	0.004	0.004	0.010	0.025	0.007	+
M5a1	LB/HPA	0.176	0.395	0.165	0.115	0.053	0.130	0.100	+
5K(pTF100)	LB	n.d.	0.007	0.006	0.010	n.d.	n.d.	n.d.	n.d.
5K(pTF100)	LB/HPA	n.d.	0.090	0.035	0.020	0.013	0.023	0.020	+
5K(pTF102)	LB	n.d.	0.037	0.012	0.003	0.004	n.d.	n.d.	n.d.
5K(pTF102)	LB/HPA	n.d.	0.110	0.051	0.020	0.020	0.034	0.020	+

n.d.=not detected

whether the clones contained a regulatory gene and if they both behaved in a similar manner. Table 4.5 shows that in the absence of an inducer the HPC enzyme specific activities are present at very low values, but are increased when HPA was present in the growth medium. Both of the clones behaved in a similar manner, showing an inducible response to the presence of HPA.

Comparison of pTF100 and pTF102.

The clones pTF100 and pTF102 both express all of the enzyme activities necessary for the conversion of HPC to SSA and pyruvate; the genes which encode those enzymes are induced in the presence of HPA. The numerical values of the enzyme specific activities are low but this is similar to an analogous clones from *E. coli* C, pJJ200 (Jenkins, 1987; Jenkins and Cooper, 1988; this work, Table 3.7). However, the restriction pattern of the DNA of pTF100 and pTF102, when restricted with *Bam*H1, is significantly different. There are several possible explanations for this series of results: the first is that more than one copy of the *hpc* genes are present on the *K. pneumoniae* chromosome; alternatively the single chromosomal copy may have been cloned in both orientations, or a rearrangement in one of the clones, outside of the *hpc* genes, may have taken place resulting in a different restriction pattern but no effect on the expression of the genes. This was investigated by cutting the DNA from each with a variety of different restriction enzymes and electrophoresing the digests through a 0.8% agarose gel. Figure 4.16a depicts the resulting gel. The gel was arranged so that two digests with the same enzyme were next to each other to aid comparison; in each case one or more of the restriction fragments are in common and at least one fragment is different. The DNA from this gel was transferred to nitrocellulose and hybridised to radiolabelled probe C. The resulting autoradiograph is pictured in figure 4.16b. In each case the same restriction fragment hybridised to probe C except for the *Eco*R1 digested DNA's - in that case a 7kbp fragment of pTF100

FIGURE 4.16: Further analysis of pTF100 and pTF102: Photograph of an agarose gel and the resulting Southern blot of plasmids pTF100 and pTF102 after digestion with several enzymes. The lanes contained:

- | | |
|---|---|
| 1) pTF100 digested with <i>Bam</i> H1 | 2) pTF102 digested with <i>Bam</i> H1 |
| 3) pTF100 digested with <i>Eco</i> R1 | 4) pTF102 digested with <i>Eco</i> R1 |
| 5) pTF100 digested with <i>Sal</i> I | 6) pTF102 digested with <i>Sal</i> I |
| 7) pTF100 digested with <i>Hind</i> III | 8) pTF102 digested with <i>Hind</i> III |
| 9) pTF100 digested with <i>Eco</i> RV | 10) pTF102 digested with <i>Eco</i> RV |
| 11) pTF100 digested with <i>Pvu</i> II | 12) pTF102 digested with <i>Pvu</i> II |
| 13) pTF100 digested with <i>Pst</i> I | 14) pTF102 digested with <i>Pst</i> I |
| 15) λ -DNA digested with <i>Hind</i> III. | |

Exposure of the filter to X-ray film was for two hours at room temperature.

FIGURE 4.16a: RESTRICTION DIGEST PATTERNS OF pTF100 AND pTF102.

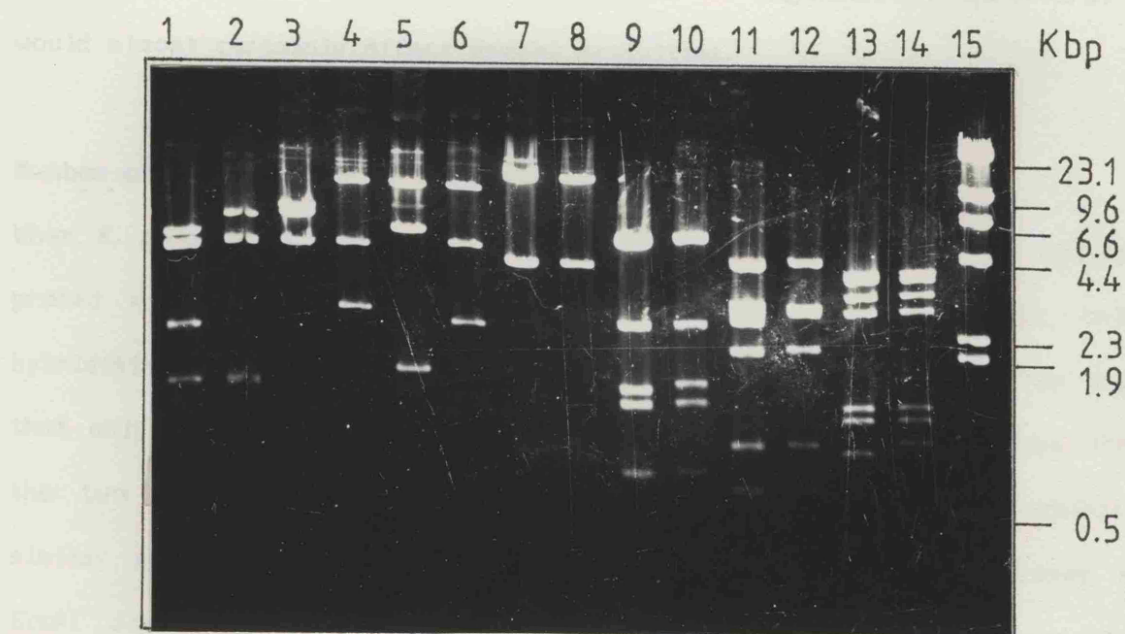
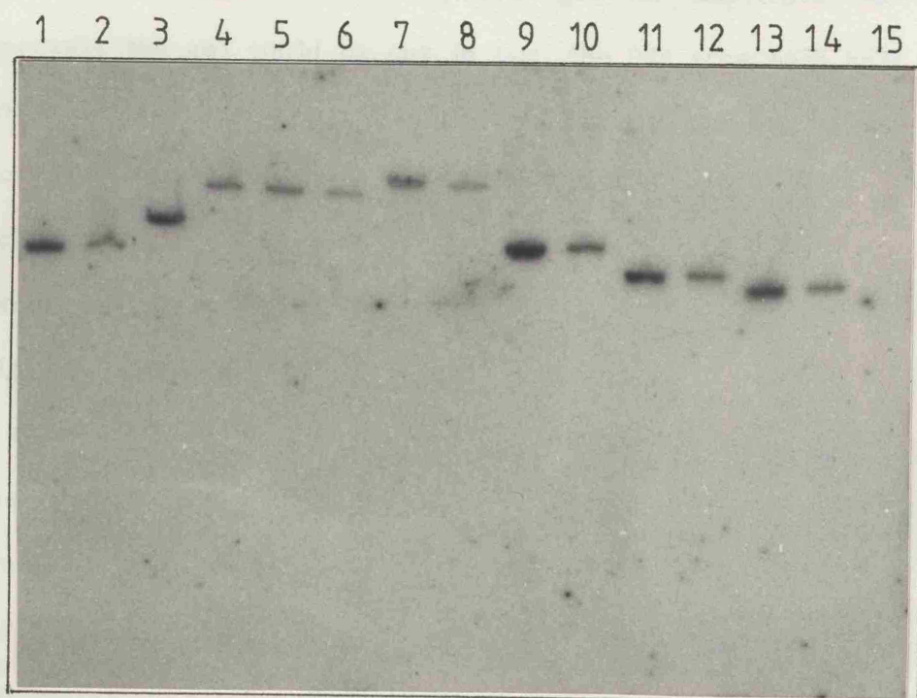


FIGURE 4.16b: SOUTHERN BLOT ANALYSIS OF THE ABOVE GEL SCREENED WITH PROBE C.



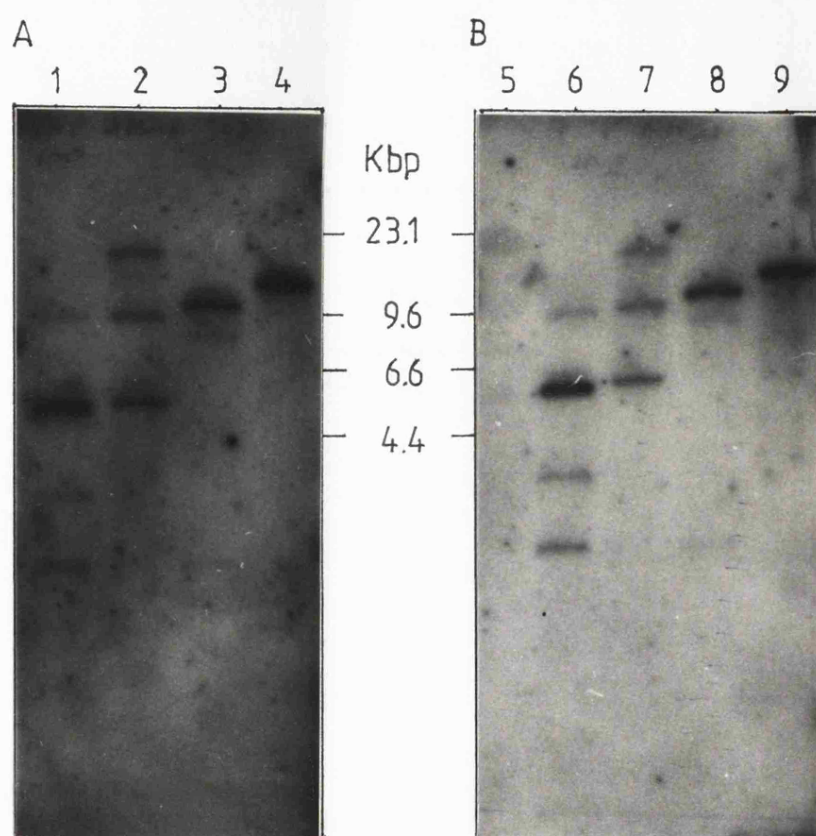
and a 10kbp fragment of pTF102 hybridised to the probe. These data suggest that a DNA rearrangement was an unlikely explanation for the different restriction patterns as disruption of the arrangement of structural genes would almost certainly affect enzyme activities.

Number of Copies of *hpc* Genes on the Chromosome.

When *K. pneumoniae* M5a1 DNA was restricted with several different enzymes and probed with radiolabelled probe C in a Southern blot experiment, only one hybridising band was seen in each track of the gel (fig. 4.12). This suggests that only a single copy of the genes are present on the chromosome. However, the two clones which possess HPC catabolic enzyme activities exhibit very similar restriction digestion patterns (fig. 4.16a) and in all cases except *Eco*R1 digestion, probe C hybridises to the same restriction fragment (fig. 4.16b). As *Eco*R1 was not used in the genomic Southern blot it was not possible to tell if one or two copies of the *hpc* genes were present on the chromosome from the present experimental evidence. To investigate this, *Klebsiella* DNA was digested with *Bam*H1, *Eco*R1, *Sal*I and *Hind*III and equal quantities of each electrophoresed through a 0.5% agarose gel in duplicate so that after electrophoresis the gel could be cut in two, the DNA from each half transferred to nitrocellulose and hybridised to the radiolabelled insert DNA of either pTF100 or pTF102 (fig. 4.17). As expected the pattern of hybridising bands in the tracks corresponding to *Bam*H1, *Sal*I and *Hind*III was the same in each experiment, however, the hybridising band in the tracks corresponding to the *Eco*R1 digested DNA's were also the same size. This experiment indicates that only a single copy of the *hpc* genes is present on the chromosome of *K. pneumoniae* M5a1 and that the clones pTF100 and pTF102 probably contain a simple inversion of the insert DNA. This can easily be verified by restricting the clones with *Hind*III and religating; if the idea of an inverted insert was true then individual clones following such an experiment would have the

FIGURE 4.17: Southern blot analysis of M5a1 DNA: Duplicate samples of restriction enzyme digested M5a1 chromosomal DNA were electrophoresed through a 0.5% agarose gel, the gel cut in two and the DNA from each half transferred to nitrocellulose. Plate A shows a photograph of a filter after hybridisation to radiolabelled insert DNA from pTF100 and plate B the second filter after hybridisation to radiolabelled insert DNA from pTF102. Lanes 1 and 6 contained DNA digested with *Bam*H1; 2 and 7 with *Eco*R1; 3 and 8 with *Sa*I1 and 4 and 9 with *Hind*III. Lane 5 contained *Hind*III digested λ -DNA.

FIGURE 4.17: SOUTHERN BLOT ANALYSIS OF M5a1 DNA HYBRIDISED TO pTF100 AND
pTF102 INSERT DNA.



restriction pattern of either pTF100 or pTF102, regardless of the source of DNA.

Oxidation of HPA and HPC by Whole Cells of 5K(pTF100).

Overnight cultures of *E. coli* 5K(pTF100) and *K. pneumoniae* M5a1 as a control, grown in LB, were each subcultured into three flasks containing either LB, LB+HPA or LB+HPC. The cultures were grown for approximately 4h in the presence of the inducers, harvested, washed twice in 0.1M phosphate buffer pH 7.5 and resuspended in the same buffer to an approximate concentration of 15mg ml⁻¹ dry weight (an OD₅₅₀ of 1.0 was assumed to be equivalent to 0.68mg dry weight ml⁻¹). The whole cells were tested for their ability to oxidise 4-HPA, 3-HPA and HPC at a concentration of 10mM. Those which did not oxidise any of these substrates were tested for their ability to oxidise glucose. When grown in the presence of HPA, *K. pneumoniae* was able to oxidise 4-HPA, 3-HPA and HPC; when grown in the presence of HPC it was able to oxidise HPC, but not 4-HPA or 3-HPA. *E. coli* 5K(pTF100) was not able to oxidise any HPA or HPC when grown in the presence of either substrate, but was able to oxidise glucose regardless of the growth conditions. Table 4.6 shows the results expressed as ng-atom O₂ mg⁻¹ h⁻¹.

Model of Regulation of *Klebsiella hpc* Genes.

The clones pTF100 and pTF102 both encode all of the enzyme activities necessary for the enzymic conversion of HPC to SSA and pyruvate. The enzymes are expressed at low rates in the absence of an inducer of the pathway and are expressed at higher rates in the presence of HPA. When transformed into a strain which does not encode any *hpa/hpc* enzymes (*E. coli* 5K) neither clone allows the organism to grow on HPC. This paradox is probably due to genes for the uptake of HPC not being present on pTF100. Some HPA must be able to enter the 5K cells in the absence of a specific uptake system as the enzymes are

TABLE 4.6: OXIDATION OF HPA AND HPC BY WHOLE CELLS.

ASSAY		GROWTH CONDITIONS				
SUBSTRATE	<i>K.pneumoniae</i> M5a1			<i>E. coli</i> 5K (pTF100)		
	LB	LB+HPA	LB+HPC	LB	LB+HPA	LB+HPC
4-HPA	n.d.	13	n.d.	n.d.	n.d.	n.d.
3-HPA	n.d.	16	n.d.	n.d.	n.d.	n.d.
HPC	n.d.	34	18	n.d.	n.d.	n.d.
GLUCOSE	79	N.T.	N.T.	75	81	88

n.d. =not detected

N.T. =not tested.

Specific activities expressed as ng-O₂ (mg dry weight)⁻¹.

expressed above the basal rate in the presence of HPA, however, this rate of uptake is presumably insufficient to support growth. This is supported by the fact that *E. coli* 5K(pTF100) whole cells cannot oxidise HPC (Table 4.6) after growth under conditions which enable HPC catabolic enzymes to be measured in cell free extracts (Table 4.5). The induced enzyme activities of 5K(pTF100) are low compared to those of M5a1 (Table 4.5); this may be due to incomplete induction of the *hpc* genes.

Klebsiella pneumoniae M5a1, when grown on LB in the presence of 4-HPA is able to oxidise 4-HPA, 3-HPA and HPC. When grown on LB in the presence of HPC it can oxidise HPC but not 3- or 4-HPA which indicates that the genes for the catabolism of HPA are present on at least two regulatory groups. One group for the conversion of HPA to HPC, inducible by HPA, and a second for the conversion of HPC to central metabolites. The second group of genes is inducible by HPC and HPA, as cloned *hpc* genes can be induced by HPA in the absence of HPA hydroxylase (Table 4.5). HPC is unable to induce the regulatory group containing the HPA hydroxylase. Therefore, it is unlikely that the gene encoding the HPC uptake system is present on the *hpa* operon otherwise when HPC was encountered in the environment it would be unable to induce its own uptake system. This suggests that at least three groups of genes are present in the *hpa/hpc* system.

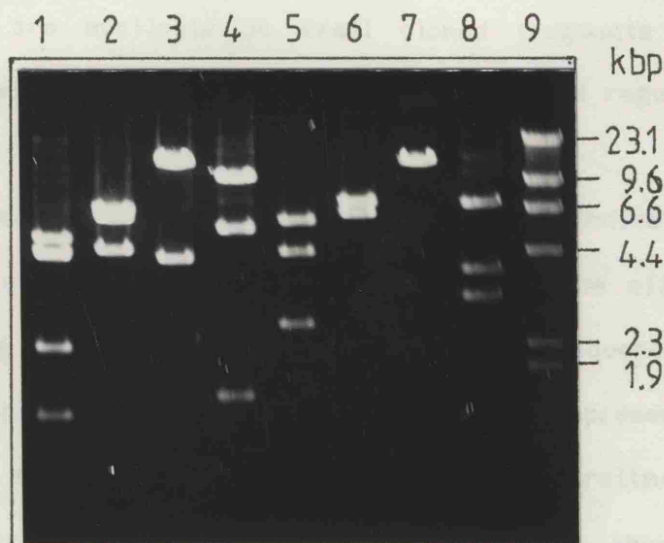
Comparison of Cloned *Escherichia* and *Klebsiella hpc* Genes.

The *Klebsiella* clone pTF100 appears to be analogous to the *E. coli* C clones pJJ200 and pJJ210 (Jenkins and Cooper, 1988). They show a regulated phenotype when present in *E. coli* 5K, express all of the enzymes necessary for the conversion of HPC to SSA and pyruvate but do not allow 5K to grow on HPC. The restriction patterns of the *Klebsiella* and *Escherichia* clones do not appear to share any common fragments (fig. 4.18) which along with the failure of cloned *E. coli* C genes and an oligonucleotide probe derived from the purified OHED

hydratase protein of *E. coli* C to hybridise to *K. pneumoniae* M5a1 DNA (Jenkins, 1987; Ferrer and Cooper, 1988) suggests that the *hpc* genes of these two organisms are generally not well conserved. The amino-terminal region of the two CHMS dehydrogenase enzymes are in stark contrast to this and may represent a region of significant structural or catalytic importance.

FIGURE 4.18: Comparison of pTF100 and pJJ200: Photograph of restriction analysis of plasmids pTF100 and pJJ200. Lanes 1-4 contained pTF100 DNA digested with 1, *Bam*H1; 2, *Eco*R1; 3, *Hind*III; 4, *Sal*I and lanes 5-8 contained pJJ200 DNA digested with 5, *Bam*H1; 6, *Eco*R1; 7, *Hind*III; 8, *Sal*I. Lane 9 contained *Hind*III digested λ -DNA.

FIGURE 4.18: RESTRICTION DIGESTS OF pTF100 AND pJJ200.



FUTURE PROSPECTS.

The present work has opened up a number of possibly interesting avenues for future research. Amongst the most interesting of these is the chance to investigate the conservation of the *hpc* pathway in *E. coli* C and *K. pneumoniae* M5a1. The genes for the conversion of HPC to SSA and pyruvate from both of these organisms are available on small cloned fragments of DNA and thus comparison of their gene order, nucleotide sequence and regulation have become possible.

A Southern blot analysis of several chromosomal DNAs indicated two regions of strong homology on the chromosome of *E. coli* W to the oligonucleotide probe constructed using information from the N-terminal sequence of the purified CHMS dehydrogenase proteins. It may be that these represent two functioning enzymes; these may have occurred by enzyme recruitment of the CHMS dehydrogenase into another pathway, by gene duplication, chromosome duplication or lateral transfer of a second homologous gene from another species. The cloning and investigation of these genes therefore would give some idea of the origin of the two hybridising bands and possibly an insight to the evolutionary mechanisms of catabolic pathways.

The position of an operator region on the *E. coli* C DNA has been implicated from this work; the presence of that region could be quickly confirmed by DNA sequencing and that region then used to purify the regulatory protein by affinity chromatography. A number of experiments using the purified protein to footprint the operator regions of both cloned sequences then becomes feasible.

REFERENCES.

- Adachi, K. Takeda, Y. Seno, S. and Kita, H. (1964). Metabolism of *p*-hydroxyphenylacetic acid in *Pseudomonas ovalis*. *Biochem. Biophys. Acta.* **93**: 483-493.
- Aiba, H., Adhya, S. and de Crombrughe, B. (1981). Evidence for two functional *gal* promoters in intact *Escherichia coli* cells. *J. Biol. Chem.* **256**: 11905-11910.
- Alonso, J. M. and Garrido-Pertierra, A. (1982). Carboxymethylhydroxymuconic semialdehyde dehydrogenase in the 4-hydroxyphenylacetate catabolic pathway of *Escherichia coli*. *Biochim. Biophys. Acta.* **719**: 165-167.
- Alonso, J. M. and Garrido-Pertierra, A. (1986). Kinetic properties of 5-carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase from *Escherichia coli*. *Biochimie* **68**: 731-737.
- Beck, C. F. and Warren, R. A. J. (1988). Divergent promoters, a common form of gene organisation. *Microbiol. Rev.* **52**: 318-326.
- Barbour, M. G. and Bayly, R. C. (1977). Regulation of *meta*-cleavage of 4-hydroxyphenylacetic acid by *Pseudomonas putida*. *Biochem. Biophys. Res. Commun.* **76**: 565-571.
- Belfaiza, J., Parsot, C., Martel, A., de la Tour, C. B., Margarita, D., Cohen, G. N. and Saint-Girons, I. (1986). Evolution in biosynthetic pathways: two enzymes catalysing consecutive steps in methionine biosynthesis originate from a common ancestor. *P. N. A. S.* **83**: 867-871.
- Blakley, E.R. (1972). Microbial conversion of *p*-hydroxyphenylacetic acid to homogentisic acid. *Can. J. Microbiol.* **18**: 1247-1255.
- Blakley, E.R., Kurz, W., Halverson, H., and Simpson, F.J. (1967). The metabolism of Phenylacetic acid by a *Pseudomonas*. *Can. J. Microbiol.* **13**: 147-152.
- Brewer, B. J. (1988). When polymerases collide: Replication and transcriptional organisation of the *E. coli* chromosome. *Cell.* **53**: 679-686.
- Buvinger, W.E. and Riley, M. (1985a). Nucleotide sequence of *Klebsiella pneumoniae lac* genes. *J. Bacteriol.* **160**: 850-857.
- Buvinger, W.E. and Riley, M. (1985b). Regulatory region of the divergent *Klebsiella pneumoniae lac* operon. *J. Bacteriol.* **163**: 858-862.

- Campbell, J. H., Lengyl, J. A. and Langridge, J. (1973). Evolution of a second gene for β -galactosidase in *Escherichia coli*. P. N. A. S. **70**: 1841-1845.
- Chow, L. T., Kahmann, R. and Kamp, D. (1977). Electron microscopy characterisation of DNAs of Mu. J. Mol. Biol. **113**: 591-609.
- Coleman, J., Green, P. J. and Inouye, M. (1984). The use of RNAs complementary to specific mRNAs to regulate the expression of individual bacterial genes. Cell **37**: 429-436.
- Cooper, R. A. and Skinner, M.A. (1980). Catabolism of 3- and 4-hydroxyphenylacetate by the 3,4-dihydroxyphenylacetate pathway in *Escherichia coli* C. J. Bacteriol. **143**: 302-306.
- Cove, D. J. (1985). Genetic similarities. Nature **314**: 44.
- Dagley, S. (1975). A Biochemical approach to some problems of environmental pollutin. Essays in biochemistry. **11**: 81-138.
- Donnelly, M. I. and Cooper, R. A. (1981). Two succinate semialdehyde dehydrogenases are induced when *Escherichia coli* K-12 is grown on γ -aminobutyrate. J. Bacteriol. **145**: 1425-1427.
- Doten, R. C., Ngai, K-L., Mitchell, D. J., and Ornston, L. N. (1987). Cloning and genetic organisation of the *pca* gene cluster from *Acinetobacter calcoaceticus*. J. Bacteriol. **169**: 3168-3174.
- Downing, R. G. and Broda, P. (1979). A cleavage map of the TOL plasmid of *Pseudomonas putida* mt-2. Mol. Gen. Genet. **177**: 189-191.
- Fawcett, T., Garrido-Pertierra, A. and Cooper, R.A. (1989). 5-Carboxymethyl-2-hydroxy-muconic semialdehyde dehydrogenases of *Escherichia coli* C and *Klebsiella pneumoniae* M5a1 show very high N-terminal homology. FEMS Microbiol. Lett. **57**: 307-312.
- Ferrer, E. and Cooper, R. A. (1988). Studies with a cloned *Escherichia coli* C 2-oxo-hept-3-ene-1,7-dioate hydratase gene. FEMS Microbiol. Lett. **52**: 155-160.
- Feinberg, A. P. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specificity. Anal. Biochem. **132**: 6-13.
- Franklin, F. C. H., Bagdasarian, M. M. and Timmis, K. N. (1981). Molecular and functional analysis of the Tol plasmid pWVO from *Pseudomonas putida* and cloning of the genes for the entire regulated aromatic *meta* cleavage pathway. P. N. A. S. **78**: 7458-7462.

- Garrido-Pertierra, A. and Cooper, R. A. (1981). Identification and purification of distinct isomerase and decarboxylase enzymes involved in the 4-hydroxyphenylacetate catabolic pathway of *Escherichia coli*. *Eur. J. Biochem.* **117**: 581-584.
- Grossman, A. D., Erickson, J. W. and Gross, C.A. (1984). The *htpR* gene product is a sigma factor for the heat shock promoters. *Cell* **38**: 383-390.
- Grunstein, M. and Hogness, D. S. (1975). Colony hybridisation: a method for the isolation of cloned DNAs that contain a specific gene. *P. N. A. S.* **72**: 8123-8136.
- Harayama, S., Lehrbach, P. R. and Timmis, K. N. (1984) Transposon mutagenesis analysis of *meta*-cleavage pathway operon genes of the TOL plasmid of *Pseudomonas putida* mt-2. *J. Bacteriol.* **160**: 251-255.
- Harayama, S., Leppik, R. A., Rekik, M., Mermod, N., Leirbach, P. R., Reineke, W. and Timmis, K. N. (1986). Gene order of the TOL catabolic plasmid upper pathway operon and oxidation of both toluene and benzyl alcohol by the *xyIA* product. *J. Bacteriol.* **167**: 455-461.
- Harayama, S., Mermod, N., Rekik, M., Leirbach, P. R. and Timmis, K. N. (1987). Roles of the divergent branches of the *meta*-cleavage pathway in the degradation of benzoate and substituted benzoates. *J. Bacteriol.* **169**: 558-564.
- Har-El, R., Silberstein, A., Kuhn, J. and Tal, M. (1979). Synthesis and degradation of *lac* mRNA in *E. coli* depleted of 30S ribosomal subunits. *Mol. Gen. Genet.* **173**: 135-144.
- Hareland, W. A., Crawford, R.L., Chapman, P. J. and Dagley, S. (1975). Metabolic function and properties of 4-HPA-1hydroxylase from *Pseudomonas acidovorans*. *J. Bacteriol.* **121**: 275-285.
- Hattori, M. and Sakiki, Y. (1985). Dideoxy sequencing method using denatured plasmid templates, *Anal. Biochem.* **152**: 232-238.
- Hegeman, G. D. and Rosenberg, S. L. (1970) The evolution of bacterial enzyme systems. *Annu. Rev. Genet.* **24**: 429-462.
- Higgins, C. F., Ames, G. F. L., Barnes, W. M., Clement, J. M. and Hofnung, M. (1982). A novel intracistronic regulatory element of procaryotic operons. *Nature* **298**: 760-762.
- Higgins, C. F. and Smith, N. H. (1985). In: Regulation of gene expression-25 years on. eds. I. R. Booth and C. F. Higgins. Cambridge University press.

- Holmes, W. M., Platt, T. and Rosenberg, M. (1983). Termination of transcription in *E. coli*. *Cell* **32**: 1029-1032.
- Horowitz, N. H. (1945). On the evolution of biochemical synthesis. *P. N. A. S.* **31**: 153-157.
- Horowitz, N. H. (1965). The evolution of biochemical syntheses-retrospective and prosoective. In: *Evolving genes and proteins*. Eds. Bryson, V. and Vogel, H. J. New York: Academic Press. 15-23.
- Inouye, S., Nakazawa, A. and Nakazawa, T. (1981a). Molecular cloning of the TOL genes *xylB* and *xylE* in *Escherichia coli*. *J. Bacteriol.* **145**: 1137-1143.
- Inouye, S., Nakazawa, A. and Nakazawa, T. (1981b). Molecular cloning of gene *xylS* of the TOL plasmid: Evidence for positive regulation of the *xylDEGF* operon by *xylS*. *J. Bacteriol.* **148**: 413-418.
- Inouye, S., Nakazawa, A. and Nakazawa, T. (1986). Nucleotide sequence of the regulatory gene *xylS* on the *Pseudomonas putida* TOL plasmid and identification of the protein product. *Gene* **44**: 235-242.
- Jenkins, J. R. (1987) Doctoral Thesis, University of Leicester.
- Jenkins, J. R. and Cooper, R. A. (1988). Molecular cloning, expression and analysis of the genes of the homoprotocatechuate catabolic pathway of *Escherichia coli* C. *J. Bacteriol.* **170**: 5317-5324.
- Keil, H., Keil, S. and Williams, P. A. (1987). Molecular analysis of regulatory and structural *xyl* genes of the TOL plasmid pWW53-4. *J. Gen. Microbiol.* **133**: 1149-1158.
- Konigsberg, W. and Godson, G. N. (1983). Evidence for use of rare codons in the *dnaG* gene and other regulatory genes of *Escherichia coli*. *P. N. A. S.* **80**: 687-691.
- Kushner, S. R. (1978). An improved method for the transformation of *Escherichia coli* with ColE1-derived plasmids. In *Genetic Engineering*. Elsevier, Holland. eds. Boyer, H.B. and Nicosia, S.
- Lewis, E. B. (1951). Pseudoallienism and gene evolution. *Cold Spring Harbour Symp. Quant. Biol.* **16**: 159-174.
- Lowry, O., Rosenbrough, N. J., Farr, R. J. and Randall, R. (1951). Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**: 265-273.

- Malan, T. P., Kolb, A., Buc, H. and McClure, W. R. (1984). Mechanism of CRP-cAMP activation of the *lac* operon transcription initiation: activation of the P1 promoter. *J. Mol. Biol.* **180**: 881-909.
- Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982). eds. In: *Molecular cloning a laboratory manual*. Cold Spring Harbor laboratory, Cold Spring Harbor, New York.
- Miller, J. H. (1972). ed. In: *Experiments in Molecular genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Muller-Hill, B. and Kania, J. (1974). Lac repressor can be fused to B-galactosidase. *Nature*. **249**: 561-563.
- Nazakawa, T. and Yokoto, T. (1973). Benzoate metabolism in *Pseudomonas putida* (arvilla) mt-2: demonstration of two benzoate pathways. *J. Bacteriol.* **115**: 262-267.
- Neidhardt, F.C. (1987) In: *Escherichia coli* and *Salmonella typhimurium* cellular and molecular biology. Vol II. ed. F.C. Neidhardt. American society of microbiology, Washington D.C.
- Neidle, E. L. and Ornston, L. N. (1986). Cloning and expression of *Acinetobacter calcoaceticus* catechol 1,2-dioxygenase structural gene *catA* in *Escherichia coli*. *J. Bacteriol.* **168**: 815-820.
- Neidle, E. L., Shapiro, M. K. and Ornston, L. N. (1987). Cloning and expression in *Escherichia coli* of *Acinetobacter calcoaceticus* genes for benzoate degradation. *J. Bacteriol.* **169**: 5496-5503.
- Postma, P. W. (1986). Catabolic repression and related processes. In: *Regulation of gene expression-25 years on*. ed. I.R. Booth and C.F. Higgins. Cambridge University Press.
- Pramanik, A., Wertheimer, S. J., Schwartz, J. J. and Schwartz, I. (1986). Expression of *Escherichia coli* *infC*: Identification of a promoter in an upstream *thrS* coding sequence. *J. Bacteriol.* **168**: 746-751.
- Ribbons, D. W. and Eaton, R. W. (1982). Chemical transformations of aromatic hydrocarbons that support growth of microorganisms. p59-84. In: *Biodegradation and detoxification of environmental pollutants*. ed. A. M. Chakrabarty. CRC press, Boca Raton, Fla.
- Riley, M. and Anilionis, A. (1978). Evolution of the bacterial genome. *Ann. Rev. Microbiol.* **32**: 519-560.

- Sala-Trepat, J. M. and Evans, W. C. (1971). The *meta* cleavage of catechol by *Azotobacter* species using the 4-oxalocrotonate pathway. *Eur. J. Biochem.* **200**: 400-413.
- Schumperli, D., McKenney, K., Sobiesky, D. A. and Rosenberg, M. (1982). Translational coupling at the intracistronic boundry of the *Escherichia coli* galactose operon. *Cell* **30**: 865-871.
- Shanley, M. S., Neidle, E. L., Parales, R. E. and Ornston, L. N. (1986). Cloning and Expression of *Acinetobacter calcoaceticus* catBCDE genes in *Pseudomonas putida* and *Escherichia coli*. *J. Bacteriol.* **165**: 557-563.
- Sinclair, M. I., Maxwell, M. C., Lyon, B. R. and Holloway, B. W. (1986). Chromosomal location of TOL plasmid DNA in *Pseudomonas putida*. *J. Bacteriol.* **168**: 1302-1308.
- Skinner, M. A. (1981). Aromatic catabolism by *Escherichia coli*. Ph.D. Thesis, Leicester University, England.
- Skinner, M. A. and Cooper, R.A. (1982). An *Escherichia coli* mutant defective in the NAD-dependant succinate dehydrogenase. *Arch. Microbiol.* **132**: 270-275.
- Soberon, X., Covarrubias, L. and Bolivar, F. (1980). Construction and characterisation of new cloning vehicles IV. Deletion derivatives of pBR322 and pBR325. *Gene* **9**: 287-305.
- Southern, E. M. (1980). Gel electrophoresis of restriction fragments. *Methods in Enzymol.* **68**: 152-176.
- Sparnins, V. L., Chapman, P. J. and Dagley, S. (1974). Bacterial degradation of 4-hydroxyphenylacetic acid and homoprotocatechuic acid. *J. Bacteriol.* **120**: 159-167.
- Spoelstra, S. F. (1978). Degradation of tyrosine in anaerobically stored piggery wastes and in pig feces. *Appl. Environ. Microbiol.* **36**: 631-638.
- Stanier, R. Y. (1947). Simultaneous adaptation: a new technique for the study of metabolic pathways. *J. Bacteriol.* **54**: 339-348.
- Stanier, R. Y., Palleroni, N. J. and Doudoroff, M. (1966). The aerobic *Pseudomonas* a taxonomic study. *J. Gen. Microbiol.* **43**: 159-160.
- Stanier, R. Y. and Ornston, L. N. (1973). β -ketoadipate pathway. *Adv. Microbial. Phys.* **9**: 89-151.
- Stern, M. J., Ames, G. F. L., Smith, N. H., Robinson, E. C. and Higgins, C. F. (1984). Repetative extrtagenic palindromic sequences: a majour component of the bacterial genome. *Cell* **37**: 1015-1056.

- Steitz, J. A. (1979). Genetic signals and nucleotide signals in messenger RNA. In Biological regulation and development. 1. Gene expression, ed. Goldberg, R. F. pp349-399. Plenum press, New York.
- Sudjadi, J. (1988). Personal communication.
- Suggs, S. V., Hirose, T., Miyake, T., Kuwashima, M. J., Johnson, T., Itakura, K. and Wallace, R. B. (1981). In: Developmental biology using purified genes. eds. D. Browe and C. F. Fox, Academic Press, New York, 1981. pp683-693.
- Ullmann, A. and Danchin, A. (1983). Role of cAMP in bacteria. Advances in Nucleotide Research. **15**: 2-53.
- Walmsley, A. R. and Lowe, A. G. (1985). MULTIFIT; A flexible non-linear least square regression programme in BASIC. Comp. Meth. Prog. Biomed. **21**: 113-118.
- Warburg, O. and Christian, W. (1941) Isolation and characterisation of enolase. Biochem. Z. **310**: 384-421.
- Wetmar, R. and Davidson, N. (1968). Kinetics of renaturation of DNA. J. Mol. Biol. **31**: 349-370.
- Williams, P. A. and Murray, K. (1974). Metabolism of benzoates and the methylbenzoates by *Pseudomonas putida* (arvilla) mt-2: evidence for the existence of a TOL plasmid. J. Bacteriol. **120**: 416-423.
- Winkler, M. E. and Yanofsky, C. (1981). Pausing of RNA polymerase during the *in vitro* transcription of the tryptophan operon leader sequence. Biochem. **20**: 3738-3744.
- Winstanely, C., Taylor, S. C. and Williams, P. A. (1987). pWW174: A large plasmid from *Acinetobacter calcoaceticus* encoding benzene catabolism by the β -ketoadipate pathway. Mol. Microbiol. **1**: 219-227.
- Wong, C. L. and Dunn, N. W. (1974). Transmissible plasmid coding for the degradation of benzoate and *m*-toluate in *Pseudomonas arvilla* mt-2. Genet. Res. **23**: 227-232.
- Worsey, M. J., Franklin, F. C. H. and Williams, P. A. (1978). Regulation of the degradative pathway enzymes encoded by the TOL plasmid (pWWO) from *Pseudomonas putida* mt-2. J. Bacteriol. **134**: 757-764.
- Yanisch-Perron, C., Vierra, J. and Messing, J. (1985). Improved M13 cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene **33**: 103-119.

Yanofsky, C. (1981). Attenuation in the control of expression of bacterial operons. *Nature*. **289**: 751-758.

Yates, J. L. and Nomura, M. (1981). Feedback regulation of ribosomal protein synthesis in *E. coli*: localisation of the mRNA target sites for receptor action of ribosomal protein L1. *Cell* **24**: 243-249.

Yeh, W. K., Davis, G., Fletcher, P. and Ornston, L. M. (1978). Homologous amino acid sequences in enzymes mediating sequential metabolic reactions. *J. Biol. Chem.* **253**: 4920-4923.

Yeh, W. K., Fletcher, P., and Ornston, L. N. (1980). Evolutionary divergence of co-selected β -ketoadipate enol-lactone hydrolases in *Acinetobacter calcoaceticus*. *J. Biol. Chem.* **255**: 6342-6346.

Yeh, W. K. and Ornston, L. N. (1980). Origins of metabolic diversity: substitution of homologous sequences into genes for enzymes with different catalytic activities. *P. N. A. S.* **77**: 5365-5369.

ABSTRACT.

Analysis of Cloned Genes for Aromatic Catabolism: the Hydroxyphenylacetate Genes of *Escherichia coli* and *Klebsiella pneumoniae*.

Tony Fawcett.

Cloned genes, from *Escherichia coli* C, for the catabolism of homoprotocatechuate were available. The organisation and expression of these genes was investigated by a number of molecular genetic techniques. The exact position of the 5' end of one of the genes, *hpcG*, was determined by a combination of Southern blotting and DNA sequencing and subsequently used as a marker to aid the positioning and elucidation of the direction of transcription of the other genes. The genes were found to be arranged in two blocks of structural genes which were both transcribed in the same direction and a separate regulatory gene. Northern hybridisation analysis of RNA from wild type cells and from cells containing cloned genes indicated the presence of three transcripts of 4.5, 2.7 and 1.6 kbp. A model for the position of the *hpc* genes, their regulation and number of transcriptional units is proposed.

One of the catalytic proteins, CHMS dehydrogenase, was purified and some of its properties investigated; homology between the amino-terminal amino acid sequence of this enzyme and the equivalent protein from *Klebsiella pneumoniae* M5a1 was very high. An oligonucleotide derived from shared protein sequence of the CHMS dehydrogenases from *E. coli* C and *K. pneumoniae* M5a1 was used to investigate homology in related species by Southern hybridisation. *E. coli* strains B and W DNA contained a hybridising fragment when tested at high stringency, but *Pseudomonas putida* DNA did not hybridise under these conditions. The same oligonucleotide was used to screen a *K. pneumoniae* M5a1 genomic library by colony hybridisation and clones carrying *hpc* genes were successfully isolated. Initial analysis of these clones suggest that little gross homology exists between the *hpc* genes of *E. coli* C and *K. pneumoniae* M5a1.