



Durham E-Theses

*The positive control of *ilvC* expression in *E. coli* K-12*

Ahmad, Khalid Dham

How to cite:

Ahmad, Khalid Dham (1989) *The positive control of *ilvC* expression in *E. coli* K-12*, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/6314/>

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

The copyright of this thesis rests with the author.
No quotation from it should be published without
his prior written consent and information derived
from it should be acknowledged.

**THE POSITIVE CONTROL OF *ilvC* EXPRESSION
IN *E. coli* K-12**

A thesis submitted to the
University of Durham
for the degree of
Doctor of Philosophy

by

KHALID DHAM AHMAD

Department of Biological Sciences
University of Durham

1989



31 OCT 1990

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Martin D. Watson who introduced me to this project and never hesitated in advising me during the work. This work is never performed unless continuous encouragements.

All the workers in the groups of Dr. C. Shaw and Dr. D. Murphy were helpful, in particular Dr. A. Ryan, who continuously gave his advice throughout this work, and I. Cummins. Thanks to Dr. N. Robinson for his recommendations during the experiments on gel filtration.

The friendly atmosphere provided by the workers in the same laboratory, Alan, Catherine, Christine and Gary.

Many people in this department made it possible for me to complete this work, in particular Prof. D. Boulter (the head of department) who allowed me to use all of the equipment and material available in the department.

Finally, I should like to thank my government for the scholarship which enables me to study this very interesting field of science.

The positive control of *ilvC* expression in *E. coli* K-12

Khalid D. Ahmad

ABSTRACT

The mechanism of *ilvC* expression in *Escherichia coli* was investigated. To carry out this work several different approaches were used. Firstly, sequencing of the *ilvY*2143 allele which carries a mutation that makes *ilvC* expression constitutive was completed. The location of the mutation was determined to be at the 5' end of the gene. It is a single base substitution (G to A) at position 87 (counted from the transcription startpoint of *ilvY*). This results in a change of the codon for one amino acid. Glutamine in wild-type *ilvY* protein is replaced by lysine in the constitutive one. This substitution in the polypeptide of the upsilon protein (product of the *ilvY* gene) was found to be solely responsible for making the upsilon protein independent of the *ilvC* gene substrates (α -acetoxybutyrate or acetolactate) needed for *ilvC* induction.

Two approaches were used to determine the direction of *ilvY* transcription. One of these employed a gene fusion technique which involves two DNA fragments of *ilvY* being fused separately to a promoterless *lacZ* gene, then monitoring the expression of *lacZ*. The other approach involved the labeling of the upsilon protein with ^{35}S -methionine after expression of *ilvY* in a T7 RNA polymerase dependent promoter system.

DNA-binding activity of upsilon protein was investigated. This was carried out in two assays, filter binding and gel retardation assays. These assays were employed to monitor purification of upsilon protein to near homogeneity. Upsilon protein has a subunit size of 35 kd and a native molecular weight of approximately 211 kd, suggesting upsilon exists as a hexamer.

Finally, *in vitro* activities of the upsilon protein were tested using transcriptional and coupled transcription-translation assays. Upsilon protein was shown to cause elevation of *ilvC* transcription.

Two models for the action of the upsilon protein in regulating the transcription of the *ilvYC* are proposed.

DECLARATION

I declare that no material in this thesis has previously been submitted for a degree at this or any other university.

The copyright of this thesis rests with the author. No quotation from it should be published without prior written consent and the information derived from it should be acknowledged.

CONTENTS

Contents

	Chapter 1 Introduction.	1
1.1	Biosynthesis of isoleucine, valine and leucine in <i>E. coli</i> .	1
1.2	Feedback or end-product inhibition.	2
1.3	Feedback inhibition for isoleucine, valine and leucine biosynthesis in <i>E. coli</i> .	3
1.4	Regulation of gene expression.	3
1.5	The Lactose operon.	4
1.6	Catabolite repression.	5
1.7	The Arabinose operon.	6
1.8	The galactose operon.	7
1.9	The tryptophan operon (A biosynthetic operon).	8
1.10	Gene-Enzyme relationships in the isoleucine-valine system in <i>E. coli</i> .	10
1.11	Physical maps of the <i>ilv</i> genes.	11
1.12	Regulation of the <i>ilv</i> operons.	12
1.13	Regulation of the <i>ilvBN</i> operon.	12
1.14	Regulation of the <i>ilvGMEDA</i> operon.	14
1.15	Regulation of the <i>ilvIH</i> operon.	16
1.16	Regulation of the <i>ilvYC</i> system.	16
1.17	The filter binding assay.	18
1.18	The Gel retardation assay.	19
1.19	Gene fusion as a tool for genetic analysis.	20

Abbreviations

- ATP** Adenosine 5'-triphosphate.
- C.P.M.** Count per minute.
- DEAE-cellulose** Diethylaminoethylcellulose.
- DNA** Deoxyribonucleic acid.
- DNase** Deoxyribonuclease.
- DTT** Dithiothreitol.
- dNTP** Deoxynucleoside triphosphate.
- ddNTP** Dideoxynucleoside triphosphate.
- DMF** Dimethylformamide.
- EDTA** Ethylenediaminetetraacetic acid.
- IPTG** Isopropylthiogalactoside.
- MOPS** (3-[N-Morpholino]propanesulfonic acid).
- RNA** Ribonucleic acid.
- mRNA** Messenger Ribonucleic acid.
- tRNA** Transfer Ribonucleic acid.
- RNase** Ribonuclease.
- TCA** Trichloroacetic acid.
- TEMED** N,N,N',N'-tetramethylethylenediamine.
- Tris** Tris(hydroxymethyl)amionmethane.
- SDS** Sodium Dodecyl Sulphate.
- UV** Ultraviolet.
- X-gal** 5-bromo-4-chloro-3-indolyl- β -D-galactoside.

1.20	The T7 expression system.	21
	Chapter 2 Materials and Methods.	23
2.1	Materials.	23
2.2	Preparation of Media.	24
2.3	Preparation of the buffers and common solutions used in this work.	25
2.4	Sterilization.	30
2.5	Extraction with Phenol.	30
2.6	Precipitation of DNA.	30
2.7	Preparation of the Dialysis tubing.	31
2.8	Spectrophotometric quantation of DNA.	31
2.9	Digestion of DNA with restriction enzymes.	31
2.10	Preparation of restriction enzyme buffers.	32
2.11	Agarose gel electrophoresis.	32
2.12	Acrylamide gels for retardation assays.	33
2.13	Preparation of SDS-PAGE gel.	33
2.14	Silver staining method for SDS-PAGE gels.	34
2.15	Protein determination.	36
2.16	Purification of DNA restriction fragments from low melting point agarose gels.	36
2.17	Dephosphorylation of DNA fragments.	36
2.18	ligation of DNA fragments.	37
2.19	Transformation and preparation of competent cells.	37
2.20	Preparation of Sephadex G-25 and G-50 column for the separation of DNA from unincorporated nucleotides.	38
2.21	Nick translation of DNA restriction fragments.	38
2.22	P ³² Labelling of restriction fragment by	39

	“filling in”.	
2.23	Labelling of restriction fragments by T4 polynucleotide kinase.	39
2.24	Northern Blotting.	40
2.25	Preparation of plasmid DNA.	40
2.26	Transformation of recombinant M13mp18 and M13mp19.	43
2.27	Preparation of single stranded M13mp DNA.	43
2.28	DNA sequencing.	44
2.29	Preparation of DNA sequencing gels and electrophoresis of samples.	45
2.30	Deoxynucleoside triphosphate (dNTP) and dideoxynucleoside triphosphate (ddNTP) mixes.	46
2.31	Preparation of S-30 extracts.	48
2.32	Solutions and media recipes for <i>in vitro</i> <i>ilvC</i> expression.	48
2.33	Bacterial strains and plasmids.	49
	Chapter 3 Characterization of the <i>ilvY^c</i> mutation.	50
3.1	Introduction.	50
3.2	Does epsilon act at the transcriptional level?.	50
3.3	Nucleotide sequencing of the <i>ilvY</i> gene.	53
3.4	Examination of the <i>ilvY</i> nucleotide sequences.	54
3.5	Is the single base change responsible for the constitutive phenotype?.	54
3.6	Discussion.	57
	Chapter 4 Transcription of <i>ilvY</i>.	60
4.1	Introduction.	60
4.2	Fusion of <i>ilvY</i> with <i>lacZ</i> .	60

4.3	Expression of <i>ilvY</i> in T7 system.	62
4.4	Preparation of radiolabelled extracts from strains A and B.	64
4.5	Discussion.	65
	Chapter 5 Purification and characterization of upsilon protein.	69
5.1	Introduction.	69
5.2	Purification of upsilon protein.	69
5.3	Gel retardation assays of the three ammonium sulphate fractions.	70
5.4	<i>in vitro</i> transcription of <i>ilvC</i> .	71
5.5	Filter binding assays of upsilon protein.	72
5.6	Purification by DEAE-cellulose column chromatography.	72
5.7	DNA cellulose chromatography.	73
5.8	Purification of radioactive upsilon protein.	75
5.9	Purification of upsilon protein by gel filtration	76
5.10	Results for gel filtration column.	77
5.11	Retardation pattern of various DNA fragments of the <i>ilvYC</i> genetic system.	78
5.12	Examination of the <i>in vitro</i> properties of upsilon protein.	79
5.13	Results for <i>in vitro</i> transcription of <i>ilvC</i> .	80
5.14	<i>in vitro</i> transcription-translation of <i>ilvC</i> .	80
5.15	Discussion.	82
	Chapter 6 General discussion.	86
	References.	96

List Of Figures

Figures

- 1.1 Isoleucine-valine biosynthetic pathway.
- 1.2 Leucine biosynthetic pathway.
- 1.3 The important features of leader the mRNA in the tryptophan operon.
- 1.4 The major *ilv* cluster.
- 1.5 The alternative secondary structures of the *ilvBN* leader mRNA.
- 1.6 The possible secondary structures of the leader mRNA in the *ilvGEDA* operon.
- 3.1 A. 1% Agarose-formaldehyde gel electrophoresis for separation of the RNA species extracted from CU152, CU827 and CU962 under different growth conditions.
B. Autoradiograph showing the results for the hybridization experiment.
- 3.2 Restriction endonuclease map of the *ilvYC* system.
- 3.3 0.7% agarose gels electrophoresis of cleaved pGMM201, pP1, pH3 and pH4 digested with EcoRI and of pDUB2400 and pDUB2401 digested with EcoRI and BglII.
- 3.4 Nucleotide sequence of the *ilvY* gene and the deduced amino acids sequence.
- 3.5 First strategy for the attempt to verify that the single base change in the small *ilvY* fragment is responsible for making the *ilvC* expression constitutive.

- 3.6** A. 0.7% agarose gel electrophoresis of pDUB2402 and pDUB2403 cleaved with EcoRI.
B. Digestion of pDUB2400 and pDUB2401 with HindIII and BglII.
C. pDUB2404 and pDUB2405 digested with EcoRI.
- 3.7** Second strategy for the verification that the small EcoRI-BglII *ilvY* fragment, is solely responsible for the constitutive phenotype.
- 4.1** The construction of the *ilvY-lacZ* fusions.
- 4.2** The two orientations of *ilvY* cloned in pBR328.
C. pMC1871 restricted with BamHI.
- 4.3** Restriction map of the fusion vector pMC1871.
- 4.4** 0.7% agarose gel electrophoresis showing the correct and incorrect fusion of *lacZ* to the large BamHI-BglII fragments of pDUB2415 and pDUB2414.
- 4.5** Illustrating the two possible orientations of *ilvY* in pGEM-Blue.
- 4.6** 0.7% agarose gels electrophoresis of plasmid DNA of transformed colonies, screening for the correct orientation *ilvY* fragment in pGEM-Blue.
in pGEM-Blue.
- 4.7** 10% SDS-PAGE gel of crude extracts from strain A and B after expressing the *ilvY* gene in the T7 system.
- 4.8** Results of electrophoresis of labelled proteins using 10% SDS-PAGE gel.
- 5.1** Gel-retardation assays of the three ammonium sulphate fractions of strain C.
- 5.2** Agarose gels electrophoresis for the purification

- of various DNA fragments of *ilvYC* system.
- 5.3** Gel retardation assay for fractions eluted from DNA-cellulose column.
- 5.4** DEAE-cellulose column chromatography.
- 5.5** DNA-cellulose chromatography.
- 5.6** 10% SDS-PAGE gel for demonstrating the purity of peak fractions eluted from the DNA-cellulose column.
- 5.7** Void volume of the gel filtration column after applying 2 mg Dextran Blue.
- 5.8** Calibration of the gel filtration column with standard protein markers.
- 5.9** Calibration curve for estimating the native protein molecular weight.
- 5.10** Gel filtration column chromatography of labelled epsilon protein.
- 5.11** 10% SDS-PAGE gel of peak fractions eluted from the gel filtration column.
- 5.12 and 5.13** Gel retardation assay of fractions eluted from the gel filtration column.
eluted from gel filtration column.
- 5.14** Autoradiograph of a 10% SDS-PAGE gel of labelled proteins made in a cell free coupled transcription-translation system.
- 6.1** Nucleotide sequence of the EcoRI-HincII fragment of *ilvY* showing the the -10 regions of the *ilvYC* genes.

List Of Tables

Tables

- 2.1 Bacterial strains.
- 2.2 Plasmids.
- 2.3 Constructed plasmids used in this work.
- 3.1 Trans-activation of *ilvC::lacZ* expression on MacConkey agar plates.
- 5.1 Percentage saturation of ammonium sulphate.
- 5.2 *In vitro* transcription of *ilvC* using the three ammonium sulphate fractions of strain C.
- 5.3 *In vitro* transcription of *ilvC* using fractions eluted from a DEAE-cellulose column after application of the 30% ammonium sulphate fraction of strain C.
- 5.4 Filter binding assays of ammonium sulphate of wild-type, deleted and constitutive strains.
- 5.5 Filter binding assays of DEAE-cellulose column fractions.
- 5.6 Filter binding assays of DNA-cellulose column fractions.
- 5.7 DEAE-cellulose chromatography of labelled epsilon protein.
- 5.8 DNA-cellulose chromatography.
- 5.9 *In vitro* transcription of the plasmid pDUB2416 and the linear PstI-BglII DNA fragment.
- 5.10 Showing the constituents of various reaction samples used in *in vitro* expression of *ilvC*.

CHAPTER 1

INTRODUCTION

1.1 Biosynthesis of isoleucine, valine and leucine in *Escherichia coli*.

Isoleucine, valine and leucine are branched amino acids. They contain aliphatic, hydrophobic R groups. The synthesis of all three amino acids in bacteria starts from pyruvate.

The biosynthesis of isoleucine and valine has been studied extensively in *E. coli* and many other organisms, in which the biosynthetic pathways appear to be identified. This biosynthesis involves parallel pathways in which a series of bifunctional enzymes catalyze the conversion of two pyruvate molecules into valine and also the last four steps of isoleucine synthesis.

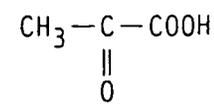
The first step in this pathway is the conversion of threonine to α -ketobutyrate and this step is unique to isoleucine biosynthesis, (see Fig.1.1). It is catalyzed by the enzyme threonine deaminase. In *E. coli*, threonine deaminase has been purified and shown to have a native enzyme molecular weight of 212,000. It is a tetramer composed of four identical subunits of molecular weight 53,000 (Calhoun *et al.*, 1973).

The first step common to the biosynthesis of both isoleucine and valine is catalyzed by three bifunctional isozymes, acetohydroxy acid synthase (AHAS) I, II and III. These isozymes catalyze the condensation of two molecules of pyruvate with the formation of α -acetolactate the initial reaction in valine biosynthesis (Strassman *et al.*, 1953). In addition, these isozymes also catalyze the formation of α -acetohydroxybutyrate by the condensation of α -ketobutyric acid and pyruvate (Adelberg 1954; Strassman *et al.*, 1954). (See Fig.1.1).

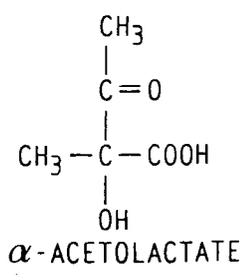
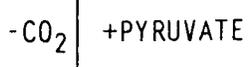
The second enzyme in the parallel pathway is acetohydroxy acid isomere-

Fig. (1.1).

Isoleucine-valine biosynthetic pathway including the metabolic intermediates and the enzymes catalyzing the various steps.

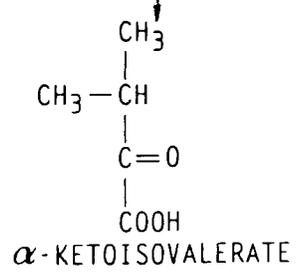
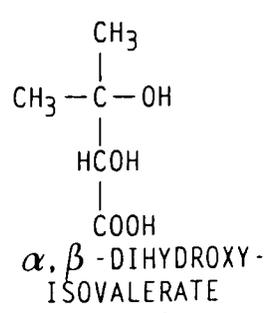


PYRUVATE



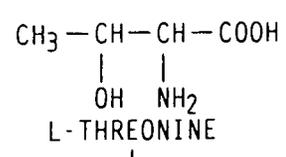
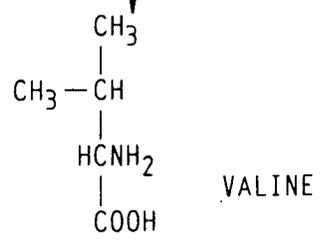
CONDENSING ENZYMES
(AHAS I, AHAS II & AHAS III)

ACETOHYDROXY ACID ISOMEROREDUCTASE

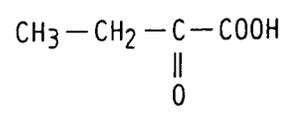


DEHYDRASE

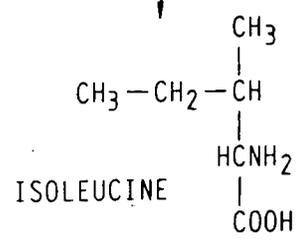
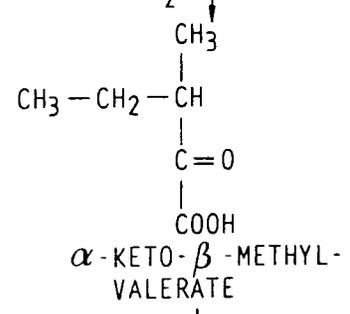
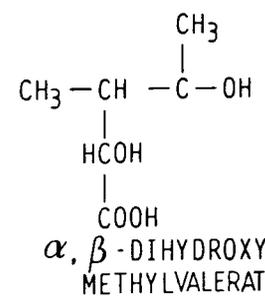
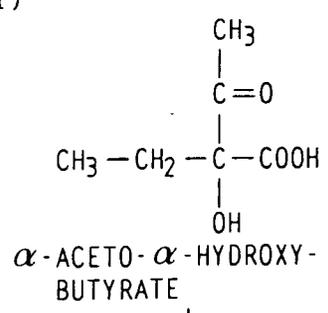
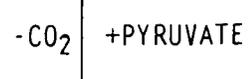
TRANSAMINASE B



L-THREONINE
DEAMINASE



α -KETO BUTYRATE



ductase which catalyzes the reductive isomerization of α -acetolactate to α - β -dihydroxyisovalerate and α -acetohydroxybutyrate to α , β ,dihydroxy- β -methylvalerate (Umbarger *et al.*, 1960). The reaction involves intramolecular migration of the R group from α to the β carbon (see Fig.1.1). It has been determined that acetohydroxy acid isomerase has a molecular weight of 54,000 (Wek and Hatfield, 1986).

The subsequent reaction is the dehydration of the α - β -dihydroxy acids to yield the corresponding α -keto acids (see Fig.1.1). This is catalyzed by the enzyme dihydroxy acid dehydrase (Myers and Adelberg, 1954). This enzyme has a molecular weight of approximately 66,000 (Uzan *et al.*, 1981; Gray *et al.*, 1981).

The last enzyme in the parallel pathways is transaminase B which catalyzes, the transamination of the α -keto acids forming the respective amino acids (Rudman and Meister, 1953). (See Fig.1.1). This enzyme was purified and characterized as a hexamer of identical subunits with a molecular weight 32,000 (Lee-Peng *et al.*, 1979).

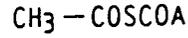
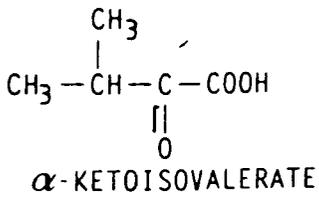
The formation of leucine begins by a condensation of α -ketoisovalerate, which is the immediate precursor of valine, with acetyl coA, to yield α -isopropylmalic acid. The subsequent steps are similar to those leading from citric acid to α -ketoglutaric acid in the tricarboxylic acid cycle (see Fig.1.2).

1.2 Feedback or end-product inhibition.

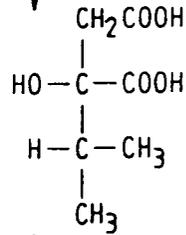
This is a mechanism by which the activity of an enzyme is regulated by the steady state concentration of the product of the pathway. In bacteria, feedback inhibition is common in biosynthetic pathways as one of the mechanisms by which flow through the pathway is regulated. In most multi-enzyme pathways, it is the first enzyme that is inhibited by the end-product. The first enzyme is usually

Fig. (1.2).

Leucine biosynthetic pathway.

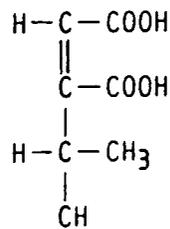


ACETYL COA



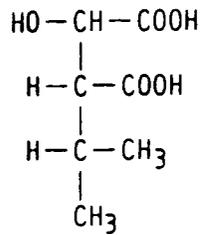
α -ISOPROPYLMALIC ACID

- H₂O



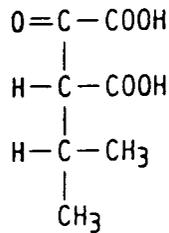
α -ISOPROPYLMALIC ACID

+ H₂O



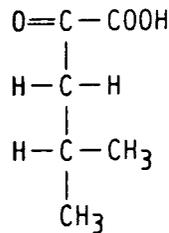
α -HYDROXY- β -CARBOXYI-SOCARPORIC ACID

NADP+



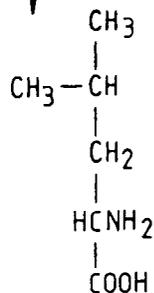
α -KETO- β -CARBOXYISOCARP-ORIC ACID

- CO₂



α -KETOISOCAPORIC ACID

TRANSAMINATION



LEUCINE

multimeric and contains separate binding sites for the substrate and regulatory effector, giving rise to allosteric reaction kinetics.

1.3 Feedback inhibition for isoleucine, valine and leucine biosynthesis in *E. coli*.

The conversion of threonine to isoleucine comprises an example of a multi-enzyme system. In this biosynthetic pathway, the first enzyme of the sequence, threonine deaminase, is strongly inhibited by the end-product of the sequence (isoleucine) when it is accumulated at high intracellular levels (Umbarger 1956).

Accumulation of valine in the cell to a high level has the effect of inhibiting two of the three acetohydroxy acid synthase isozymes, AHAS I and AHAS III, that catalyze the first step of the parallel pathways (Guardiola *et al.*, 1977). The third isozyme AHAS II is resistant to end-product inhibition by valine.

Additionally, leucine is also involved in feedback inhibition mechanism, A high level of leucine inhibits the activity of AHAS III (Haughn *et al.*, 1985).

1.4 Regulation of gene expression.

There are many mechanisms for genetic regulation in bacteria. All these mechanisms share the principle that a system is expressed only when it is needed. This type of on and off activity is carried out by regulating transcription.

Regulation of transcription in bacteria exists in several patterns and these patterns depend on the type of metabolic activity of the system being regulated. For example, in a degradative system (catabolic pathway), the concentration of the initial substrate in the sequence often determines whether the enzymes in the pathway are synthesized. In an anabolic pathway the final product is the

regulatory substance. On the other hand, in a system in which a single type of protein molecule is translated from a monocistronic mRNA, the protein may be autoregulated.

There are two major categories of molecular mechanism for genetic regulation called negative and positive regulation. In negative regulation the transcription is turned off by an inhibitor protein present in the cell and an inducer (anti-inhibitor) is needed to turn the system on. In positive regulation, an activator protein is required to stimulate transcription. There are some genetic systems which are both positively and negatively regulated.

1.5 The lactose operon.

Lactose is a sugar that can serve as a carbon source for metabolism. A mechanism for the regulation of the lactose genes, the *lac* operon, was proposed by Jacob and Monod (1961). The *lac* operon includes three structural genes *lacZ*, *lacY* and *lacA* which encode β -galactosidase, permease and acetylase respectively. There is an operator region (*lacO*), recognized by a repressor protein (encoded by the *lacI* gene), and a promoter region (*lacP*) to which RNA polymerase attaches before transcription of the structural genes. Two states of the *lac* operon are established, an off state (in the absence of lactose) and on state (in the presence of lactose). In the case of wild-type *E. coli* growing in the absence of lactose, RNA polymerase binds to the repressor gene (*lacI*) and transcribes the *I* gene. The synthesis of the repressor is constitutive. The repressor binds with high affinity to a particular base-pair sequence in the operator. When the repressor is bound to the operator, RNA polymerase can not bind to the operon's promoter and hence transcription of the structural genes will not occur. There is in fact a low level of transcription of the structural genes, that results in the presence of a few molecules of

each structural protein. This is because repressor protein is in equilibrium bound and unbound forms. This occasionally allows access of RNA polymerase to the promoter. When wild-type *E. coli* is grown in the presence of lactose as the sole carbon source, the cells must make large amounts of the lactose utilization enzymes to grow. Some of the lactose transported into the cells is converted by the few existing molecules of β -galactosidase into allolactose which in turn induces the production of the *lac* operon enzymes (thus allolactose, not lactose is the actual inducer of the *lac* operon). In addition to having a recognition site for the *lac* operator the *lac* repressor protein also has a recognition site for allolactose. When allolactose binds to the repressor it causes a conformational change in the repressor. As a result, the repressor loses its affinity for the *lac* operator and dissociates from it (free repressor protein is also altered so that it can not bind to the operator). In the absence of repressor, RNA polymerase is now able to bind to the promoter and initiate the transcription of a single polycistronic mRNA.

1.6 Catabolite repression.

Glucose, being the preferred carbon source, plays an important role in the regulation of expression of inducible enzymes. In this regulation the rate of the synthesis of inducible enzymes is reduced to about a third when bacteria are grown in a medium containing glucose compared with the same strain growing in a medium without glucose. The reason for this reduction is that glucose metabolism exerts an effect in lowering the level of 3'-5'-cyclic AMP in the cell (Makman and Sutherland, 1965). Cyclic AMP is essential for expression of the *lac* operon. A protein called the catabolite activator protein (CAP) forms a complex with cAMP. It is now established that the cAMP-CAP complex binds to a base sequence in the *lac* promoter region. This binding is a prerequisite for the correct RNA polymerase binding. Hence, the cAMP-CAP complex is a positive regulator. The cyclic AMP-

CAP complex acts not only on the *lac* operon but also needed for the expression of many inducible operons.

Thus, the *lac* operon responds to both positive and negative control.

1.7 The Arabinose operon.

Arabinose is a sugar which can serve as a carbon source for metabolism. In *E. coli* the arabinose operon includes three structural genes, *araB*, *araA* and *araD* which form an operon. Adjacent to these there is a complex promoter region and a regulatory gene *araC*. Additionally, there are other unlinked genes *araE* and *araF* which are necessary for the transport of arabinose across the cell membrane.

The arabinose operon is another inducible system and the inducer is arabinose itself. The *lac* operon is under the control of independent positive and negative systems, whereas, the arabinose operon is under dual positive and negative control mediated by the same protein encoded by the *araC* gene (Englesberg 1971; Lee 1978). This double action of *araC* protein was explained by proposing that there are two isomeric forms of the protein, inducing and repressing form. In the absence of arabinose in the media, the repressor form of *araC* protein predominates but when arabinose is present it binds to the *araC* protein causing a change to the activator form. The *ara* system is also subject to catabolite repression.

The synthesis of *araC* protein is also regulated. It is one of a class of proteins that regulate their own synthesis (autogenous regulation). It acts as a negative regulator at the promoter of the *araC* gene. Thus, the expression of *araC* is repressed by its own product and stimulated by the CAP system (Casadaban 1976b).

1.8 The Galactose operon.

Galactose is a sugar and has two roles in cellular metabolism. It serves as a carbon source and also in the synthesis of cell wall components. This operon includes three structural genes, *galK*, *galT* and *galE*, an operator and a promoter region. The repressor gene *galR* is not linked to the structural genes. The idea of existence of the two overlapping promoters in this operon was concluded from *in vitro* transcription studies with the *E. coli gal* operon. One promoter (P1) functions only in the presence of the cyclic AMP-receptor protein (cAMP-CRP) whereas the other promoter (P2) operates without these factors and is inhibited by them (Musso *et al.*, 1977).

Four *gal* operon promoter mutants have been isolated by Busby (1982). The biochemical analysis of these mutants argues strongly for the existence of two functional promoters *in vivo* and their different regulation. These mutants were classified into two classes. One class of mutant has the property that high level synthesis of the enzymes fails to occur when glucose is present. The other mutant class is defective in enzyme synthesis when glucose is absent. Glucose fails to inhibit the induction of the *gal* operon because one of the promoters does not require cAMP-CAP and thus remains active even in the presence of glucose.

The start points S1 and S2 for the two *gal* mRNA molecules differ functionally. Transcription from S1 occurs when glucose is absent because it is dependent upon cAMP-CAP. While transcription from the S2 occurs when glucose is present. The mechanism by which cAMP-CAP stimulates S1 transcription and inhibits S2 transcription is not clearly known. It is thought that stimulation at the S1 site occurs in a way that is similar to that in the *lac* operon.

The mechanism of action of the *gal* repressor is also obscure. It is thought that the *gal* repressor binds to the cAMP-CAP binding site and blocks the the initiation

at S1 in one of two ways. Either by interfering with the binding of cAMP-CAP to DNA or by altering the stimulatory effect of cAMP-CAP on RNA polymerase binding. Repression at S2 seems to have a totally different mechanism. Initiation still occurs at S2 if the *gal* repressor is present but transcription is terminated 10-20 bases downstream, how this happens is not known.

1.9 The tryptophan operon (A biosynthetic system).

Tryptophan is synthesized by a group of five enzymes encoded by the *trp* operon consisting of *trpE*, *trpD*, *trpC*, *trpB* and *trpA*. The first gene to be transcribed is *trpE*. Adjacent to this are the promoter, the operator, the leader and the attenuator regions.

Regulation of tryptophan biosynthesis were thought to be carried out solely by the repressor-operator control system of gene expression and feedback inhibition of enzyme activity. Tryptophan acts as a corepressor it combines with a repressor protein and together they bind to the operator that precedes the *trp* operon and thus represses the synthesis of *trp* mRNA.

The discovery of attenuation adds an additional level of control to the operon. Attenuation is a mechanism involved in the control of several amino acid biosynthetic operons and was first described in the *trp* operon of *E. coli*. This system is effected by premature termination of transcription before it reaches the first structural gene and regulation of the frequency of this termination is by the concentration of relevant amino acids.

The *trp* operon initiates transcription 162 base pairs before the start codon of the first structural gene *trpE*. This segment of mRNA is called the leader, the translation of the leader *in vivo* to a small peptide was established (Miozzari and Yanofsky, 1978). Within this leader there is a sequence of bases (123 through 150)

which if deleted, cause a sixfold increase in the synthesis of *trp* enzymes either in derepressed cells or constitutive mutants. Thus, this sequence of bases must have a regulatory activity. It has been established that transcription termination occurs in this region at a site called the attenuator. The attenuator itself consists of a G-C rich stem and loop structure followed by a stretch of poly-U residues. The formation of this G-C rich structure causes transcription termination.

Yanofsky *et al.*, (1978) proposed a model for attenuation which was modified by Keller and Calvo (1979) using the *trp* operon as an example. The regulation of transcription termination is concerned with allowing or preventing the G-C rich terminator to form. When the terminator fails to form, transcription will continue through the poly U stretch into the structural genes. Preceding the terminator there are several overlapping stem and loop structures. The structure overlapping the terminator is called the pre-emptor, if it is allowed to form it will prevent the formation of the terminator allowing transcription to continue. Preceding the pre-emptor is another structure called the protector, which will allow the formation of the terminator. Such structures form in order of transcription, thus the protector will form first preventing pre-emptor formation and thus allowing the formation of the terminator (see Fig.1.3). Overlapping the protector is the coding sequence for a small (14 amino acid residues) polypeptide, called the leader peptide. The leader polypeptide has an interesting feature namely, at positions 10 and 11 there are two adjacent tryptophan codons.

The attenuation control system works as follows, in the presence of excess tryptophan there is also a large amount of the corresponding charged tRNA thus the leader peptide will be rapidly synthesized. As transcription and translation are coupled in bacteria, the ribosome will follow closely behind the RNA polymerase, translating the message as soon as it is formed. When the ribosome reaches the stop codon of the leader peptide, all of protector stem and loop can form, whereas

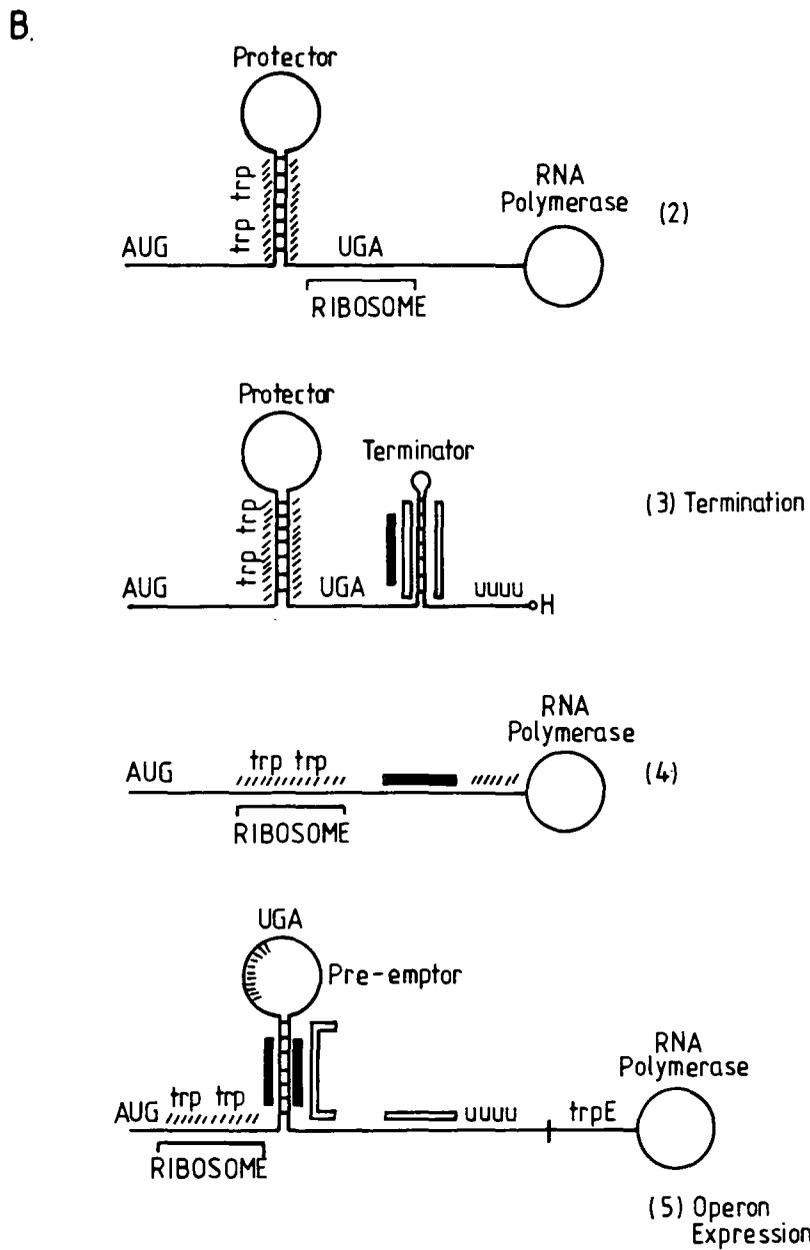
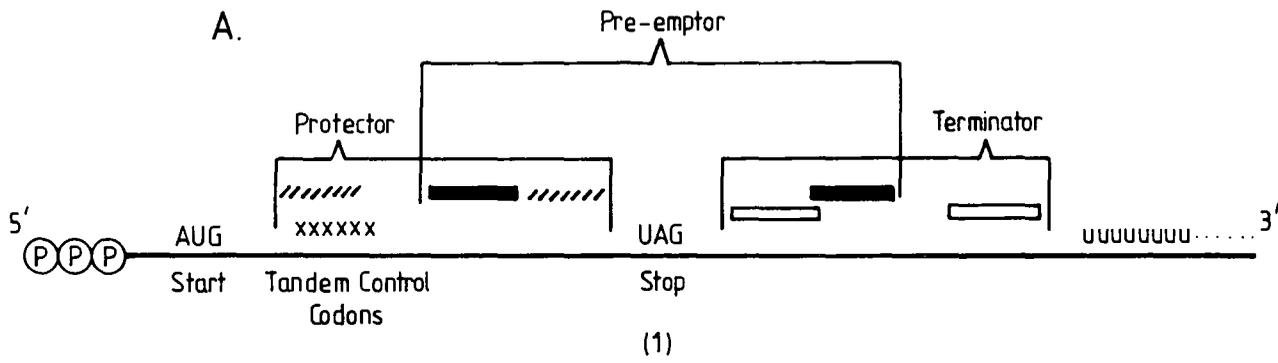
Fig. (1.3)

A. The important features of leader mRNA in the tryptophan operon. Regions involved in stem and loop formation are:- Hatched area is the protector stem. Closed box is the pre-emptor stem. Open box is the terminator stem. While xxxxxxxx refers to the positions of the tryptophan regulatory codons.

B. Effect of various growth conditions on the formation of secondary structures of the *trp* leader transcript.

2 & 3 under excess tryptophan. 4 & 5 under limiting amounts of tryptophan.

After M. D. Watson (1981).



the pre-emptor cannot, the formation of protector stem and loop will result in formation of a terminator (see Fig.1.3). On the other hand, when tryptophan is limiting there will be insufficient charged tRNA to complete the translation of the leader peptide. The ribosome will thus stall at the tandem tryptophan codons and prevent formation of the protector. Thus, the first secondary structure formed will be the pre-emptor. As the result of pre-emptor formation it will prevent subsequent formation of the terminator. In the absence of the terminator the RNA polymerase will transcribe through the attenuator into the structural genes.

1.10 Gene-enzyme relationships in the isoleucine-valine system of *E. coli*.

Threonine deaminase, the first enzyme in isoleucine biosynthesis, is encoded by the *ilvA* gene. The three AHAS isozymes are encoded by *ilvB* (AHAS I), *ilvG* (AHAS II) and *ilvH* (AHAS III).

In *in vitro* complementation studies by DeFelice *et al.*, (1974), it has been suggested that the enzyme AHAS III composed of a regulatory subunit (encoded by *ilvH*) and a catalytic subunit (encoded by *ilvI*). Similarly, AHAS II has a large (*ilvG*) and a small (*ilvM*) subunit, as does AHAS I (*ilvB* and *ilvN*). AHAS I and AHAS III are sensitive to valine feedback inhibition, whereas AHAS II is resistant to end-product inhibition by valine. The original *E. coli* K-12 strains express only AHAS I and AHAS III activities because of a naturally occurring frameshift mutation in *ilvG*. (Favre *et al.*, 1976; Lawther *et al.*, 1982). This results in the well described valine-sensitivity of wild-type *E. coli* strains. Later, spontaneous mutations have been isolated in the *ilvG* gene which remove this frameshift site. These were originally called *ilvO*⁻ mutations (Ramakrishnan and Adelberg, 1965). These mutations result in expression of AHAS II and confer valine resistance to the bacteria and result in a higher expression of the *ilvGMEDA*

operon.

The second step in the parallel pathway is catalyzed by acetohydroxy acid isomeroreductase which encoded by the *ilvC* gene.

Dihydroxy dehydrase is the third enzyme catalyzing the third reaction in the parallel pathway. This enzyme is encoded by the *ilvD* gene.

The last gene involved in the above biosynthesis is *ilvE* which encodes the enzyme transaminase B.

1.11 Physical Maps of the *ilv* genes.

The physical location of the *ilvGMEDAC* genes on the restriction map of the *ilv* region of *E. coli* K-12 was determined by McCorkle *et al.*, (1978) by two methods. Firstly, heteroduplex and endonuclease cleavage analysis of hybrid phages carrying genetically defined parts of the *ilv* cluster. Secondly, complementation analysis and enzyme assays to determine *ilv* genes expression from hybrid plasmids containing DNA restriction fragments of the transducing phage $\lambda 80d_{ilv}$.

The results obtained by McCorkle *et al.*, (1978) revealed the location of the *ilvEDA* transcriptional unit and *ilvC*. The direction of transcription was deduced for both units. The exact location of *ilvGM* could not be determined in their studies. Subrahmanyam *et al.*, (1980) eventually showed that *ilvG* is adjacent to *ilvE* and forms part of the same operon.

Thus, from above description, it can be concluded that the major transcriptional unit of the *ilv* cluster is the *ilvGMEDA* operon. This operon encodes four of the five enzymes involved in the parallel pathway of isoleucine-valine biosynthesis. (See Fig.1.4).

The location of the *ilvY* gene is between the *ilvA* and *ilvC* genes. This was

Fig. (1.4).

The major *ilv* cluster. The position of the structural genes are demonstrated within this cluster. The size of these genes are drawn to scale. Arrows represent the direction of transcription within this cluster.

P[G]

ilvG

ilvM

ilvE

ilvD

ilvA

ilvY

ilvC



0

1000

2000

3000

4000

5000

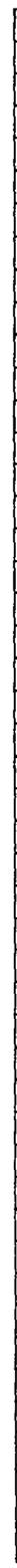
6000

7000

8000

9000

9600



Overlapping
P[Y] & P[C]



shown by the fact that only plasmids containing the entire region between *ilvA* and *ilvC* were able to restore to strain CU827, which deleted for *ilvDAYC*, the ability to induce *ilvC* expression (Watson *et al.*, 1979).

The *ilvGMEDAYC* cluster is located at 85 minutes on the *E. coli* linkage and the *ilvBN* and *ilvIH* are located at 82 and 2 minutes respectively on the linkage map (Bachmann 1983).

1.12 Regulation of the *ilv* operons.

Genes and operons concerned with amino acid biosynthesis are generally regulated by two mechanisms in *E. coli* and other enteric bacteria. The first mechanism involves negative control and repression of transcription by an operator-repressor system while the second one is attenuation. In *E. coli* K-12, the genes and operons involved in the biosynthesis of isoleucine-valine are generally regulated by attenuation, however, there are exceptions to this outlined in the following sections.

1.13 Regulation of *ilvBN* operon.

The *ilvBN* operon is regulated both by catabolite repression (Freundlich 1977; Sutton and Freundlich, 1980), who carried out a number of *in vivo* experiments that suggest that the *ilvBN* operon is positively regulated by cAMP-CAP, and attenuation (Friden *et al.*, 1982; Hauser and Hatfield, 1983).

Friden *et al.*, (1982) found that cyclic AMP was involved in the control of the *ilvBN* operon via the usual complex. The bound cAMP-CAP (catabolite activator protein) protects an area of 26 nucleotides positioned around the -70 region of the *ilvBN* promoter. The above *in vitro* studies showed that transcription from *ilvB* promoter is strongly increased by addition of cyclic AMP and CAP, thus

confirming that *ilvB* is regulated by cyclic AMP.

The second mechanism of regulation is through attenuation in a manner similar to that proposed for the regulation of the tryptophan operon. Friden *et al.*, (1982) determined the DNA sequences of the promoter-regulatory region of the *ilvBN* operon of *E. coli*. A sequence starting at the ATG at position 36 and ending at position TAG at position 134 that could code for a polypeptide. This potential leader polypeptide contains 32 amino acids, 9 of which are valine at positions 19 through 28 and 3 leucine residues at positions 6, 10 and 11. These two amino acids are involved in the multivalent control of the *ilvBN* operon.

Five stem and loop structures can be potentially formed by the *ilvB* leader mRNA. So, in this regulation more stem and loop structures can be formed than that of the *trp* leader attenuation. The initial stem and loop structure (pre-emptor defender) forms prior to and precludes the protector and thus allow the pre-emptor to form while the other extra stem and loop structure (protector defender) precludes the pre-emptor defender and permits protector formation. Thus, the function of these protector and pre-emptor defenders is to safeguard the formation of protector and pre-emptor respectively. Without these initial structures the tandem Leu codons of the leader peptide could not participate in regulation because these codons are situated too far upstream from the protector and the Val regulatory codons. A ribosome stalls at the tandem Leu codons when there is a shortage of leucyl-tRNA^{Leu} and this stalling prevents the formation of the protector defender. Thus, pre-emptor defender and pre-emptor form and preclude termination which results in transcription readthrough and increased expression of the *ilvB* operon (see Fig.1.5). On the other hand during a limitation of valyl-tRNA^{Val}, stalling of ribosomes at the multiple tandem Val codons prevents formation of the protector itself allowing the pre-emptor to form and so again precluding transcription termination (see Fig.1.5). The model proposed above

. Fig. (1.5).

The alternative secondary structures of the leader mRNA in the *ilvBN* operon and the effects of growth conditions on formation of these structures.

A. Possible secondary structures in the absence of ribosome.

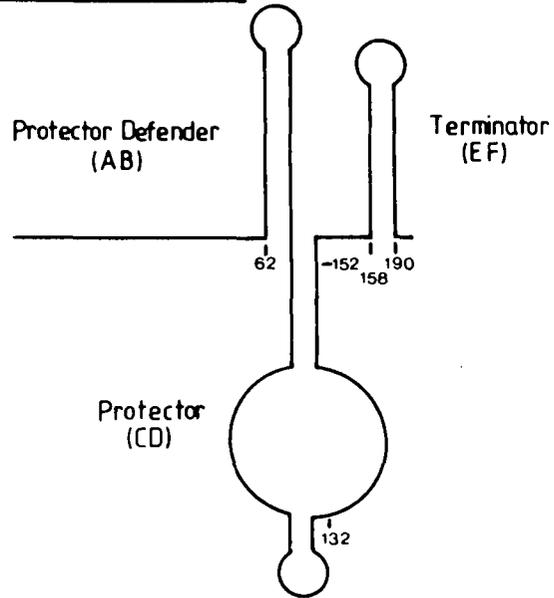
B. Excess leucine and valine which cause transcription termination.

C. limiting Valyl-tRNA.

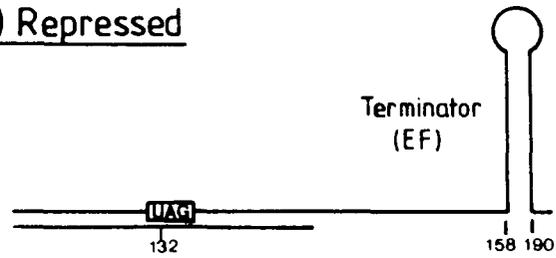
D. limiting leucyl-tRNA. Both conditions in C and D cause transcription readthrough.

After Friden, P. *et al.*, (1982).

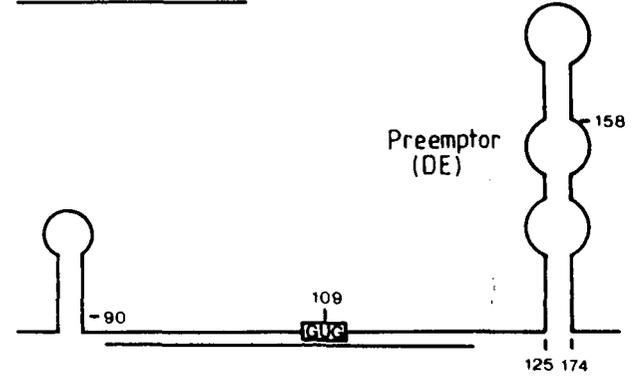
(a) No Ribosomes



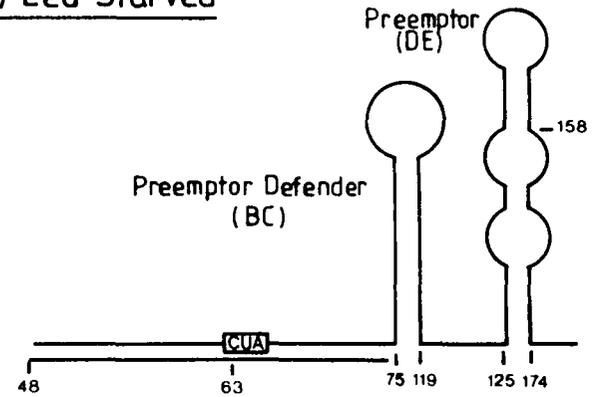
(b) Repressed



(c) Val Starved



(d) Leu Starved



also predicts that ribosomes stalling at the 3rd, 4th or 5th Val codons results in the partial relief of attenuation. It is surprising that the 1st and 2nd Val do not participate in regulation.

1.14 Regulation of the *ilvGMEDA* operon.

The *ilvGMEDA* operon is regulated by a multivalent mechanism (Lawther and Hatfield, 1980; Adams *et al.*, 1985). The expression of enzymes encoded by the above operon is inhibited by the presence of all three branched amino acids, isoleucine, leucine and valine. When the growth rate of the cell is limited by the availability of any one of these amino acids the rate of synthesis of these enzymes increases. On the other hand, there is no direct evidence for the involvement of a repressor in the regulation of this operon.

Lawther and Hatfield (1980) determined and characterized the regulatory region for the *ilvGMEDA* operon of *E. coli* K-12. This region includes a transcription promoter, 186 nucleotide leader RNA encoding a 32 residue polypeptide containing 5 isoleucine, 5 leucine and 6 valine codons and a transcription termination site preceding the first structural gene. Based on the secondary structure prediction of the nucleotide sequence of the leader RNA, it is proposed that the leader RNA can form alternative structures which control the level of transcription into the structural genes of this operon. Three stem and loop structures can be formed by the leader RNA.

Lawther and Hatfield (1980) proposed a model for the multivalent attenuation of the *ilvGEDA* operon on the basis of analysing the possible secondary structures of the its leader RNA. This model proposes that under *in vivo* conditions of coupled transcription-translation, a ribosome stalling at initial tandem codons for leucine at positions 4 and 5 results in disruption of the base pairing between re-

gions 1A and the 3' half of region 2. This free 3' half of region 2 can now base pair with complementary sequences in region 5 of the stem-loop terminator structure. This event destabilizes the top half of the terminator and results in prevention of transcription termination (see Fig.1.6B).

Stalling the ribosome at the tandem valine codons at positions 11 and 12 and the three valine codons followed by the three isoleucine codons at positions 15 through 20 disrupt the base pairing between regions 1B and the 5' half of region 2. This disruption results in inhibition of the formation of the bottom half of the stem-loop terminator structure and as a result transcription of the structural genes occurs (see Fig.1.6C). A ribosome stalling at these latter codons exposes the 5' end of the leader message sufficiently to load another ribosome. From Fig.1.6D it is clear that the loading of additional ribosomes allows both halves of region 2 to base pair with region 5 and result in complete disruption of the terminator structure. To explain the observation that the *ilvGMEDA* operon is deattenuated by isoleucine limitation, the stalling of a ribosome at the last two isoleucine codons allows the attachment of another ribosome. This is necessary because if there was no ribosome at region 1A then the entire stem-loop terminator structure would form.

Recently, the necessity for transcription of certain sequences of DNA more than 100 bp upstream of the transcription initiation site of the *ilvGMEDA* operon has been reported. This observation suggests that there may be a new class of promoters in enteric bacteria and that regulation of this operon cannot be solely explained by attenuation (Haughn *et al.*, 1985).

. Figure (1.6)

The possible secondary structures of the leader mRNA in the *ilvGEDA* operon and the effect of ribosome positioning at different regulatory codons. Thick bars represent base-paired regions.

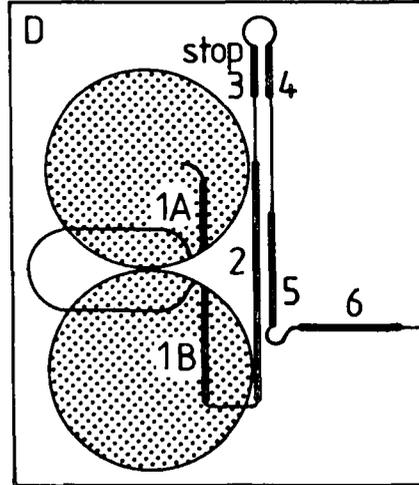
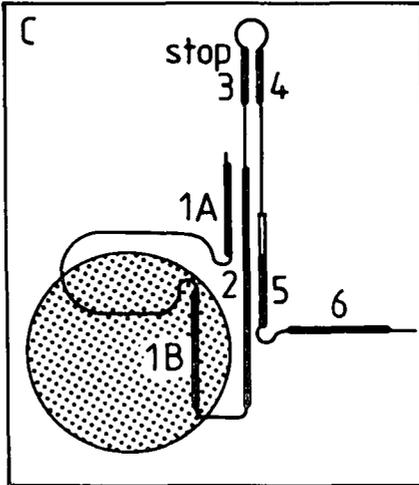
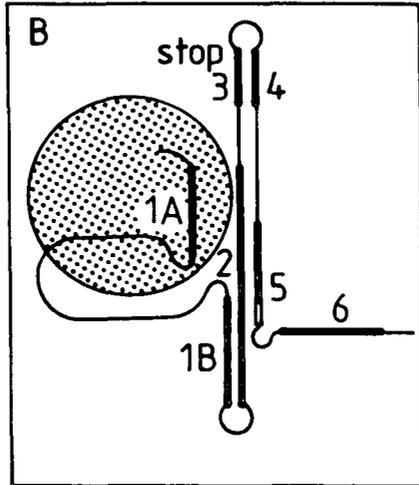
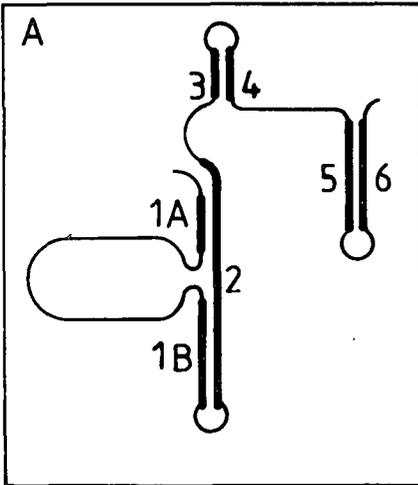
A. No translation.

B. Ribosome paused at a tandem leucine codons.

C. Ribosome paused at tandem valine or isoleucine codons.

D. Ribosome paused at tandem isoleucine codons with attachment of the second ribosome.

After Lawther & Hatfield (1980).



1.15 Regulation of the *ilvIH* operon.

The expression of the *ilvIH* operon is negatively controlled by leucine at the level of transcription. The mechanism by which leucine represses the transcription of the *ilvIH* operon remains unclear. Haughn *et al.*, (1985) studied the mechanism by which *ilvIH* operon is regulated. Their results reveal that sequences more than 200 bp upstream of the presumed site of transcription initiation appear to be necessary for optimal expression of *ilvIH* operon. Transcription initiated near position -30 bp was repressed by leucine only if these upstream sequences were present. This appears to rule out a mechanism of regulation based solely on attenuation.

1.16 Regulation of the *ilvYC* system.

Arfin *et al.*, (1969) concentrated their studies on the regulation of expression of the *ilvC* gene. They produced evidence which strongly supports the view that acetohydroxy acid isomerase is induced directly or indirectly by its substrate. This finding is quite in accord with the fact that this enzyme is not controlled by either the operator locus that controls the three structural genes for the enzymes transaminase B, dehydrase and threonine deaminase or by the operator locus that controls the genes governing the AHAS enzymes (Ramakrishnan and Adelberg, 1965).

Another attempt to study the mechanism of *ilvC* induction in *E. coli* strain K-12 was carried out by Wild *et al.*, (1977) who developed an *in vitro* protein synthesizing system using the DNA from a plaque forming phage carrying an *ilvC::lacZ* fusion as a template for β -galactosidase synthesis. They stated that unless the *ilvC* inducer (acetohydroxybutyrate or acetolactate) was present during the transcription period, the level of β -galactosidase formed was very low.

The characterization of *ilvC* induction has been described in more detail by Watson *et al.*, (1979). They identified the existence of a regulatory gene, designated *ilvY*, which codes for a positive control factor necessary for *ilvC* induction. The product of the *ilvY* gene is called the epsilon protein. The experiments of Watson *et al.*, involved, firstly, the construction of an *ilvC::lacZ* fusion using the method of Casadaban (1976a). This fusion puts *lacZ* expression under the control of the *ilvC* promoter. Secondly, the preparation of S-30 extracts from several strains of *E. coli* including CU827 which carries the *ilvDAC115* deletion. S-30 extracts prepared from the above strains were examined for their ability to induce the *ilvC* directed *lacZ* expression in the presence of acetohydroxybutyrate. The results obtained from these experiments revealed that only the wild-type S-30 extracts were competent in *ilvC* directed β -galactosidase synthesis. The S-30 extracts prepared from the strains carrying the deletions failed to show any *lacZ* expression although acetohydroxybutyrate was available. Hence, the two deleted strains CU827 and CU838 lack a factor needed for *ilvC* induction. The gene for this factor deleted in the above strains, was designated *ilvY*. When the deleted strains were transformed with plasmids carrying parts of *ilvDAC* region the induction of *ilvC* was restored. Thus, the induction of *ilvC* is under the positive control of the product. Genetic analysis of *ilvY^c* mutations constitutive for *ilvC* expression confirms this prediction and map position (Biel and Umbarger, 1980).

In *E. coli* the existence of an inducible gene under positive control is unique for one involved in amino acid biosynthesis, which is more usually controlled by repressible negative control and /or attenuation.

A model has been proposed for the regulation of the *ilvYC* system (Wek and Hatfield, 1986). This model proposes that the *ilvY*-encoded activator protein, epsilon, in the presence of either acetohydroxybutyrate or acetolactate, the substrates for the *ilvC* encoded acetohydroxy acid isomeroreductase, bind to an

operator site located between *ilvY* and *ilvC* and activate *ilvC* transcription. The proposed operator site is a sequence that is conserved between both *E. coli* and *S. typhimurium*.

Assay methods for identification and characterization of DNA binding proteins.

1.17 The Filter binding assay.

Several techniques have been employed to identify and characterize DNA binding proteins. These assays in general exploit the specific binding of these proteins to certain DNA sequences. The filter binding assay is one such method which is at present widely used (Riggs *et al.*, 1970). This assay depends on the fact that proteins have the ability to bind to nitrocellulose filters. DNA molecules rarely bind to nitrocellulose filters under physiological conditions. However, when a particular DNA binding protein is present with its target DNA, it will cause retention of the DNA on the nitrocellulose filter.

Many protein-nucleic acids interactions have been studied utilizing a membrane filter binding technique, of these, aminoacyl tRNA synthase-tRNA (Yarus and Berg, 1967), lac repressor-*lacO* (Riggs *et al.*, 1968) and mRNA-ribosome (Nirenberg and Leder, 1964), interactions are good examples.

Due to some weaknesses in the filter binding assay, complete retention cannot be expected. These weaknesses can be explained by three possible reasons. Firstly, some of the DNA molecules are broken and thus they will not possess an intact binding site. Secondly, washing, which is usually done after applying the binding mixture to the nitrocellulose filter, causes loss or dissociation of some complexes. Lastly, the binding of the protein itself may not be 100% efficient.

The second weakness prevents this technique from providing an accurate analysis of the composition of the binding reaction product.

1.18 The Gel retardation assay.

Regulatory proteins have the ability to bind to a particular DNA sequence. This protein-DNA binding requires more characterization and analysis. In addition to the filter binding assay, the gel retardation assay is a powerful method for studying DNA-protein interactions. The principle of this technique relies on the fact that a protein bound to a particular DNA sequence has a greater molecular mass than either of the two components alone. The product of this interaction is analysed on either an agarose or polyacrylamide gel electrophoresed in low ionic strength buffer which stabilize the complex against dissociation. The greater mass of the complex results in a lower mobility compared with the migration of the same DNA fragment without bound protein.

This technique was first introduced by Fried and Crothers (1981) who characterized the interaction of *E. coli lac* repressor with DNA restriction fragments containing specific repressor binding sites. The protein is mixed with one or more DNA fragments and protein-DNA complexes are resolved as discrete bands by polyacrylamide gel electrophoresis, mobilities decrease as a function of the number of proteins bound to a given DNA fragment.

Garner and Revzin (1981) applied this assay to the study of the *E. coli* lactose operon regulatory system. They showed that the CAP-cAMP complex forms a long-lived complex with wild-type *lac* promoters. Additionally, this gel retardation assay has been used by Zerbib *et al.*, (1987) to identify the specificity of IS1 DNA binding of particular protein in the cell extracts from *InsA* overproducers. They show that the extracts display a DNA binding activity specific for the ends of IS1

only. This activity was identified as the InsA protein. It is probable that the InsA protein recognizes and binds sequences within the inverted repeats which reside at the ends of IS1.

1.19 Gene fusions as a tool for genetic analysis.

Fusions of the *E. coli lacZ* gene to other genes produces a powerful tool for molecular biology due to the stability of β -galactosidase and the ease by which it can be assayed (Bassford *et al.*, 1978). Two types of fusions have been used, the first, a transcriptional fusion, is where the operon of interest is fused to a *lacZ* gene lacking its own promoter but containing its own translation start site. The second type is a translational fusion, in which a *lacZ* gene lacking both transcription and translation initiation signals is fused in frame to the coding sequence of the target gene. Both types of fusion mentioned above express the *lacZ* gene from the promoter of the target gene and allow the genetic and biochemical methods developed for the *lac* system to be used to study the regulation of any gene.

Additionally, the protein fusion, produced by the translational fusion, produces a hybrid protein whose N terminus is encoded by the target gene and is fused to an enzymatically active β -galactosidase. These hybrid proteins have additional applications beyond the study of gene expression. For instance, hybrid proteins can be used to raise antibodies against the exogenously encoded N terminus which in turn can be used to identify the product of the wild type target gene. Such, hybrid β -galactosidase proteins have been useful in the study of protein function.

As a result of the development of gene fusion technology much effort has been spent on constructing plasmids which can be used as cloning vectors which directly fuse a target gene to *lacZ*. Two classes of these cloning vectors have been

created. The first class contains various restriction enzyme cleavage sites in a variety of translational reading frames adjacent to a promoterless *lacZ* gene. The second class retains an active *lac* promoter and *lacZ* translational initiation region which can direct hybrid protein synthesis from DNA fragments that do not have functional promoters or translation initiation sequences (Shapira *et al.*, 1983).

1.20 The T7 expression system.

Sometimes it is necessary to express particular genes to a high level because their products normally comprise only a small fraction of total bacterial protein. For example, regulatory genes which encode regulatory proteins are normally expressed at only low levels.

Studier and Moffatt (1986) developed a gene expression system based on the bacteriophage T7 RNA Polymerase. They demonstrated that T7 RNA Polymerase is highly selective for its own promoter which does not occur naturally in *E. coli*. Transcription by T7 RNA Polymerase is very active. A relatively small amount of T7 RNA Polymerase, provided from a cloned copy of T7 gene 1, is sufficient to transcribe a gene placed next to a T7 RNA polymerase specific promoter and this transcription occurs at high level. T7 RNA polymerase and *E. coli* RNA polymerase recognize different promoter sequences (Dunn and Studier, 1983; Hawley and McClure, 1983). Thus, it is possible to clone a gene under the control of T7 RNA polymerase without effecting its expression by *E. coli* RNA polymerase.

The antibiotic rifampicin can be used to establish the high level expression of a particular gene by T7 RNA polymerase due to its inactivation of *E. coli* RNA Polymerase. Thus, all transcription in the cell will be due to T7 RNA polymerase. The rifampicin is useful in reducing the background in a pulse-labelling experiment to label either specific RNA or protein whose synthesis is under the control of T7

RNA Polymerase.

CHAPTER 2

MATERIALS AND METHODS

2.1 Materials.

All chemicals and reagents were purchased from Sigma except following, the restriction enzymes, modifying enzymes, X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactoside) and IPTG (isopropylthiogalactoside) were from Northumbria Biological Ltd., Cramlington, Northumberland.

Agarose and low melting point agarose were purchased from GIBCO/BRL, Paisley, Scotland and from ICN biochemical Ltd., High Wycombe, respectively.

All the radioactive isotopes were purchased from Amersham International p.l.c., Amersham, Bucks.

Nitrocellulose filters were from Schleicher and Schuell, Dassel.

3MM paper was from Whatman Ltd., Maidstone, Kent.

Sephadex G-25, G-50 and G-200 were from Pharmacia Fine Chemicals, Uppsala, Sweden.

Bacto agar was from Difco Labs, Detroit, Michigan, U.S.A.

BBL trypticase peptone was from Becton Dickinson and Co, Cockeysville, M.D., U.S.A.

Caesium chloride and sodium chloride were from Kochlight Ltd., Haverhill, Suffolk.

High vacuum grease was from Dow Corning S.A., Seneffe, Belgium.

Polaroid 667 film was from Polaroid (U.K.) Ltd., St Albans, Hertfordshire.

2.2 Preparation of media.

Luria broth

10g Bacto-Tryptone

5g Bacto-Yeast extract

5g NaCl

1 Litre H₂O

2x YT media

16g Bacto-Tryptone

10g Bacto-Yeast-Extract

5g NaCl

1 Litre H₂O

L-agar plates

15g Bacto-agar per 1 litre of L-broth.

MacConkey-lactose agar plates

40g MacConkey agar was dissolved in 1 litre water by heating to boiling followed by addition of 10g lactose then autoclaved.

Top-agar

It has the same composition of L-agar plates except that the agar is 0.66%.

M9 Media

0.6g Na₂HPO₄

0.3g KH₂PO₄

50mg NaCl

0.1g NH₄Cl

90 ml H₂O.

These components were mixed, autoclaved, cooled to room temperature then

0.2 ml 1 M MgSO₄

1.0 ml 20% Glucose

0.01 ml 1 M CaCl₂

H₂O to 100 ml.

2.3 Preparation of the buffers and common solutions used in this work.

10x Alec's buffer

Tris 44.4g

EDTA 3.7g

Ethidium Bromide 0.5 ml (10 mg/ml)

pH 7.7 with glacial acetic acid

H₂O to 1 litre.

10x Binding buffer

Tris-HCl pH 7.4 500 mM

KCl 100 mM

EDTA 10 mM

β -mercaptoethanol 10 mM

MgCl₂ 70 mM

CaCl₂ 30 mM

To 1 ml by sterile water.

10x TBE buffer

Tris 121.1g

Boric acid 51.35g

EDTA 3.72g

H₂O to 1 litre.

10x PAGE buffer

Glycine 141g

Tris 30g

SDS 10g

H₂O to 1 litre.

2x Sample buffer

Tris-HCl pH 6.8 0.2 M

SDS 0.2%

Sucrose 10%.

Prehybridization buffer

20x SSC 18 ml

0.5% SDS 3 ml

5x Denhardt's solution 6 ml

Denatured salmon sperm DNA (100 µg/ml) 200 µl

H₂O to 60 ml.

Hybridization buffer

This has the same components as prehybridization buffer in addition to 0.01 M EDTA and 1 ml of radioactive probe.

20x SSC solution

NaCl 175.3g

Trisodium citrate 88.2g

Dissolved in 800 ml water and the pH adjusted to 7.0 with 10 M NaOH and the volume made up to 1 litre.

5x Denhardt's solution

Ficoll 400 0.5g

Polyvinyl Pyrrolidone 0.5g

BSA 0.5g

Dissolved in 100 ml sterile water and stored at -20°C.

Gel stain solution

Coomassie Brilliant Blue 0.2 %

Methanol 50 %

Glacial acetic acid 7%.

Gel destaining solution

Methanol 40%

Glacial acetic acid 10%.

TE buffer pH 8.0

Tris-HCl pH 8.0 10 mM

EDTA pH 8.0 1 mM.

TEN buffer

Tris-HCl pH 8.0 10 mM

EDTA pH 8.0 1 mM

NaCl 10 mM.

TENS buffer

Tris-HCl pH 8.0 10 mM

EDTA pH 8.0 1 mM

NaCl 10 mM

SDS 0.1%.

Transformation buffer

CaCl₂ 50 mM

Tris-HCl pH 8.0 10 mM.

10x Transcription buffer

MgCl₂ 100 mM

Tris-HCl pH 8.0 0.4 M

EDTA pH 8.0 10 mM

β -mercaptoethanol 60 mM

BSA 5 mg/ml

Glycerol 2%

DTT 30 μ M.

Nucleotide mixture solution

15 μ l of 10 mM of ATP, GTP, CTP and UTP were mixed and the volume completed to 100 μ l by sterile water.

10x Calf intestinal phosphatase (CIP) buffer

Tris-HCl pH 9.0 0.5 M

MgCl₂ 10 mM

ZnCl₂ 1 mM

Spermidine 10 mM.

10x Ligation buffer

Tris-HCl pH 7.6 0.66 M

MgCl₂ 50 mM

DTT 50 mM

ATP 10 mM.

10x Blunt end-kinase buffer

Tris-HCl pH 9.5 0.5 M

MgCl₂ 0.1 M

DTT 50 mM

Glycerol 50%.

10x T4 kinase buffer

Tris-HCl pH 7.6 0.5 M

MgCl₂ 0.1 M

DTT 50 mM

Spermidine 1 mM

EDTA 1 mM.

10x nick translation buffer

Tris-HCl pH 7.2 0.5 M

MgCl₂ 0.1 M

DTT 1 mM

BSA 500 $\mu\text{g/ml}$.

Benedict's Reagent.

Composed of two solutions:-

Sol.I

Trisodium citrate 173g

Na_2CO_3 100g

The two above components were dissolved in 800 ml water with warming.

Sol.II

CuSO_4 17.3g in 100 ml water.

Sol.I was added to Sol.II and the volume was completed to 1 litre.

Amino acid stock solutions

Stock solutions of individual amino acids were prepared at concentrations of 2 mg/ml then autoclaved. They were used at a final concentration 20 $\mu\text{g/ml}$.

Vitamin stock solution

A stock solution of 1 mg/ml thiamine was prepared. It was used at a final concentration 10 $\mu\text{g/ml}$.

IPTG

A stock solution of 100 mM was prepared and used at a final concentration 0.05 mM.

Antibiotic stock solutions

Stock solutions of ampicillin 25 mg/ml in 70% ethanol, tetracycline and 5 mg/ml in absolute ethanol, kanamycin 25 mg/ml in sterile H_2O , chloramphenicol 20 mg/ml in absolute ethanol and rifampicin 10 mg/ml in DMF were prepared. They were used at the final concentrations 50 $\mu\text{g/ml}$, 10 $\mu\text{g/ml}$, 50 $\mu\text{g/ml}$, 10 $\mu\text{g/ml}$ and 200 $\mu\text{g/ml}$ respectively.

Gel loading dye solution

1 ml 10% SDS, 2 ml of 250 mM EDTA pH 8.0, 0.2 ml 1M Tris-HCl pH 8.0, 5 ml 100% glycerol and 10 mg of Bromophenol blue and the volume was completed to 10 ml by sterile water.

Loading buffer for gel retardation assay

To 1 mg/ml BSA, 50% glycerol and 0.01% xylene cyanol were added then the volume was made up to 1 ml by D.water.

Sequencing stop buffer(formamide-dye mix)

To 98% deionized formamide (w/w), 10 mM EDTA pH 8.0, 0.2% bromophenol blue (w/v) and 0.2% xylene cyanol (w/v) were added and the volume was completed to 1 ml.

2.4 Sterilization.

All the media and glassware used in this work were autoclaved at 15 p.s.i., 123°C for 15 minutes.

2.5 Extraction with Phenol.

The removal of proteins from nucleic acids solutions was carried out as follows:- An equal volume of phenol was mixed with the DNA sample and then vortexed for few seconds. After standing for a few minutes and vortexing again it was centrifuged for 3 minutes. The aqueous layer was removed into a new tube. An equal volume of chloroform was added (chloroform is always 24:1 mixture of chloroform and isoamylalcohol), vortexed then centrifuged. The Upper layer was transferred into a fresh tube.

2.6 Precipitation of DNA.

In order to concentrate DNA in a sample or to change buffers, ethanol precipitation was carried out as follows:- DNA volume was estimated, 0.1 volume of

3M sodium acetate pH 4.8 and 2 volumes of ethanol were added, mixed and kept at -80°C for at least 15 minutes. DNA was recovered by centrifugation at 12000 rpm for 10 minutes. The pellet was washed with 70% ethanol. Finally the pellet was dried under vacuum and resuspended in TE buffer.

2.7 Preparation of the Dialysis tubing.

Dialysis tubing was boiled for 10 minutes in a solution containing 2% Na_2CO_3 (w/v) and 1 mM EDTA (final concentration) pH 8.0. Then the tubing was rinsed five times in cold water, followed by another 10 minutes boiling in water. Finally, the tubing was rinsed several times in water and then stored at 4°C .

2.8 Spectrophotometric quantation of DNA.

The concentration of nucleic acid solutions was measured spectrophotometrically at 260 nm using an Ultrospec 4050. The concentration was calculated using the following formula:-

Optical density at 260 nm of 1 = 50 $\mu\text{g}/\text{ml}$.

2.9 Digestion of DNA with restriction enzymes.

All the digestions with restriction enzymes in this work were carried out overnight unless it is stated elsewhere. The digestion reactions were set up as follows:- 0.2-1 μg of DNA sample was mixed with distilled water to give a volume of 18 μl . 2 μl of the appropriate 10x digestion buffer was added, followed by 1 μl of a restriction enzyme (5 unit/ μl). The contents were mixed and then centrifuged for a few seconds to recover the components to the bottom of the tube followed by incubation at the appropriate temperature. The digestion reaction was stopped by the addition 3 μl of a gel loading dye then loaded into the gel slot. Sometimes, it is necessary to purify the restricted DNA by extraction with phenol/chloroform

and concentrate the DNA by ethanol. Throughout this work, digestion with two restriction enzymes are required for certain purposes, in this case the digestion of DNA sample was performed by adding the two restriction enzymes simultaneously.

2.10 Preparation of restriction enzyme buffers.

Three different salt concentrations are used for the digestion of DNA molecules with restriction enzymes. These are low, medium and high buffers and are prepared in 10 times concentrations.

10x Low salt buffer solution:-

100 mM Tris-HCl pH 7.5, 100 mM MgCl₂ and 10 mM DTT.

10x Medium salt buffer solution:-

500 mM NaCl, 100 mM Tris-HCl pH 7.5, 100 mM MgCl₂ and 10 mM DTT.

10x High salt buffer solution:-

1000 mM NaCl, 500 mM Tris-HCl pH 7.5, 100 mM MgCl₂.

2.11 Agarose gel electrophoresis.

Various percentages of agarose gel could be prepared depending on the size of DNA fragment of interest. Normally, 0.7% agarose is used but concentrations up to 1.5% can be used for the separation of small DNA fragments. The required amount of agarose was added to 200 ml solution of 1x Alec's buffer and boiled for 10 minutes. The gel solution was cooled down to approximately 45°C and poured onto a glass plate (180x150 mm) surrounded by a gel former, the comb was inserted and the gel allowed to set.

The comb and surrounding former were removed and the gel soaked in a gel tank containing 2 litres of 1x Alec's buffer. The DNA samples containing gel dye were loaded into the wells. Then the gel apparatus was connected to the power and then electrophoresed for 3 hours at 80 volts. After electrophoresis, the gel

was taken out from the tank and visualized on a UV transilluminator. The gels were photographed using Polaroid film.

2.12 Acrylamide gels for the retardation assays.

Polyacrylamide gels were used in this work to establish the binding of upsilon protein to particular *ilvYC* DNA sequences. 8% polyacrylamide gels were utilized for this because they show effective separation of DNA fragment in the size range of 60-300 bp used in this assay. The gel was prepared as follows:- To 26.6 ml of 30% acrylamide (29% acrylamide, 1% bisacrylamide), 61.3 ml water, 2.1 ml 3% ammonium persulphate and 10 ml 10x TBE buffer were mixed, Next, the gel solution was degassed, 30 μ l TEMED was added, mixed and the gel solution poured. After polymerization the comb was removed and the gel mounted onto the apparatus. The two reservoirs of the apparatus were filled with 1x TBE buffer. The sample reactions were loaded and the apparatus connected to the power pack. The gel was electrophoresed at 10v/cm for 3-4 hours. After electrophoresis the gel was removed and directly dried after being wrapped with cellophane viscous membrane. The dried gel was then exposed to X-ray film.

2.13 Preparation of SDS-PAGE Gel.

This gel involves preparation of two gels, the acrylamide gel and the stacking gel. Various percentages of this gel could be prepared and it depended upon the molecular weight of the monomer protein. In this work 10% SDS gels were used and prepared as follows:-

Acrylamide gel

22.5 ml 1 M Tris-HCl pH 8.8

20 ml 30% acrylamide solution

15.4 ml sterile water

The solution was degassed then

0.6 ml 10% SDS

150mg Ammonium persulphate

20 μ l TEMED

Then, the solution was poured between two clean glass plates (175x190mm) with spacers until the level of fluid reached four-fifths up the glass plate. Followed by addition of 1 ml isopropanol to obtain an even surface.

Stacking gel

2.5 ml Tris-HCl pH 6.8

2.0 ml 30% acrylamide

14.8 ml sterile water

Degassed then add

0.2 ml 10% SDS

100mg Ammonium persulphate

20 μ l TEMED

The stacking gel solution was poured over the polymerized acrylamide gel then the comb was inserted and the gel left for polymerization. Then, the comb and the bottom spacer were removed and the gel mounted to the apparatus. The two reservoirs were filled with 1x PAGE buffer.

The protein samples were mixed with an equal volume of 2x sample buffer and then heated in a boiling water bath for 3-5 minutes and finally loaded alongside appropriate protein markers. The gel was electrophoresed at 40 volts, overnight.

2.14 Silver staining method for SDS-PAGE gels.

This method of staining has the ability to detect protein bands at very low concentrations. It is much more sensitive than the Coomassie staining method. After the protein gel was removed from the glass plate it was soaked in 50%

MATERIALS AND METHODS

methanol solution for at least 2 hours. Next, the gel was washed with water and stained with solution C for 15 minutes with constant shaking. This was followed by washing the gel in water for 5 minutes then soaking in solution D (developer). It was left in developer until the protein bands appeared then washed with water and placed in 50% methanol solution to stop the developer reaction. Alternatively, developer action could be slowed down by the addition of 5-10% methanol to the developer solution. The gel was then washed with water and dried under vacuum.

Sometimes, the development of the protein bands is too fast and the gel becomes a uniform dark brown in colour. In this case destaining of the gel was carried out by soaking the gel in destain solution containing equal volumes of solutions E and F diluted with 4 volumes of water.

Solution A

0.8g silver nitrate in 4 ml D.water.

Solution B

1 ml of 0.36% NaOH and 0.7 ml of 0.88% ammonia.

Solution C

Solution A mixed with solution B with continuous stirring. A few drops of 0.88% ammonia solution were added to clear the brown colouration, then the volume was completed to 100 ml.

Solution D (Developer)

2.5 ml of 1% citric acid and 0.25 ml 38% formaldehyde were mixed and the volume completed to 500 ml.

Solution E

1.5 M sodium thiosulphate.

Solution F

0.15 M CuSO_4 , 0.6 M NaCl and 0.9 M NH_4OH .

2.15 Protein determination.

Two methods were used to determine protein concentration. These are the Biuret and Bio-Rad microassay procedures.

Biuret assay :- 750 μ l of 1M NaOH was mixed with 50 μ l Benedict's reagent and 50 μ l of protein sample and the volume made to 1 ml. The reaction was left at room temperature for 15 minutes then the protein concentration was estimated using an Ultraspec 4050 at 330 nm. A Sample reaction without protein was used as a blank. A calibration curve of BSA was used.

The Bio-Rad microassay method was used to estimate the concentration of proteins especially in fractions from column chromatography. To 300 μ l of each sample was added 600 μ l diluted dye reagent (1 volume of dye mixed with 4 volumes of water), followed by mixing. The O.D. were measured at 595 nm using an Ultraspec 4050. A calibration curve of BSA was made.

2.16 Purification of DNA restriction fragments from low melting point agarose gels.

The preparation and running of a 0.7% gel was as described in (section 2.11) except that low melting point Agarose was used. The DNA fragment of interest was excised from the gel and transferred to an eppenderoff tube after removing as much as possible of the gel from the fragment. This was then placed in a water bath at 68°C until melted. After that, 2 volumes of TE buffer was added, and left for 5 minutes at 37°C. This was followed by phenol extraction at least two times, then chloroform extraction and the DNA was finally precipitated by ethanol. The pellet was resuspended in an appropriate volume of TE buffer.

2.17 Dephosphorylation of DNA fragments.

Sometimes, it is necessary to remove 5' terminal phosphate groups from DNA

fragments to avoid self ligation of the DNA molecules. The process was carried out as follows:- To 1 μg of restricted DNA fragment 1/10 volume of 10x CIP buffer and 1 unit of alkaline phosphatase were added and the volume made to 18 μl . The sample was incubated for 30 minutes at 37°C. Then, 4.8 μl of 100 mM NTA (nitrilotriacetic acid) was added followed by heating for 15 minutes at 70°C in a water bath. The sample was then phenol/chloroform extracted and finally ethanol precipitated. The pellet was resuspended in an appropriate volume of TE buffer.

2.18 Ligation of DNA fragments.

Ligation of restricted DNA molecules was achieved as follows:- Insert DNA and up to 0.5 μg of vector DNA, 3 μl of 10x ligase buffer and 1 μl of T4 DNA ligase were mixed in a total volume of 30 μl . The ligation reaction was carried out at 15°C, overnight.

A molar ratio of the insert to vector DNA of 2:1 was used to reduce self ligation of the vector. Sometimes, the overnight ligation reaction was followed by phenol/chloroform extraction to increase the efficiency of transformation.

2.19 Transformation and preparation of competent cells.

Transformation of bacterial cells by the calcium chloride technique was carried out as described by Mandel and Higa (1970). Competent cells were prepared as follows:- 50 ml L-broth was inoculated with 1 ml of an overnight culture and then incubated with shaking at 37°C for 2 hours to reach log phase. Cells were harvested by centrifugation for 10 minutes at 6000 rpm at 4°C. The pellet was resuspended in 40 ml of cold transformation buffer and then placed on ice for 1 hour. They were then centrifuged as above and finally resuspended in 1 ml of the same buffer. The cells are now competent for transformation.

To carry out transformation, 100 μ l of competent cells were mixed with at least 0.01 μ g of DNA and kept on ice for 20 minutes. Following a heat shock for 2 minutes at 42°C, 1 ml of L-broth was added and the cells left for 1 hour at 37°C to express antibiotic resistance. Aliquots were then plated on selective media.

2.20 Preparation of Sephadex G-50 and G-25 columns for the separation of DNA from unincorporated nucleotides.

30g of Sephadex G-50 or G-25 (medium) was suspended in 250 ml TE buffer pH 8.0. The suspension was then autoclaved to swell the gel. The solution was then left to cool down to room temperature. The supernatant was removed and replaced with an equal volume of TE fresh buffer pH 8.0. The swollen gel was stored at 4°C.

A disposable 5 ml glass pipette was plugged with a sterile siliconized glass wool. The Sephadex suspension was applied to the column after preloading 5 ml of TEN buffer to the plugged column. The outlet was then opened and the application of Sephadex was continued until the bed level filled four-fifths of the column. The column was then washed with 10 ml TENS buffer. Next, the sample was applied to the column, more TENS buffer was added and fractions of 0.5 ml were collected.

The detection of the radioactivity in the fractions was by liquid scintillation counting. Two peaks were obtained from chromatography through Sephadex. The leading peak consists of the labelled DNA while the trailing peak consists of unincorporated nucleotides. Finally, the peak fractions containing the DNA were mixed and stored at -20°C.

2.21 Nick translation of DNA restriction fragments.

Amersham's nick translation kit was used to label *in vitro* DNA restriction

fragments. Solution I is a nucleotide buffer which contains a mixture of dATP, dGTP and dTTP in a concentrated nick translation buffer. Whereas, solution II is a buffer solution containing DNA polymerase I and DNase (deoxyribonuclease). The reaction was carried out as follows:- To 3 μg purified DNA, 12 μl solution I (nucleotide buffer), 5 μl of $\alpha\text{-}^{32}\text{P}$ -dCTP (specific activity 410 Ci/mmol) and 5 μl solution II (enzyme mixture) were added and the volume completed to 30 μl . The reaction components were mixed and centrifuged for few seconds. The sample was incubated at 15°C in a water bath for 90 minutes. The nick translation reaction was applied to a Sephadex G-50 column to separate the DNA fragment from unincorporated nucleotides as described above.

2.22 ^{32}P Labelling of restriction fragments by "filling" in.

Restriction fragments were labeled by filling the recessed 3' end of double stranded DNA as follows:- To 2 μg DNA fragment, 3 μl of 10x nick translation buffer, 2 μl of each 2 mM unlabeled dGTP, dCTP and dTTP solutions, 2 μl of $\alpha\text{-}^{32}\text{P}$ -dATP (specific activity 410 Ci/mmol) and 2 μl of DNA polymerase I Klenow enzyme (1 unit/ μl), the final volume was completed to 25 μl . The components were mixed, centrifuged for a few seconds and incubated at room temperature for 30 minutes. The DNA was separated from unincorporated deoxynucleotides as described above.

2.23 Labelling of restriction fragments by T4 polynucleotide kinase.

Fragments with protruding 5' termini were labeled as follows:- 3 μg of DNA was dephosphorylated by treatment with Alkaline phosphatase then to this 5 μl 10x kinase buffer, 5 μl of $\gamma\text{-}^{32}\text{P}$ -ATP (specific activity 5000 Ci/mmol) and 20 units of T4 polynucleotide kinase were added and the volume was completed to 50 μl . The contents were mixed well, centrifuged for a few seconds to recover the

components to the bottom of the tube and then incubated for 60 minutes at 37°C. The sample was then applied to a Sephadex G-50 column to separate the DNA from the unincorporated nucleotides as described above.

Fragments with blunt-ends were labeled by the following method. After dephosphorylation of 2 µg of restriction fragment, 4 µl of a solution containing 0.2 M Tris-HCl pH 9.5, 10 mM spermidine and 1 mM EDTA was added and the volume completed to 40 µl. Next, the samples reactions were heated at 70°C a water bath for 5 minutes and then chilled on ice. Next, 5 µl 10x blunt end kinase buffer, 5 µl of $\gamma\text{-}^{32}\text{P}\text{-ATP}$ (specific activity 5000 Ci/mmol) and 20 units of T4 polynucleotide kinase were added to the sample, mixed, centrifuged for a few seconds and incubated for 30 minutes at 37°C. After that, the reaction samples were applied to a Sephadex G-25 column to separate DNA from the unincorporated nucleotides.

2.24 Northern Blotting.

The gel was soaked in 20x SSC for 1 hour. Then over a glass plate a large piece of 3MM paper, soaked in 10x SSC, was placed such that the ends dipped into a tray containing 10x SSC. The paper was then smoothed over to remove any air bubbles. The gel was placed on it and was covered with a wet nitrocellulose filter already soaked in 10x SSC without allowing any air bubbles trapped. Next, 3 layers of wet 3MM paper soaked in 10x SSC and 3 layers of disposable nappies were placed on then a glass plate and a heavy weight. The blott was then left overnight. The nitrocellulose filter was then transferred to 3MM paper and allowed to dry at room temperature, then in a vacuum oven for 3-4 hours at 80°C.

2.25 Preparation of plasmid DNA.

“Mini-Prep” method

The method used was as described by Birnboim and Doly (1979). A 5 ml

MATERIALS AND METHODS

overnight culture of a plasmid bearing strain was centrifuged and resuspended in 180 μ l of solution I then 8 μ l of stock lysozyme solution (50 mg/ml). 400 μ l of solution II was added and vortexed gently. The mixture was placed on ice for 5 minutes after which 300 μ l of 3M sodium acetate was added followed by gentle vortexing. After a further 10 minutes on ice the cells were centrifuged for 5 minutes. 800 μ l of supernatant was removed and mixed with 500 μ l of isopropanol. The mixture was kept at -20°C for 15 minutes, then centrifuged for 5 minutes. The pellet was washed with cold 70% ethanol and then resuspended in 200 μ l TE buffer. After repeated phenol and chloroform extraction the DNA was once more ethanol precipitated and finally resuspended in 100 μ l TE buffer.

Solutions

Lysozyme stock solution

50 mg lysozyme in 1 ml water. This was used at a final concentration of 2 mg/ml.

Solution I

2 % glucose, 80 mM EDTA pH 8.0 and 50 mM Tris-HCl pH 8.0.

Solution II

Freshly made 0.2 M NaOH and 1 % SDS.

50 ml-Scale preparation of plasmid DNA.

This essentially a modified scaled up of the method of Birnboim and Doly (1979). 50 ml L-broth of was inoculated with a single colony in the presence of a suitable antibiotic and incubated overnight with shaking at 37°C . The cells were then centrifuged at 8000 rpm for 10 minutes at 4°C . The pellet was resuspended in 2 ml of solution I with lysozyme (5 mg/ml), left for 10 minutes at room temperature then 4 ml of solution II was added, mixed by inversion and left on ice for 10 minutes. This was followed by adding 3 ml of cold 5M potassium acetate pH 4.8 and then left on ice for another 10 minutes. The lysed bacteria

were then centrifuged as above. The supernatant was removed and extracted with phenol/chloroform (equal volume). The aqueous layer was transferred to a clean corex tube and the DNA was precipitated with 2 volumes of ethanol. After centrifugation the pellet was washed with 70% ethanol, dried then resuspended in 500 μ l TE buffer.

Large scale preparation of plasmid DNA.

In order to obtain more purified plasmid DNA and in high yield, Caesium chloride-Ethidium bromide purification was used. A 500 ml culture of bacteria was grown at 37°C with vigorous shaking after inoculation with a 5 ml overnight culture. After the culture reached stationary phase the cells were harvested by centrifugation in an MSE 18 at 7000 rpm for 10 minutes. The supernatant was removed and the pellet resuspended in 10 ml solution I. The suspension was transferred to a clean 250 ml beaker, followed by addition of 50 mg lysozyme and was left for 10 minutes at room temperature. Then 20 ml of solution II was added and left on ice for 10 minutes then finally 15 ml of 3M sodium acetate was added and left for another 10 minutes. The lysed cells were clarified by centrifugation in an MSE 18 for 45 minutes at 15000 rpm at 5°C. The supernatant was retained and 30 ml of isopropanol was added, mixed and left for 30 minutes at room temperature. The precipitated nucleic acids were recovered by centrifugation for 30 minutes at 12000 rpm at room temperature. The pellet was retained and resuspended in 15 ml TE buffer. Following that purification through a Caesium chloride-Ethidium bromide gradient was carried out as follows. One gram of Caesium chloride for each 1 ml of resuspended pellet was added and mixed until the Caesium chloride dissolved completely. Then 0.3 ml of 10 mg/ml solution of ethidium bromide was added and the solution transferred to a measuring cylinder to complete the volume to 27 ml with TE buffer containing the same concentration of Caesium chloride

used above. The 27 ml solution was transferred into a Quickseal centrifuge tube and centrifuged overnight at 44000 rpm at 15°C using a Sorvall ultracentrifuge and a VTi 50 rotor.

Two bands appeared in the gradient, the lower one comprises the plasmid DNA. This band was removed using a syringe and the solution obtained was extracted with an equal volume of isopropanol saturated with Caesium chloride to remove the ethidium bromide. This process was repeated until no ethidium bromide remained in the solution. Finally, the extracted DNA solution was dialysed against TE buffer for 4-5 changes at 4°C, overnight.

Sometimes it is necessary to recentrifuge the DNA through a CsCl-EtBr gradient to obtain highly purified plasmid DNA and this was performed as follows:- After removing the plasmid DNA from the gradient it was transferred to another a Quickseal tube and the volume completed to 27 ml with TE buffer plus CsCl as before, followed by centrifugation under the same conditions described above.

2.26 Transformation of recombinant M13mp18 and M13mp19 vectors.

Competent cells from JM101 were prepared and transformed following the general method for transformation (see section 2.19) except the buffer is 50 mM CaCl₂ only.

Plating the transformed cells was carried out as follows:- to each tube containing 3 ml of top agar at 42°C, the transformed cells (after heat shock), 200 µl of fresh log-phase of JM101, 10 µl of 100 mM IPTG and 20 µl of 10% X-gal were added. These components were mixed and the entire contents plated onto L-agar plates. After the top agar had set the plates were incubated at 37°C.

2.27 Preparation of single-stranded M13mp DNA.

50 ml of 2x YT media was inoculated with 100 µl of an overnight JM101

MATERIALS AND METHODS

culture. Then 2 ml of this was transferred into 12 sterilized universal bottles. 11 white and 1 blue (which was used as a standard) plaques were picked and resuspended into the 12x2 ml bottles. These were incubated at 37°C.

1.5 ml was transferred from each bottle to an eppenderoff tube and the cells pelleted. The supernatant was taken from each tube and recentrifuged to remove any remaining cells. 1 ml from each sample was transferred to a new tube and 200 μ l of 20% polyethylene glycol 2.5 M NaCl was added, then left at room temperature for 20 minutes followed by centrifugation for 10 minutes at 13000 rpm. The supernatant was discarded and the tubes recentrifuged for 10 seconds, the remaining polyethylene glycol was removed by a micropasteur pipette. The pellets were resuspended in 100 μ l TE buffer, then phenol extracted followed by chloroform extraction. The DNA was precipitated by ethanol. Finally, the DNA pellet was resuspended in 20 μ l TE buffer.

To ensure that single-stranded DNA prepared from recombinant plaques (white) contained inserts before using them as template DNA in the sequencing reactions the following test was done. A 2 μ l of the DNA samples from each set was loaded alongside M13 single stranded DNA purified from a blue plaque, which acts as a size control, to a 0.7% agarose gel, electrophoresed and then examined under UV light. The migration of the DNA fragment from the samples prepared should be slower than the DNA fragment that acts as a control. This difference in migration depends on insert size.

2.28 DNA Sequencing.

For each sequencing reaction there are four incubations. Each incubation contains all the deoxynucleoside triphosphates, α -³⁵S -dATP and one of the four dideoxynucleoside triphosphates. The sequencing reactions was carried out as follows:- In a 1.5 ml eppenderoff tube the following components were mixed, 5 μ l

(1 μg) of the DNA template (M13 mp18 or M13 mp19 single-stranded DNA with insert), 2 μl (4ng) universal M13 primer (15 bases), 1 μl 10x polymerase reaction buffer and 4.5 μl sterile water. The total volume is 12.5 μl . These components were centrifuged for a few seconds to collect the components in the bottom of the tube. Then, the sample reaction was heated at 85-95°C in a water bath for 5 minutes. The sample reaction was then slowly cooled down to room temperature to allow annealing of the primer to the DNA template. 2 μl of A⁰ G⁰, C⁰ and T⁰ mixes were added to four fresh tubes. Then, to the annealed template-primer mixture, 1 μl of α -³⁵S-dATP (specific activity 410 Ci/mmol), 1 μl of 0.1 M DTT and 1 μl of Klenow enzyme (1 unit/ μl) were added and 4 μl aliquots from this mixture were dispensed to the four eppendorff tubes (A⁰, G⁰, C⁰ and T⁰). The contents of each tube were mixed and centrifuged for a few seconds. These incubation mixtures were left for 15 minutes at 30°C. Then 1 μl of a chase solution was added to each incubation. This chase solution serves to extend the polynucleotide chains which have prematurely terminated due to low dATP concentration. After a further 15 minutes incubation at 30°C, 5 μl of formamide-dye mix was added to each incubation.

2.29 Preparation of DNA sequencing gels and electrophoresis of samples.

An 8% Polyacrylamide sequencing gel was prepared as follows:-

Two 20x42 cm glass plates were used for preparing the sequencing gel. These plates were washed first with teepol and water, wiped dry then washed with ethanol. Further, one surface of one plate was siliconized with "Repelcote". The two plates with 0.4 mm spacers were taped around the edges with gel sealing tape. Then, 42g Ultrapure urea, 20 ml 40% acrylamide (19:1 acrylamide-bisacrylamide), 20 ml 10x TBE buffer and 35 ml water. The above components were mixed and

urea was dissolved to completion. Fines were removed by filtration. Following degassing, to remove air from the gel solution, 0.8 ml of 10% ammonium persulphate and 20 μ l of TEMED were added.

The gel solution was poured between the glass plates using a 50 ml syringe. The gel plates were held at an angle of 45⁰ so that the gel solution flows evenly into the plates. After addition of the well-forming comb, the gel was allowed to set.

The tape around the bottom and front edges was removed from the gel, then it was fixed to the apparatus along with an aluminum plate which ensures an even distribution of temperature. The upper and lower reservoirs were filled with 1x TBE buffer and the gel was pre-electrophoresed for 30 minutes before loading the sequencing reaction samples.

The gel wells were rinsed with the running buffer and the DNA samples were heated at 95-100°C for 3-5 minutes then loaded immediately into the gel wells in the order A, G, C and T. The gel was electrophoresed at 1500v for about 3 hours until the bromophenol blue reached the bottom. Further samples were loaded into empty tracks and the gel electrophoresed for an additional 2 hours. After electrophoresis the gel was transferred to 3MM paper and dried for at least 3 hours at 80°C on a vacuum gel drier.

After the sequencing gel was dried it was exposed to preflashed X-ray film. Autoradiography was carried out overnight at room temperature. X-ray films were developed in Phenisol developer (Kodak) and fixed in Kodafix (Kodak).

2.30 Deoxynucleoside triphosphate (dNTP) and dideoxynucleoside triphosphate (ddNTP) mixes.

The dNTPs are provided in stock solutions of 10 mM. Working solutions of 0.5 mM were prepared from these stock solutions then A, G, C and T mixes

prepared from these working solutions as follows:-

A mix

20 μ l 10x polymerase reaction buffer, 20 μ l from each working solution (0.5 mM) of dGTP, dTTP and dCTP.

G mix

20 μ l 10x polymerase reaction buffer, 1 μ l of 0.5 mM dGTP and 20 μ l from each working solution (0.5 mM) of dTTP and dCTP.

C mix

20 μ l 10x polymerase reaction buffer, 1 μ l 0.5 mM dCTP and 20 μ l from each working solution (0.5 mM) of dGTP and dTTP.

T mix

20 μ l 10x polymerase reaction buffer, 1 μ l 0.5 mM dTTP and 20 μ l from each working solution (0.5 mM) of dGTP and dCTP.

The above mixes were designed for the use with α -³⁵S dATP.

The dideoxynucleoside triphosphate are provided in stock solutions of 10 mM.

Working solutions of ddNTPs of various concentrations were prepared and from these solutions the ddNTP mixes were prepared as follows:-

ddATP mix

1 μ l 0.125 mM ddATP was diluted to 80 μ l by sterile water.

ddGTP mix

1 μ l 0.5 mM ddGTP was diluted to 20 μ l by sterile water.

ddCTP mix

1 μ l 0.5 mM ddCTP was completed to 20 μ l by sterile water.

ddTTP mix

1 μ l 1 mM ddTTP was diluted to 10 μ l by sterile water.

New mixes were prepared by mixing equal volumes of dNTPS and their corresponding ddNTPs mixes. These mixes were given the following symbols:- A⁰, G⁰,

C⁰ and T⁰. Additionally, chase solution was prepared by diluting 10 mM dATP stock solution to 0.5 mM with sterile water.

2.31 Preparation of S-30 extracts.

The preparation of S-30 extracts was initially that of Wild *et al.*, (1977) and was as follows :- 1 litre of S-30 media was inoculated with a 5 ml culture of CU827 and grown at 30°C, overnight with vigorous shaking. The cells were harvested at 4°C. The weight of the cells was determined and then resuspended in buffer III (10 ml/g cells), centrifuged and resuspended in buffer III (4 ml/g cells), centrifuged and resuspended in buffer III (1.3 ml/g cells). All centrifugation was at 4°C.

The cells were lysed in a precooled cell under 1500 p.s.i. and the lysed cells were collected on ice. The lysed cells were passed through a fine needle to shear the DNA and the passage through the pressure cell repeated. After the addition of 10 µl of 0.1 M DTT for each ml of lysate the suspension was clarified by centrifugation at 15000 rpm in an MSE18 centrifuge using a 8x50 rotor for 30 minutes at 4°C. The supernatant was retained and centrifuged again under the same conditions.

To the supernatant 0.16 ml of preincubation mixture per gm of the original pellet was added and incubated in a light protected vessel at 37°C for 80 minutes. This allows translation of endogenous mRNA and frees ribosomes. The lysate was then dialysed for 3 changes in buffer III at 4°C then distributed into aliquots in a screw cap vials, frozen in liquid nitrogen and stored at -80°C.

2.32 Solutions and media recipes for *in vitro* *ilvC* expression.

S-30 media

5.6g KH₂PO₄ (anhydrous)

28.3g K₂HPO₄ (anhydrous)

10g Bacto-Yeast extract

10-15mg thiamine

These components were mixed and dissolved in 1 litre water, autoclaved then 40 ml of 25% glucose was added.

Buffer III

Tris-acetate pH 8.2 10 mM

Mg-acetate 14 mM

K-acetate 60 mM

DTT 1 mM

Preincubation mixture

Tris-acetate pH 8.2 0.75 mM

Mg-acetate 21 mM

DTT 7.5 mM

A solution of 20 amino acids 75 μ M each

ATP 6 mM

Phosphoenol pyruvate 67.5 mM

Pyruvate kinase 20 μ g/ml

Solution 2 (Supplement solution from the kit)

This contains sufficient nucleotides for transcription, tRNA for translation, an energy-generating system and inorganic salts.

Solution 3

It contains an equimolar mixture of amino acids minus methionine.

Solution 5

It is a dilution buffer.

2.33 Bacterial strains and plasmids.

Bacterial strains and plasmids used in this work are listed in Tables 2.1 and 2.2 respectively. Constructed plasmids used in this work are listed in Table 2.3.

Table 2.1 Bacterial Strains

Strain	Description	source of reference
A	it is C600 A ⁻ contains pGP1-2 and pDUB2411.	This work
B	it is C600 A ⁻ contains pGP1-2 and pDUB2412.	This work
C	It is JM83 contains the <i>ilvY^c</i> gene.	This work
CU152	<i>F⁻, ara Δ(lacpro) thi.</i>	Cold Spring Harbor Laboratory.
CU176	<i>F⁻, ara Δ(lacpro) thi, Sm^R.</i>	Cold Spring Harbor Laboratory.
CU827	<i>Δ ilvDAC115 ara thi, Δ(proAB-lac).</i>	M. Watson et. al. (1979).
CU962	<i>ilvY2143 rbs-215 ara thi, Δ(pro-lac).</i>	Biel and Umbarger (1980).
JM83	<i>ara, Δ(lac-proAB) rpsL, (=strA), φ80, ΔM15.</i>	Vieira and Messing (1982).
JM101	<i>supE, thi, Δ(lac proA,B), F', traD36, proA,B, lac^RZ M15.</i>	Yanish-Perron et. al. (1985).
MC4100	<i>F' araD139 Δ(argF-lac)U169 rpsL150 relA1 deoC1 ptsF25 rbsR flbB5301.</i>	University of Texas.

Table 2.2 Plasmids

Plasmids	Description	Source of Reference
pBR328	Wild-type plasmid <i>bla</i> ⁺ , <i>tet</i> ⁺ , <i>Cm</i> ⁺ .	Soberon et. al. (1980).
pDUB2007	<i>RP4::ilvC::lacZ</i> , <i>Km</i> ⁺ .	M. Watson (1979).
pDUB2012	<i>RP4::ilvC::lacZ</i> and <i>ilvY::Tn7</i> .	M. Watson (1979).
pDUB2013	<i>RP4::ilvC::lacZ</i> , <i>ilvY</i> ⁺ , <i>Tn7</i> .	M. Watson (1979).
pGEM-Blue	it is the first SP6/T7 pGEM vector, <i>bla</i> ⁺	"Promega" Biotech."
pGP1-2	it contains two promoters, <i>P_{lac}</i> and <i>P_L</i> , <i>Km</i> ⁺ .	Tabor and Richardson (1985).
pGMM201	<i>pBR322</i> Ω10[3.612kb:λ80 <i>ilv ilvAYC-0.27-4.35kb(+)</i>].	McCorkle et. al. (1978).
pMC1871	It is a fusion vector, <i>Tet</i> ⁺ , contains a promoterless <i>lacZ</i> fragment.	"Promega" Biotech."
pUC19	wild-type plasmid <i>bla</i> ⁺ .	Yanisch-Perron et. al. (1985).

Table 2.3 Constructed plasmids used in this work.

Plasmid	Description	Source of reference
pDUB2400	EcoRI <i>ilvY^C</i> fragment in pUC19.	This work
pDUB2401	EcoRI <i>ilvY⁺</i> fragment in pUC19.	This work
pDUB2402	Small EcoRI-BglIII fragment of wild-type <i>ilvY</i> gene and large BglIII-EcoRI fragment of constitutive <i>ilvY</i> gene into EcoRI of pUC19.	This work
pDUB2403	Small EcoRI-BglIII fragment of constitutive <i>ilvY</i> gene and large BglIII_EcoRI fragment of wild-type <i>ilvY</i> gene into EcoRI site of pUC19.	This work
pDUB2404	BglIII-HindIII fragment from pDUB2401 into pDUB2400 after removing its BglIII-HindIII fragment.	This work
pDUB2405	BglIII-HindIII fragment from pDUB2400 into pDUB2401 after removing its BglIII-HindIII fragment.	This work
pDUB2406	BamHI Kan ^R fragment from pDUB2033 into pDUB2405.	This work
pDUB2407	BamHI Kan ^R fragment from pDUB2033 into pDUB2404.	This work
pDUB2408	BamHI Kan ^R fragment from pDUB2033 into pDUB2401.	This work
pDUB2409	BamHI Kan ^R fragment from pDUB2033 into pDUB2400.	This work

Plasmid	Description	Source of reference
pDUB2411	EcoRI <i>ilvY^C</i> fragment in pGEM-Blue. (Orientation 1)	This work
pDUB2412	Opposite orientation to pDUB2411.	This work
pDUB2414	EcoRI <i>ilvY⁺</i> fragment in pBR328. (Orientation 1)	This work
pDUB2415	Opposite orientation to pDUB2414.	This work
pDUB2416	PstI-BglIII <i>ilvC</i> fragment from pGMM201 in pUC19.	This work
pDUB2417	<i>lacZ</i> fragment from pMC1871 in pDUB2414.	This work
pDUB2418	<i>lacZ</i> fragment from pMC1871 in pDUB2415.	This work

Bacteriophages

Bacteriophages	Description	Source of reference
M13mp18	Multiple cloning sites.	BRL
M13mp19	Multiple cloning sites.	BRL

CHAPTER 3

Characterization of the *ilvY^c* mutation**3.1 Introduction.**

This chapter seeks to characterize the *ilvY*²¹⁴³ mutation that makes *ilvC* expression in the strain CU962 constitutive (Biel and Umbarge, 1980). The first question addressed was, does *ilvY* regulate *ilvC* expression at the level of transcription. Secondly, the nature of the mutation in the *ilvY^c* mutant gene was determined by sequencing both *ilvY⁺* and *ilvY^c* and comparing the nucleotide sequences. Lastly, experiments were carried out to confirm that the observed single base change is responsible for making *ilvC* expression constitutive due to the *ilvY* product (upsilon) no longer needing the substrates α -acetoxybutyrate or acetolactate for induction of *ilvC*.

3.2 Does upsilon act at the transcriptional level?

The experiment involves the extraction of RNA from different bacterial strains (CU152 wild-type, CU827 $\Delta ilvDAYC$, CU962 *ilvY^c*) grown under repressing and derepressing conditions. The RNA was characterized by Northern blotting then the hybridization of these RNA species to a radiolabelled *ilvC* probe.

The following sets of cultures in different media were prepared :-

1. 10 ml cultures of CU962, CU827 and CU152 in 2x YT media.
2. 10 ml cultures of CU962, CU827 and CU152 in M9 minimal media containing isoleucine, leucine and valine (repression media).
3. 10 ml cultures of CU962, CU827 and CU152 in M9 minimal media grown as follows. After growth in M9 containing isoleucine, leucine and valine, the cultures were centrifuged then resuspended in minimal media without the above amino acids, except CU152 in which valine was included. This acts to derepress the

expression of all *ilv* genes. The cultures were grown for at least 3 hours at 37°C after derepression.

The RNA was extracted from the cultures by the following method:- 1 ml was transferred from each culture to a 1.5 ml eppendorff tube, centrifuged and resuspended in 100 μ l lysis buffer (20% sucrose, 20 mM Tris-HCl pH 7.6, 10 mM EDTA and 50 mM NaCl) plus 10 μ l of 10 mg/ml lysozyme. The cells were kept on ice until the suspensions became viscous (at least 10 minutes). The lysis was completed by the addition of 100 μ l of 2% SDS. Next, 10 μ l of 5 mg/ml Proteinase K was immediately added and the samples were vortexed. Finally, the lysate samples were frozen and thawed twice by using liquid nitrogen and a water bath at 45°C then 50 μ l of loading dye (40% glycerol, 100 mM EDTA and 0.1% bromophenol blue) was added.

To determine the integrity of RNA in the lysate samples. 6 μ l aliquotes from each sample were electrophoresed in 1% agarose gels. Two bands were visible in the gel which represent the ribosomal RNA. Then the lysate samples were used for RNA analysis.

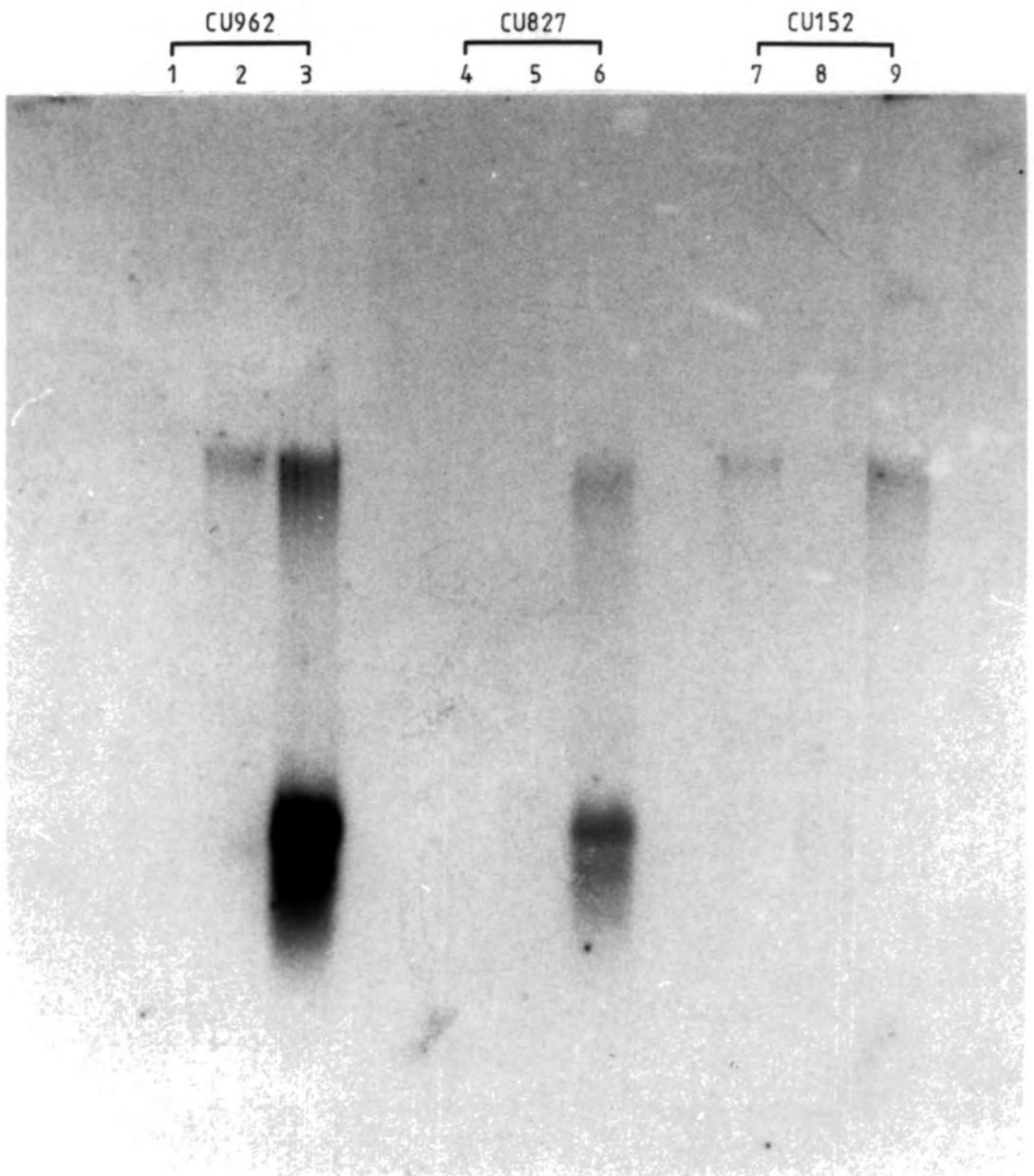
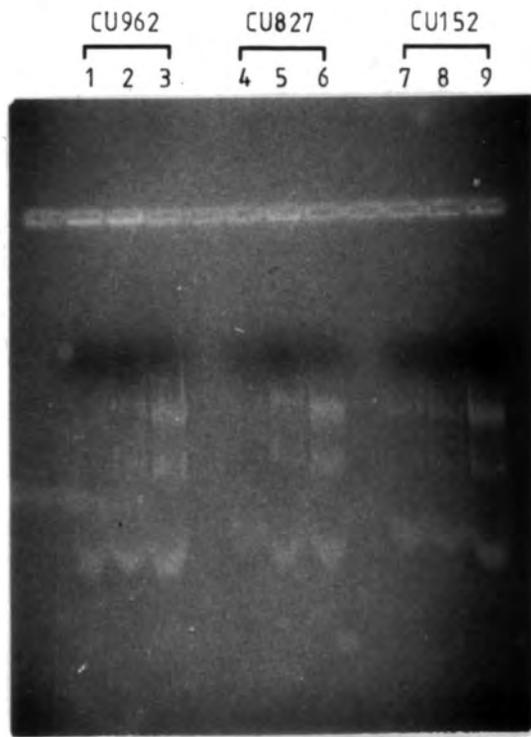
The RNA species were separated on an agarose-formaldehyde gel which was prepared as follows:- 1g agarose in 70 ml water was autoclaved then cooled to 65°C and the following components were added, 20 ml MOPS running buffer (0.2 M MOPS, 50 mM sodium acetate, 5 mM EDTA) and 10 ml formaldehyde, mixed and poured into a horizontal gel appartus. After the gel had set it was immersed in 1x TBE buffer. Before loading, the lysates samples were denatured at 65°C in a water bath for 15 minutes and the gel wells were rinsed. Electrophoresis was carried out in a fume hood. After electrophoresis the gel was viewed with a UV transilluminator and photographed (see Fig.3.1).

The nucleic acids were transferred to a nitrocellulose filter by Northern Blotting (see section 2.24).

Fig. (3.1).

A. 1% agarose-formaldehyde gel electrophoresis for the separation of RNA species extracted from three bacterial strains, CU152 , CU827 and CU962 under different growth conditions.

B. No hybridization was detected with any RNA extracted from cells grown with M9 media under both repressing and non-repressing conditions (lanes 1, 2, 4, 5, 7, & 8). No hybridization was detected with RNA from CU152 grown with 2x YT (lane 9). Hybridization was detected to RNA from strain CU962 (lane 3) and CU827 (lane 6) grown with 2x YT.



RESULTS

In order to carry out the hybridization, a radiolabelled probe the EcoRI *ilvC* fragment was prepared by a technique known as random priming. 1 μg of the EcoRI fragment was denatured by heating it for 10 minutes at 95°C and subsequent cooling on ice. The DNA was made radioactive by synthesising the second strand on the denatured DNA using sheared salmon sperm DNA as random primers and incorporating $\alpha\text{-P}^{32}\text{-dCTP}$. Unincorporated nucleotides were removed by applying the reaction sample to a Sephadex G-50 column. The probe was labelled to a specific activity of 1.4×10^7 cpm/ μg .

The baked nitrocellulose filter was floated on the surface of 6x SSC until wet from beneath, then it was immersed in the solution for 2 minutes then transferred to a plastic bag with 0.2 ml of prehybridization buffer warmed at 68°C for each cm square of the filter. The filter was incubated for 2-4 hours at 68°C in a water bath. Next, the prehybridization solution was exchanged with hybridization buffer without allowing any air bubbles to be trapped, the bag sealed and incubated further at 68°C overnight. Then the filter was washed to remove unbound radioactive material, in 200 ml buffer containing 2x SSC and 0.5% SDS for 30 minutes at 68°C then in 200 ml containing 0.1x SSC and 0.5% SDS for 30 minutes at 68°C . The filter was then placed on 3MM paper for drying, wrapped in cling film and applied to X-ray film for autoradiography.

After hybridization the filter was washed, dried and exposed to X-ray film. Examination of the autoradiograph (see Fig.3.1) shows that hybridization has occurred. The extent of hybridization is proportional to the amount of mRNA loaded. Most hybridization occurs with RNA from strain CU962 (bearing the *ilvY^c* mutation). Less occurs to CU827 ($\Delta ilvDAYC$) and very little to CU152 (wild-type).

3.3 Nucleotide sequencing of the *ilvY* gene.

In order to sequence the *ilvY^c* gene, it was first subcloned from plasmids pP1, pH3 and pH4 (M. Watson, unpublished results). These carry either PstI or HindIII genomic fragments of CU962 in pBR322. They were characterized as being able to complement *ilvA⁻* strains and also to induce an *ilvC::lacZ* fusion strain in the absence of inducers (*ilvY^c* phenotype) (M. Watson, unpublished results). pP1, pH3 and pH4 were digested with EcoRI and electrophoresed on a 0.7% agarose gel. The *ilvY* EcoRI fragment (see Fig.3.2) was identified by comparison with digested pGMM201 (see Fig.3.3A), purified and cloned into pUC19. The resulting plasmid is pDUB2400. pDUB2401 is the wild-type EcoRI *ilvY* fragment from pGMM201 into pUC19.

Both the *ilvY* wild-type and mutant genes were sequenced using the Sanger chain termination method (1977). The two recombinant plasmids, pDUB2400 and pDUB2401, were restricted with EcoRI and then BglII and the digested molecules were loaded onto a 0.7% low melting point agarose gel, electrophoresed and the generated *ilvY* DNA fragments (small 220 bp and large 1000 bp) from each digestion were purified (see Fig.3.3B). 4 fragments in total.

Then, each of the small and large fragments were ligated to EcoRI plus BamHI restricted M13mp18 and M13mp19. These ligated plasmids were transformed into JM101 (see section 2.26), screened for recombinant plaques followed by preparation of recombinant single stranded M13mp18 and M13mp19 (see section 2.27). Sequencing of each different insert cloned into the phage vectors (M13mp18 and M13mp19) was then carried out. The entire sequence of the small 200 bp BglII-EcoRI fragment was determined in both directions using just these clones.

Fig. (3.2).

Restriction endonuclease map of the *ilvYC* system.

300 bp

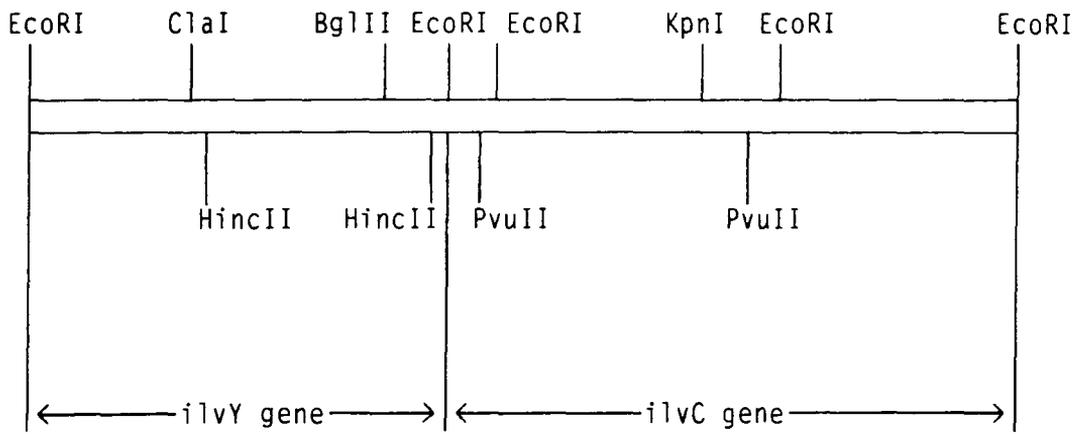


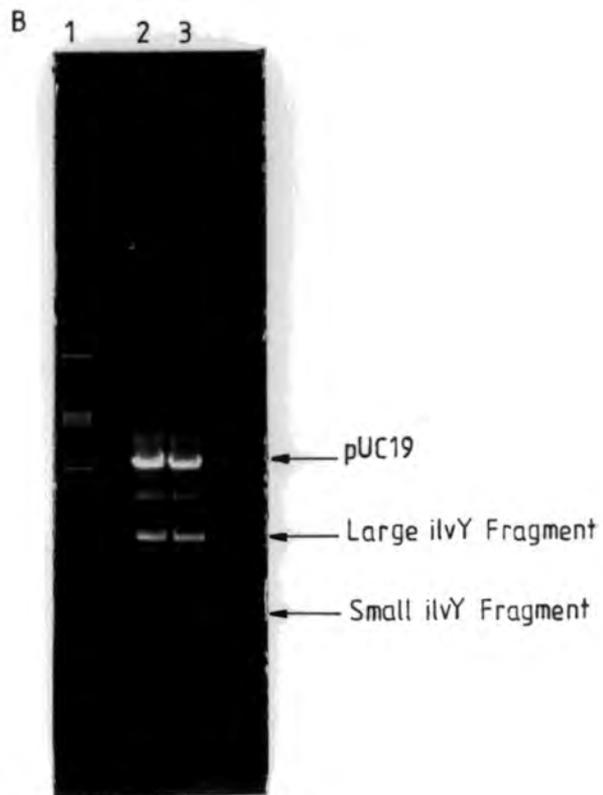
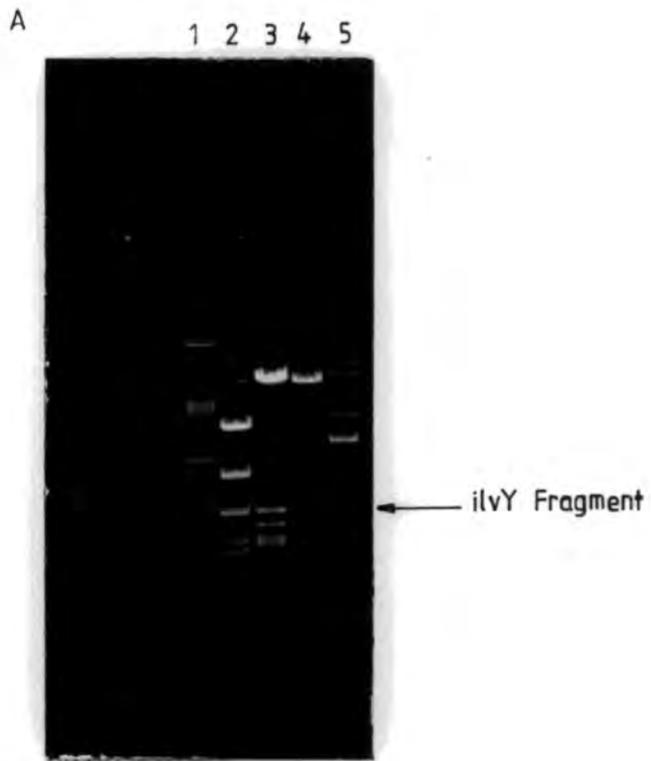
Fig. (3.3).

A. 0.7% agarose gel electrophoresis of cleaved pGMM201 (lane 2) acting as a size control, pP1, pH3 and pH4, which carry *ilvY^c* gene(lanes 3, 4 & 5), digested with EcoRI. Lane 1 is λ DNA digested with PstI.

The Arrow indicates the position of the *ilvY* fragment.

B. Lane 2 and 3 represent digestion of pDUB2400 and pDUB2401 with EcoRI plus BglII.

Lane 1 is λ DNA marker.



3.4 Examination of the *ilvY* nucleotide sequences.

After comparing the nucleotide sequences of the non-transcribed DNA strand for the EcoRI-BglII small fragment from the wild-type and constitutive mutation, a single base change was detected between them. The wild type sequence was found to be identical to that reported by Wek and Hatfield (1986). This base change involves a guanine (G), in position 87 (counted from the transcription start point) of the *ilvY*⁺ gene, to an adenine (A) transition in the *ilvY*^c mutant gene (see Fig.3.4). No additional base changes were observed. This base change alters the codon GAA, which codes for glutamine, to the codon AAA that codes for lysine in the *ilvY*^c gene (see Fig.3.4). The reading frame remains intact. The base change which has been detected in the *ilvY*^c gene is located towards the 5' end.

No differences between the sequences obtained from the large EcoRI-BglII fragment for either gene or with that found by Wek and Hatfield were observed.

3.5 Is the single base change responsible for the constitutive phenotype?.

To verify that the base substitution that occurs in the small EcoRI-BglII fragment of the *ilvY*^c mutant gene is totally responsible for making the the gene product independent of substrate α -acetoxybutyrate or acetolactate for induction of *ilvC* expression, the following experiments were conducted.

The first experiment, illustrated in Fig.3.5, involved exchanging the small and large EcoRI-BglII fragments of both wild-type and mutant *ilvY* genes with each other and transforming the reconstructed plasmids into *ilvC*⁺ strains in an attempt to induce constitutive *ilvC* gene expression.

pDUB2400 (*ilvY*^c) and pDUB2401 (*ilvY*⁺) were restricted separately with EcoRI to generate the *ilvY* fragments. The total digests were dephosphorylated

Fig. (3.4)

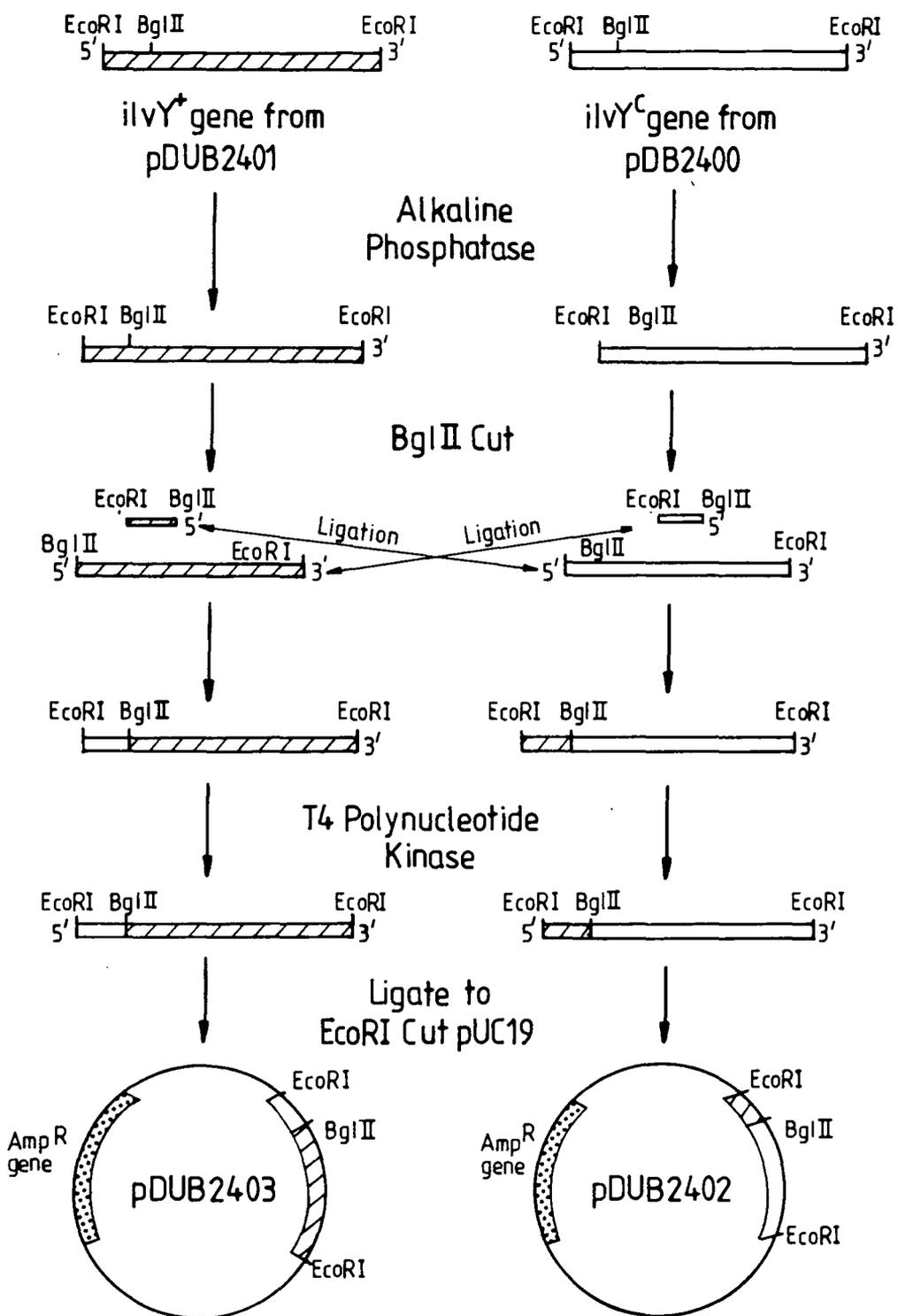
Nucleotide sequence of the wild-type *ilvY* gene and the deduced amino acid sequence. The base substitution which changes the codon from glutamine to lysine in the mutant *ilvY* gene is indicated. Location of restriction enzymes cleavage sites, ribosome binding site and the transcription startpoint for either *ilvC* & *ilvY* are underlined.

GCT	GCC	TAC	AGC	CAT	CTG	CCG	CCG	ATT	110
Ala	Ala	Tyr	Ser	His	Leu	Pro	Pro	Ile	CTG
									Leu
GAT	CGC	TTC	CGC	GCG	GAA	CAC	CCG	TCG	120
Asp	Arg	Phe	Arg	Ala	Glu	His	Pro	Ser	GTG
									Val
GAG	ATT	AAA	CTT	ACT	ACT	GGT	GAT	GCG	130
Glu	Ile	Lys	Leu	Thr	Thr	Gly	Asp	Ala	GCA
									Ala
GAT	GCG	ATG	GAA	AAG	GTG	GTC	ACT	GGT	140
Asp	Ala	Met	Glu	Lys	Val	Val	Thr	Gly	GAA
									Glu
GCG	GAT	CTG	GCG	ATT	GCG	GGT	AAA	CCG	150
Ala	Asp	Leu	Ala	Ile	Ala	Gly	Lys	Pro	GAA
									Glu
ACC	TTG	CCC	GGC	GCA	GTG	GCG	TTT	TCG	160
Thr	Leu	Pro	Gly	Ala	Val	Ala	Phe	Ser	ATG
									Met
CTG	GAG	AAT	CTG	GCA	GTA	GTG	CTG	ATT	170
Leu	Glu	Asn	Leu	Ala	Val	Val	Leu	Ile	GCC
									Ala
CCC	GCG	CTG	CCC	TGC	CCG	GTG	CGT	AAT	180
Pro	Ala	Leu	Pro	Cys	Pro	Val	Arg	Asn	CAG
									Gln
GTG	TCG	GTA	GAG	AAG	CCG	GAC	TGG	TCA	190
Val	Ser	Val	Glu	Lys	Pro	Asp	Trp	Ser	ACG
									Thr
GTG	CCG	TTT	ATT	ATG	GCC	GAT	CAG	GGG	200
Val	Pro	Phe	Ile	Met	Ala	Asp	Gln	Gly	CCG
									Pro
GTA	CGC	CGC	CGC	ATT	GAA	CTG	TGG	TTT	210
Val	Arg	Arg	Arg	Ile	Glu	Leu	Trp	Phe	CGA
									Arg
CGC	AAT	AAA	ATC	AGT	AAC	CCG	ATG	ATT	220
Arg	Asn	Lys	Ile	Ser	Asn	Pro	Met	Ile	TAC
									Tyr
GCC	ACG	GTT	GGC	GGG	CAT	GAA	GCG	ATG	230
Ala	Thr	Val	Gly	Gly	His	Glu	Ala	Met	GTA
									Val
-- Clal --									240
TCG	ATG	GTG	GCA	CTC	GGC	TCT	GGC	GTG	GCA
---	---								
Ser	Met	Val	Ala	Leu	Gly	Cys	Gly	Val	Ala

TTG	TTG	CCG	GAA	GTG	GTG	CTG	GAA	AAC	250
Leu	Leu	Pro	Glu	Val	Val	Leu	Glu	Asn	AGC
									Ser
CCC	GAA	CCG	GTG	CGT	AAC	CGC	GTG	ATG	260
Pro	Glu	Pro	Val	Arg	Asn	Arg	Val	Met	ATT
									Ile
TTA	GAG	CGC	AGC	GAT	GAG	AAA	ACG	CCG	270
Leu	Glu	Arg	Ser	Asp	Glu	Lys	Thr	Pro	TTT
									Phe
GAG	CTG	GGC	GTG	TGT	GCG	CAG	AAA	AAG	280
Glu	Leu	Gly	Val	Cys	Ala	Gln	Lys	Lys	CGG
									Arg
CTA	CAT	GAG	CCG	CTA	ATT	GAG	GCA	TTC	290
Leu	His	Glu	Pro	Leu	Ile	Glu	Ala	Phe	TGG
									Trp
AAG	ATT	TTG	CCG	AAC	CAC	297	TGA	CGTTGTCGC	
Lys	Ile	Leu	Pro	Asn	His	AAA	STOP		
						Lys			
GCGGGTAGG		CCTGATAAG		CGAAGCGCT		ATCAGGCAT		TTTTCCCTA	
ACCCGCCAA		AAAGAACCT		GAACGCCGG		GTTATTGGT		TTCGTCGTG	
GCAATCGTA		GCCCAGCTC		ATTCAGCCG		GGTTTCGAA		ATCCGGTTC	
ATGGTCGCC		AAGTTCGAA		CGCCGCCAG		TACGCGCCC		GTAGTCGGT	
GCCATGGCT		CCGATAGTG		GAACAAAGA		AATGTTCCA		GTACGTACC	
CAGCGTGTT		GAGGAAGCG		CAGCAGCGC		GCCCGGTGA		TTCCGG	
-- EcoRI--									
GAA	TTC.								
---	---								

Fig. (3.5).

Scheme for the strategy of the first attempt to verify that the single base change in the small *ilvY* fragment is responsible for making the *ilvC* expression constitutive.



(see section 2.17) and then restricted with BglIII to generate the small and large DNA fragments from each fragment. These DNA fragments were electrophoresed and then purified from a low melting point agarose gel (see section 2.16).

Ligation was now carried out as follows. Firstly, between the small fragment (EcoRI-BglIII) of *ilvY*⁺ gene and the large fragment (BglIII-EcoRI) of *ilvY*^c. Secondly between the small fragment of *ilvY*^c and the large one of *ilvY*⁺. The two ligation mixtures were loaded onto a low melting point agarose gel alongside the EcoRI *ilvY* fragment as a size control. The two bands from both ligation mixtures which migrated at the same rate as the control DNA fragment were excised and purified (see section 2.16). These purified DNA fragments were phosphorylated by treatment with T4 polynucleotide kinase in order to generate 5' terminal phosphate groups. Then the two phosphorylated fragments were cloned separately into the plasmid pUC19 to create new recombinant plasmids pDUB2402 and pDUB2403 and transformed into JM83 (see section 2.19).

The screening for recombinant colonies was carried out on L-agar plates containing ampicillin, IPTG and X-gal. A few white colonies appeared from both sets of transformation. Mini preparations of plasmid DNA from all white colonies of both groups were made (see section 2.25) and the DNA restricted with EcoRI. The digested molecules of the two plasmids pDUB2402 and pDUB2403 were loaded onto a 0.7% agarose gel alongside an EcoRI *ilvY* fragment as a control to make sure that the two plasmids pDUB2402 and pDUB2403 carry the correct size *ilvY* gene. Unfortunately, none of the recombinant colonies from both experiments contained the proper size of *ilvY* gene fragment (see Fig.3.6A).

The second experiment involved a different strategy outlined in Fig.3.7. pDUB2401 and pDUB2400, that carry the wild type and constitutive *ilvY* genes respectively, were digested with HindIII plus BglIII. The digested samples were loaded onto a low melting point agarose gel, electrophoresed and the two fragments from each

Fig. (3.6).

A. 0.7% agarose gel electrophoresis of pDUB2402(lane 2 to 6) and pDUB2403(lane 7 to 12) cleaved with EcoRI to show that they have the correct size of reconstructed *ilvY* fragment. Lane 1. EcoRI digested pDUB2400. with the *ilvY* fragment to act as a size control.

B. Digestion of pDUB2400 and pDUB2401 with HindIII and BglII (lane 1 to 4 & 5 to 7) respectively then electrophoresed on a 0.7% agarose gel to separate the fragments.

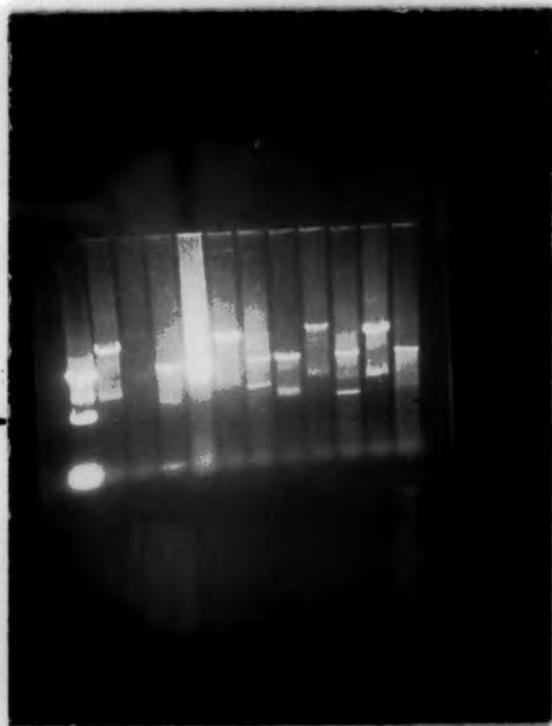
I refers to the large HindIII-BglII fragment.

II refers to the small HindIII-BglII fragment.

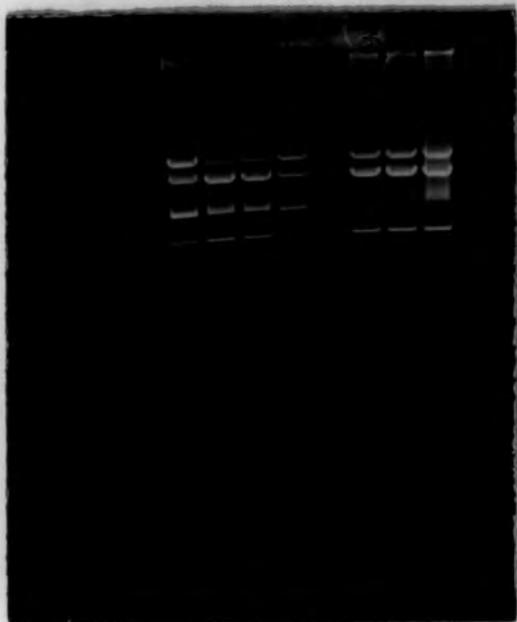
C. 0.7% agarose gel electrophoresis of pDUB2404 and pDUB2405 digested with EcoRI(lane 2 to 7 & 8 to 11 respectively), to examine the correct size of *ilvY*. All the *ilvY* inserts show an identical migration rate with a size control *ilvY* fragment (III) from pDUB2400(lane 1).

A 1 2 3 4 5 6 7 8 9 10 11 12

ilvY Fragment →



B 1 2 3 4 5 6 7



← I

← II

III →

C 1 2 3 4 5 6 7 8 9 10 11

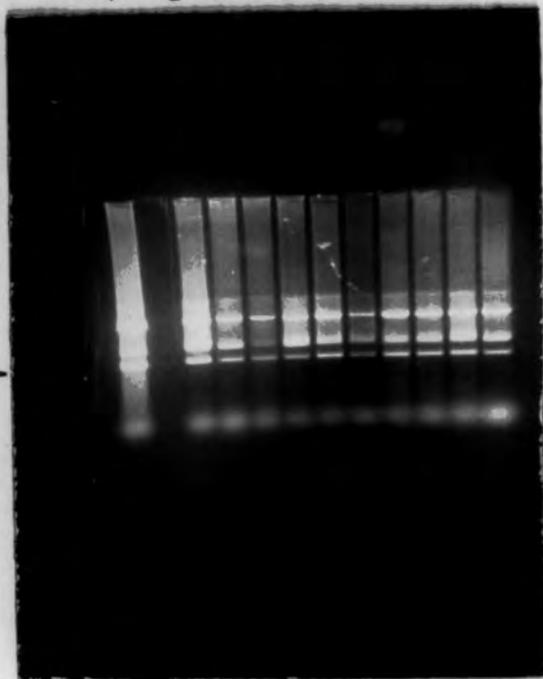
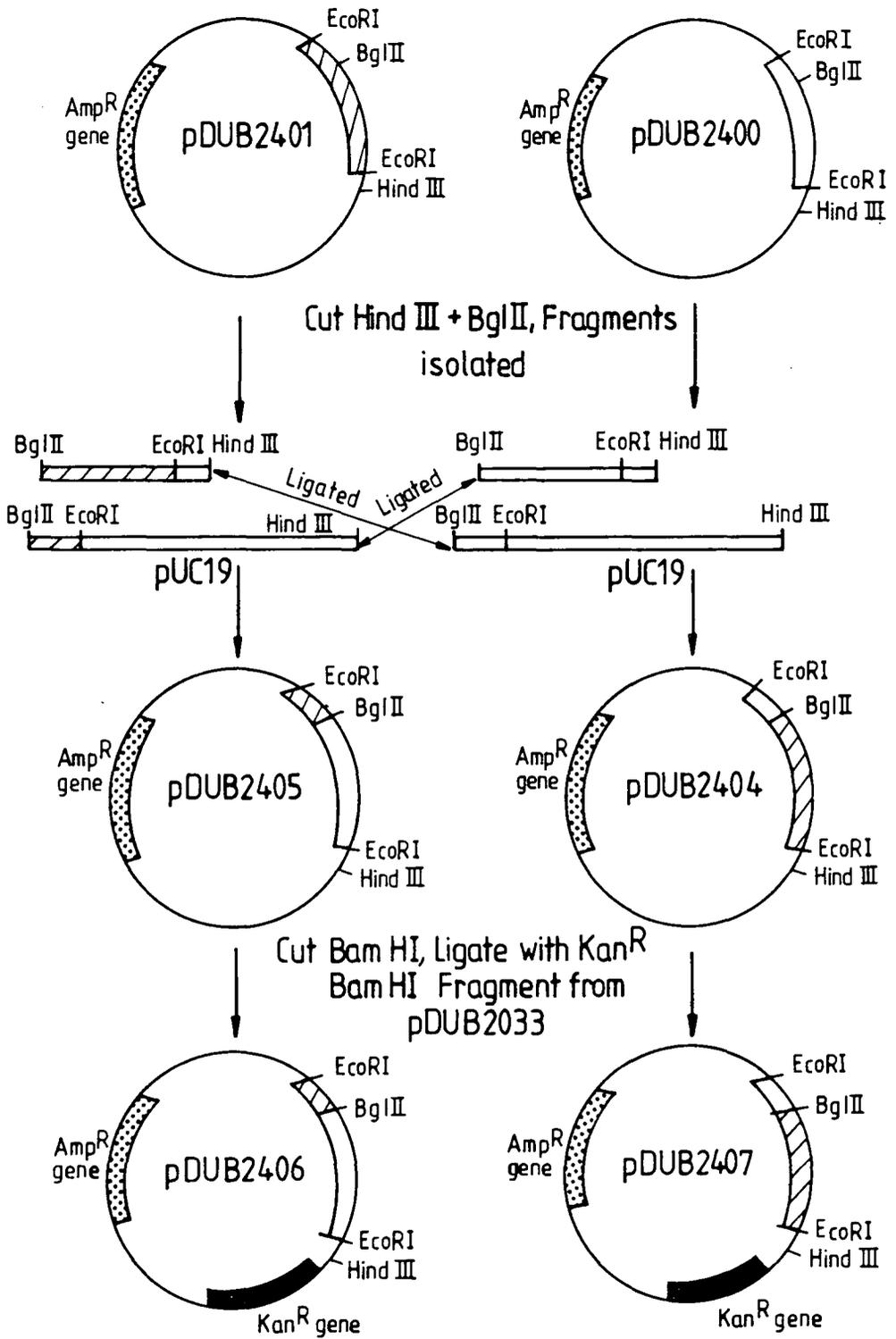


Fig. (3.7).

An alternate strategy for verifying that the small EcoRI-BglIII *ilvY* fragment, is solely responsible for the constitutive phenotype.



digestion purified (see section 2.16). The large fragment represents most of the vector and part of the *ilvY* gene (EcoRI-BglII). The smaller fragment represents the complementary half of the *ilvY* gene plus the remainder of the vector (see Fig.3.6B). The large and small fragments from each digestion were exchanged and ligated. This produces two plasmids containing hybrid *ilvY* genes. One contains the large EcoRI-BglII fragment from the wild type ligated to the small BglII-EcoRI fragment from the constitutive gene. The other is the complement of this.

The two ligation mixtures, representing pDUB2404 and pDUB2405, were transformed into JM83 (see section 2.19). Mini-preparations of plasmid DNA from the transformed colonies (see section 2.25) and digestion of the DNA with EcoRI followed. The digested molecules were loaded onto a 0.7% agarose gel alongside an EcoRI *ilvY* fragment serving as a size control. All of the digested DNA samples of pDUB2404 and pDUB2405 showed that they contained the proper reconstructed *ilvY* gene (see Fig.3.6C).

To proceed further a new selectable marker had to be added to the original and new recombinant plasmids. The four plasmids pDUB2400, pDUB2401, pDUB2404 and pDUB2405 have a single BamHI site in the pUC19 vector part. A BamHI fragment that encodes kanamycin resistance was purified from the plasmid pDUB2033 and ligated to the BamHI cut plasmids to create recombinant plasmids pDUB2409, pDUB2408, pDUB2407 and pDUB2406.

The next step involved the transformation of the new plasmids pDUB2406, pDUB2407, pDUB2408 and pDUB2409 into CU827($\Delta ilvDAYC$, Δlac) (see section 2.19), containing the following individual plasmids:-

pDUB2007 which is *RP4::ilvC::lacZ*

pDUB2012 which is *RP4::ilvC::lacZ* and *ilvY::Tn7*

pDUB2013 which is *RP4::ilvC::lacZ*, *ilvY*⁺, Tn7.

The above plasmids encode ampicillin and tetracycline resistance. Thus, the

transformed cells were plated on L-agar plates containing ampicillin , tetracycline and kanamycin. The transformed colonies obtained from each transformation set were purified on the above agar plates and then transferred to MacConkey agar containing lactose in addition to the above antibiotics to examine the inducibility of the *ilvC::lacZ* fusion. The results are displayed in Table 3.1.

From the Table 3.1, it is clear that *ilvC::lacZ* strains that contain pDUB2407 appear red when they are plated onto the lactose MacConkey agar plates. This colony colour is an indication of *ilvC* inducibility independent of any substrate due to the *ilvY^c* mutation carried on pDUB2407. The same colour phenotype was obtained for the strains transformed with pDUB2409. The strains that contain pDUB2406 and pDUB2408 show white colonies when they are plated on the same media.

3.6 Discussion.

It is clear from the autoradiograph (see Fig.3.1B) that the degree of hybridization to the labelled *ilvC* fragment is greater to the mRNA extracted from the whole-cell lysate of CU962 than to extracts from the other strains. This provides evidence that the *ilvY^c* allele, which exists in strain CU962, makes the expression of *ilvC* constitutive due to an increase in the amount of transcription. Strain CU152 was included as the wild-type control and shows almost no hybridization. Strain CU827 was included as a negative control as it is deleted for the *ilvDAYC* region. In fact the deletion starts within *ilvC* (Umbarger, H. E., personal communication) and probably fuses *ilvC* to the *ilvGMEDA* operon. Thus, message complementary to the probe is produced by this strain. A better control would have been CU838 which carries the *ilv-2049* deletion and lacks all of *ilvC*.

Among the three strains used in this experiment no clear hybridized band was detected from the RNA species extracted from strains grown in minimal

Table (3.1).

Trans-activation of *ilvC::lacZ* expression on MacConkey agar plates.

	pDUB2406	pDUB2407	pDUB2408 <i>ilvY</i> [†]	pDUB2409 <i>ilvY</i> ^c
1. pDUB2007 RP4:: <i>ilvC::lacZ</i> , <i>ilvY</i> [†]	White colonies	Red colonies	White colonies	Red colonies
2. pDUB2012 RP4:: <i>ilvC::lacZ</i> , <i>ilvY::Tn7</i>	White colonies	Red colonies	White colonies	Red colonies
3. pDUB2013 RP4:: <i>ilvC::lacZ</i> , <i>ilvY</i> [†] , Tn7.	White colonies	Red colonies	White colonies	Red colonies

media with limiting amounts of isoleucine, leucine. Thus the expected theoretical result of no hybridization to CU827; maximal hybridization to CU962 under all conditions and CU152 under derepressing conditions; and low hybridization to CU152 under repressing conditions was not seen. It is clear that there is higher levels of *ilvC* mRNA in CU962. Fig.3.1A indicates that the total amount of RNA isolated from each strain is roughly equivalent. There being slightly more RNA from growth in 2x YT media (lanes 3, 6 and 9). If the levels of RNA are roughly equivalent then the increased hybridization to CU962 RNA, surely represents an increase in the level of transcription of *ilvC*. However, this increase should also have been reflected in RNA extracted from CU962 grown in M9 minimal media. It is difficult to state with certainty that upsilon protein acts at the level of transcription based on this experiment alone. However, evidence presented later indicates that this is indeed the case.

An investigation was carried out to determine the precise nature of the *ilvY^c* mutation. Thus, both *ilvY⁺* and *ilvY^c* were sequenced. The single base change which is detected in the nucleotide sequence of *ilvY^c* results in the creation of a codon for a different amino acid from that which exists in the wild-type. However, the change is fairly conservative substituting a neutral to basic amino acid (glutamine) for a basic amino acid (lysine) (see Fig.3.4). The substitution of lysine for glutamine in the constitutive upsilon protein makes the protein independent of acetohydroxy acid isomeroreductase (*ilvC*) substrates (α -acetohydroxybutyrate and acetolactate) in the induction of *ilvC* expression. Thus, the expression of *ilvC* gene becomes constitutive.

An amino acid substitution does not always create a phenotypically detectable mutation, for instance, a hydrophobic cluster might be virtually unaffected by replacement of one leucine by another nonpolar amino acid such as isoleucine. Such a conservative change which has no effect on the phenotype is called a silent

mutation. Sometimes an amino acid substitution is only partially disruptive, for instance, the replacement with a more bulky amino acid such as phenylalanine might cause a reduction rather than a loss of activity of an enzyme. Such a mutation is called a leaky mutation. On the other hand, in other mutations the effect of amino acid substitution is very profound like that which occurs in haemoglobin molecules obtained from patients with sickle-cell anaemia, which differs from normal haemoglobin only by a single glutamine to valine change in the mutant form.

Two attempts were carried out to verify that the observed change of glutamine to lysine in *ilvY^c* is responsible for constitutive phenotype. The first attempt failed to obtain the correct size of hybrid *ilvY* gene cloned into pUC19. Theoretically, the strategy for the attempt was correct (see Fig.3.5). The *ilvY*²¹⁴³ constitutive mutation (*ilvY^c* allele) is trans-dominant affecting both chromosomes in a partial diploid cell. This is a diagnostic feature of positive control. In a cell that contains the *ilvC::lacZ* fusion and *ilvY^c* on the chromosome and *ilvC⁺* on *F'* plasmid, the expression of both enzymes, *ilvC*-directed β -galactosidase and isomeroreductase are elevated due to the action of the mutant gene product (Biel and Umbarger, 1980).

The second attempt used a different strategy and was successful (see Fig.3.7). Four derivative plasmids were produced, pDUB2406 (small *ilvY⁺*, large *ilvY^c*), pDUB2407 (small *ilvY^c*, large *ilvY⁺*), pDUB2408 (*ilvY⁺*) and pDUB2409 (*ilvY^c*). Small and large refer to the two BglIII-EcoRI sub-fragments of *ilvY*. These four plasmids were tested for trans-dominance with a low copy number plasmid carrying the *ilvC::lacZ* fusion. pDUB2407 and pDUB2409 showed a trans-dominant constitutive phenotype. This provides conclusive evidence that the single base change that occurs in the small EcoRI-BglIII *ilvY* fragment is totally responsible for making *ilvC* expression constitutive.

CHAPTER 4

TRANSCRIPTION OF *ilvY*

4.1 Introduction.

Two groups of workers concentrated their studies on determining the direction of transcription of the *ilvYC* system in different enteric bacteria using different techniques. The first group includes Wek and Hatfield (1986) who used *E. coli*. Their studies revealed that the two promoters of *ilvY* and *ilvC* are adjacent and transcribed in opposite directions. The other group (Blazey and Burns, 1984), used *S. typhimurium*. Their results were completely different from that of Wek and Hatfield. They stated that *ilvY* transcribes in the same direction as *ilvC*. This chapter provides evidence in support of the results of Wek and Hatfield.

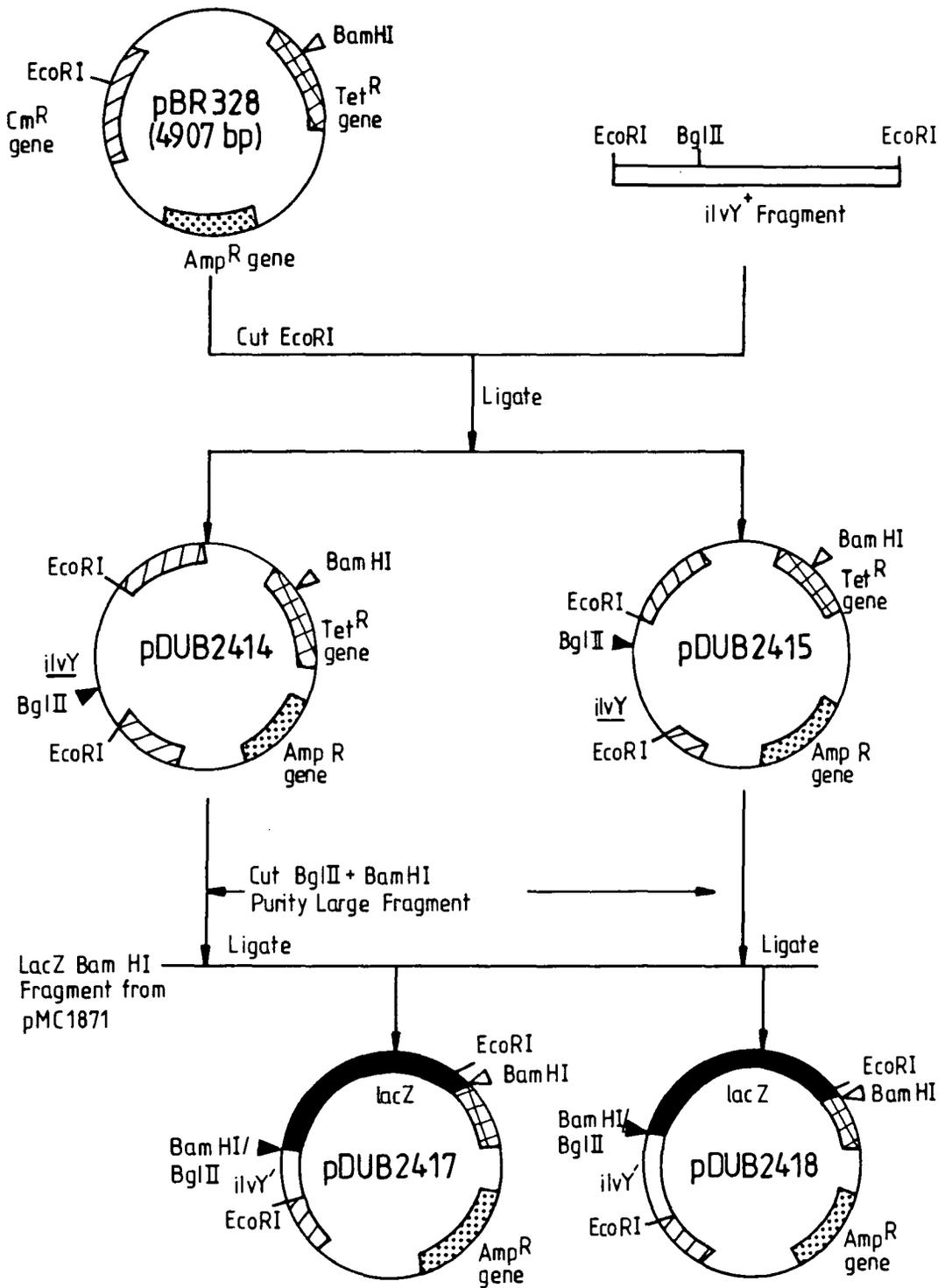
The direction of *ilvY* transcription was determined by two approaches. The first method involves using a novel technique which is currently widely used to study the regulation of gene expression. This technique involves fusing the *lacZ* gene, that encodes β -galactosidase, to *ilvY*, such that it is under the control of the *ilvY* promoter, and then observing the expression of the *lacZ* gene. The second approach is to express *ilvY* in a T7 expression system and characterizing the labelled products.

4.2 Fusion of *ilvY* with *lacZ* .

The direction of transcription of *ilvY* can be determined by making an *in vitro* fusion between the *ilvY* and *lacZ* genes, so that the later will be under the control of the *ilvY* promoter. The strategy for this approach is illustrated in Fig.4.1. Firstly, the plasmid pDUB2401, containing the wild-type *ilvY* gene was digested with EcoRI and the *ilvY* containing fragment purified after electrophoresis through low melting point agarose (see section 2.16). This fragment

Fig. (4.1).

Showing the steps in the construction of the fusion between *ilvY* fragments and *lacZ* for the purposes of determining the direction of *ilvY* transcription.



was ligated to EcoRI cut pBR328 and transformed into JM83. The transformed colonies were purified and then screened for resistance to chloramphenicol. Next, mini-preparations of plasmid DNA from those colonies which were sensitive to chloramphenicol were made (see section 2.25). The presence of the correct *ilvY* fragment was confirmed in the plasmids by restriction with EcoRI (see Fig.4.2A)

To examine the orientation of the *ilvY* gene cloned into pBR328 the plasmid DNA from twelve isolates was digested with BamHI and BglII. The digested molecules were loaded onto 0.7% low melting point agarose gel and electrophoresed. Two orientations for the *ilvY* gene are possible in pBR328 (see Fig.4.2B). These differ in the position of the *ilvY* BglII site with respect to the BamHI site of pBR328. The larger BglII-BamHI fragment of both orientations retains most of pBR328, including the ampicillin-resistance gene. This retains the large EcoRI-BglII *ilvY* fragment in pDUB2415 and the small EcoRI-BglII *ilvY* fragment in pDUB2414. The large BglII-BamHI fragments, from both pDUB2414 and pDUB2415 were purified from the low melting point agarose gel (see section 2.16).

The next step was to purify the *lacZ* fragment from the pMC1871 (see Fig.4.3). The plasmid was digested with BamHI and electrophoresed on a low melting point gel (see Fig.4.2C). The *lacZ* fragment migrates close to the vector fragment of pMC1871 so the gel was electrophoresed until the two close fragments separated. Next, the *lacZ* fragment was purified from the gel and ligated separately to the large BamHI-BglII fragments from pDUB2414 and pDUB2415. The ligated molecules were transformed into JM83. The cells were plated onto L-agar plates plus ampicillin.

Mini-preparations of plasmid DNA from at least 15 transformed colonies of each group were prepared (see section 2.25) and digested with EcoRI to examine the orientation of the *lacZ* gene which must be fused to the *ilvY* fragment in

Fig. (4.2).

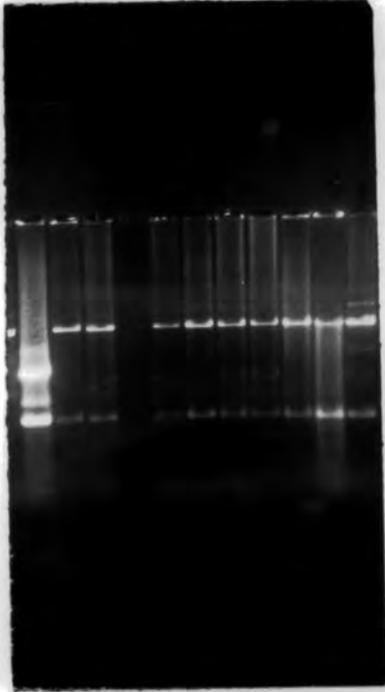
A. 0.7% agarose gel electrophoresis of plasmid DNA from transformed colonies sensitive to chloramphenicol cleaved with EcoRI to test the presence of cloned *ilvY* fragment (lane 2 to 10). Lane 1 is pDUB2400 digested with EcoRI to generate *ilvY* fragment acting as a size control.

B. Two orientations of cloned *ilvY* fragment with respect to BamHI site on the pBR328.

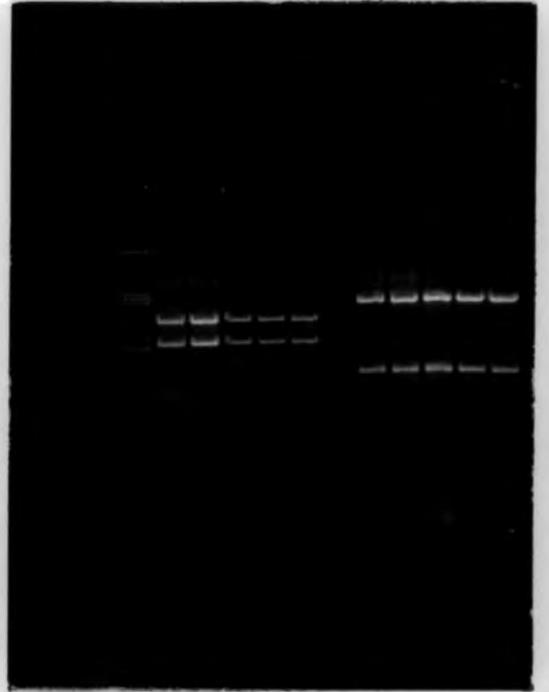
Lane 1 is λ DNA marker. Lanes 7 to 11 represent one orientation in which the large BamHI-BglII fragment carries the large *ilvY* sub-fragment (pDUB2415). Lanes 2 to 6 show the other orientation in which the large BamHI-BglII fragment carries the small *ilvY* sub-fragment (pDUB2414).

C. 0.7% agarose gel electrophoresis of digested pMC1871 with BamHI to generate the lacZ fragment.

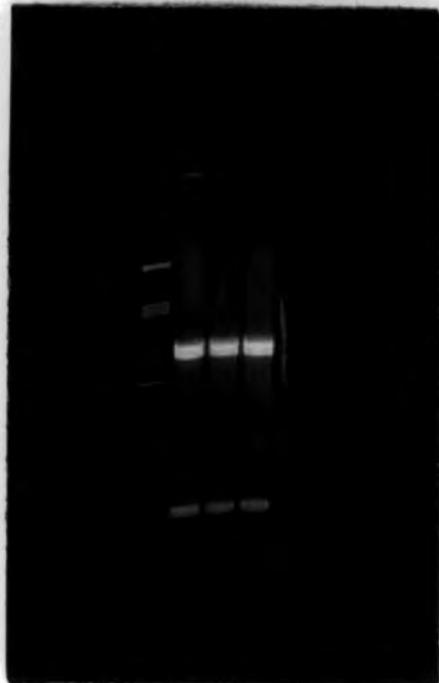
A 1 2 3 4 5 6 7 8 9 10



B 1 2 3 4 5 6 7 8 9 10 11



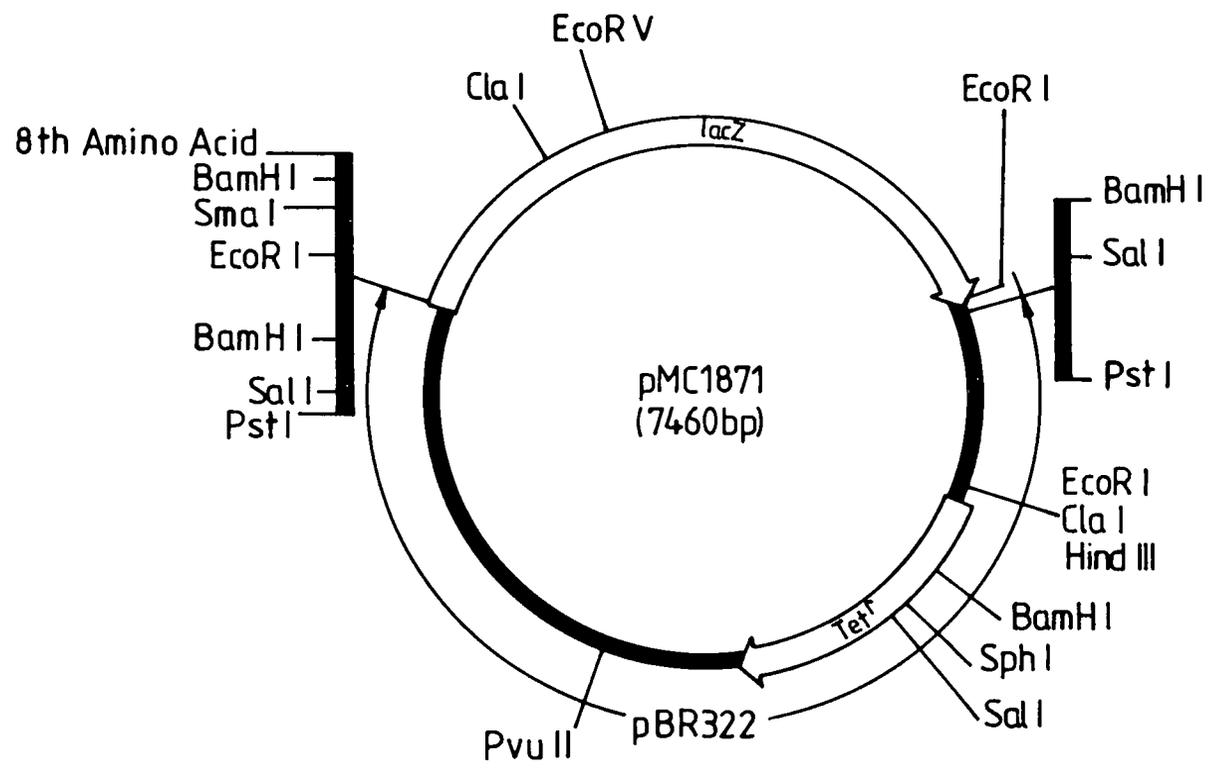
C 1 2 3 4



← lacZ Fragment

Fig. (4.3).

Restriction map of the fusion vector pMC1871.



	PstI	SalI	BamHI	EcoRI	SmaI
pBR322	CT GCA GGT CGA CGG ATC CGG GGA ATT CCC				
	8	9	1021		
	BamHI	pro val	lys stop stop stop		
GGG	GAT CCC GTC	lacZ	AAA TAA TAA TAA CCG GGC		
AGG GGG	GAT CCG TCG ACC TGC AG				pBR322
	BamHI	SalI	PstI		

the correct reading frame (see Fig.4.4A & 4.4B). The correct orientation creates plasmids with the *lacZ* EcoRI site adjacent to the BamHI site of pBR328 rather than adjacent to the BglII site of *ilvY* (see Fig.4.1).

The plasmid containing pBR328 with the small EcoRI-BglII fragment of *ilvY* fused in the correct orientation to *lacZ* was designated pDUB2417 whereas the plasmid containing pBR328 with large EcoRI-BglII fragment of *ilvY* fused in the correct orientation to the *lacZ* gene was designated pDUB2418. These two recombinant plasmids were transformed into the bacterial strain MC4100 (see section 2.19) and the cells were plated onto L-agar plates containing ampicillin and X-gal as an indicator for *lacZ* expression.

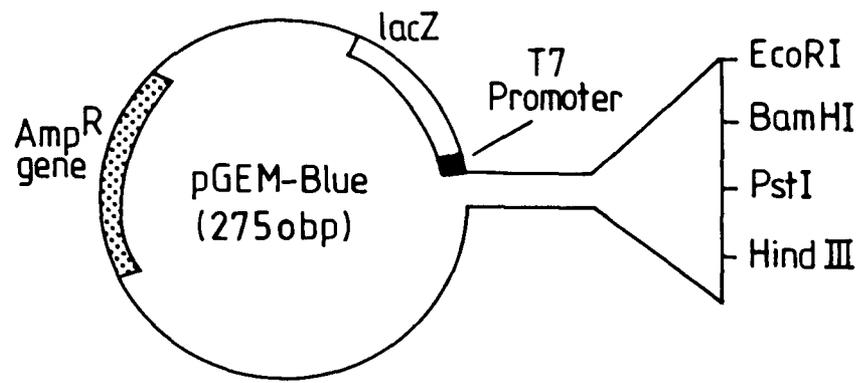
All the transformed colonies containing the plasmid pDUB2417 appear blue in colour while those containing pDUB2418 were white. Thus, the small EcoRI-BglII fragment of *ilvY* definitely contains a promoter region.

4.3 Expression of *ilvY* in the T7 system.

An alternative method to determine the direction of transcription is to express *ilvY* under the control of a foreign promoter, and to determine which orientation of *ilvY* with respect to the promoter gives the correct protein. The system chosen was the powerful T7 promoters, which respond only to T7 RNA polymerase. The *ilvY^c* gene was expressed in this system (see Fig.4.5).

The expression of the *ilvY^c* gene in the T7 system involved many steps. Firstly, the DNA fragment encoding the *ilvY^c* gene was purified from pDUB2400 after digestion with EcoRI and then ligated to pGEM-Blue adjacent to the T7 promoter. The ligated molecules were transformed into JM83 (see section 2.19) and plated on L-agar plates containing ampicillin, IPTG and X-gal. The white colonies were screened for the presence of the inserted *ilvY* fragment (see Fig.4.6A).

The cloned *ilvY* gene in pGEM-Blue could be in either of two orientations



↓ Cut EcoRI, Ligate EcoRI ilvY^C Fragment

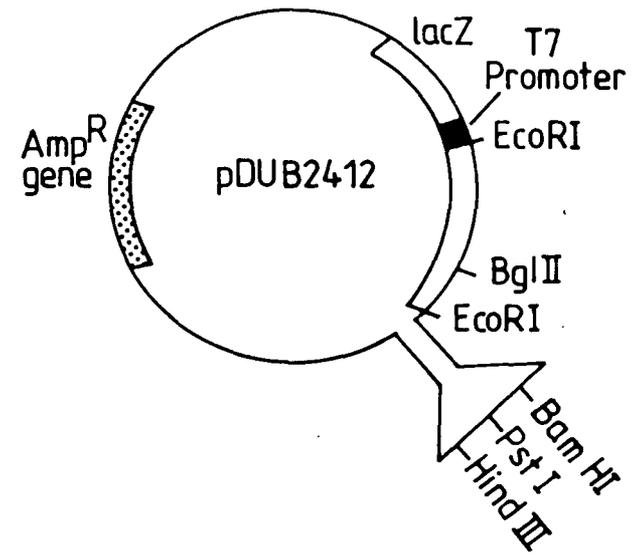
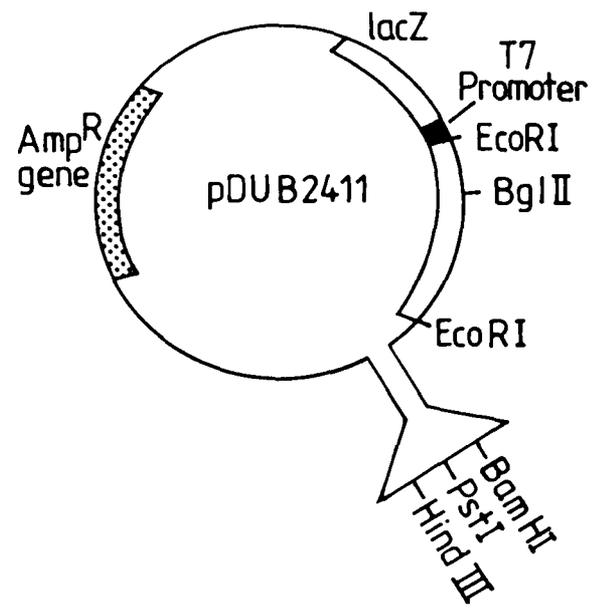
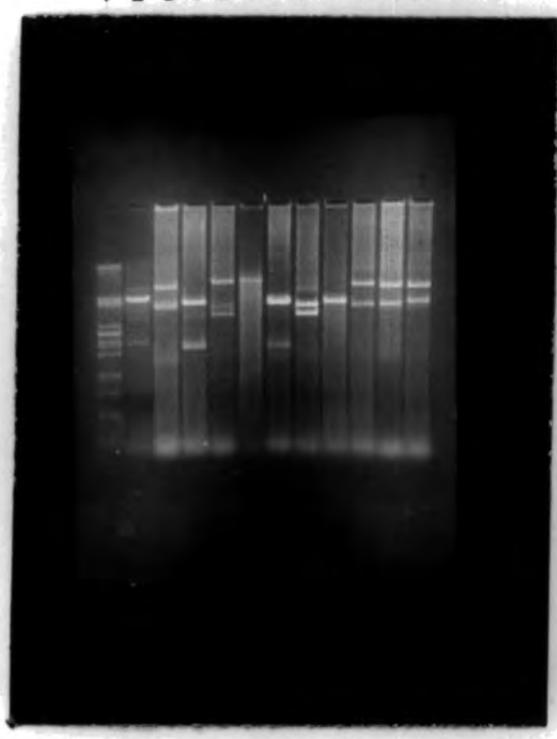


Fig. (4.5).

Illustrating the two possible orientations of *ilvY* in pGEM-Blue used in the T7 expression system to determine the direction of transcription of *ilvY*, and secondly to obtain high levels of expression of upsilon protein (see chapter 5).

A

1 2 3 4 5 6 7 8 9 10 11 12



B

1 2 3 4 5

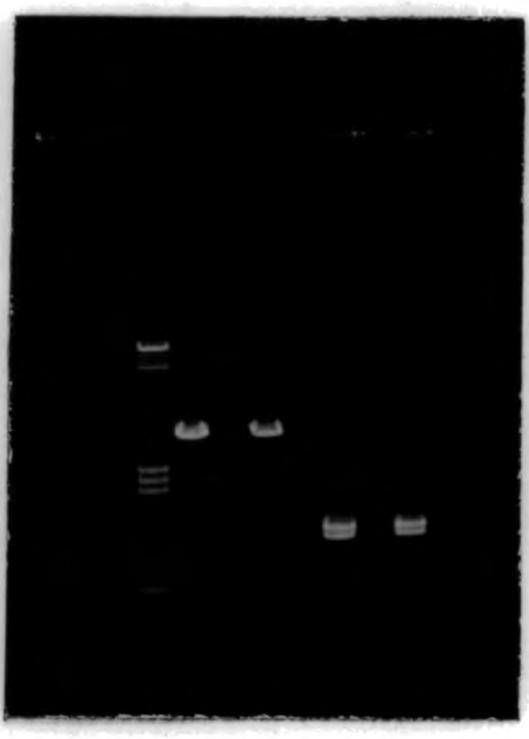


Fig. (4.4).

A. 0.7% Agarose gel showing the correct and incorrect fusion of the *lacZ* fragment to the large BamHI-BglII fragment of pDUB2415 after restriction with EcoRI.

Lane 1 is λ DNA marker. Lane 8 represents the correct fusion of the *lacZ* fragment to the large fragment of *ilvY*. Other lanes represent the wrong fusion of *lacZ* fragment to the large fragment of *ilvY*.

B. BamHI *lacZ* fragment was fused to the large BamHI-BglII fragment of pDUB2414 then restricted with EcoRI then loaded on a 0.7% agarose gel.

Lane 1 is λ DNA marker. Lanes 4 & 5 show the correct orientation of *lacZ* to the small fragment of *ilvY*. Lanes 2 & 3 represent the wrong orientation of *lacZ* to the small fragment of *ilvY*.

Fig. (4.6).

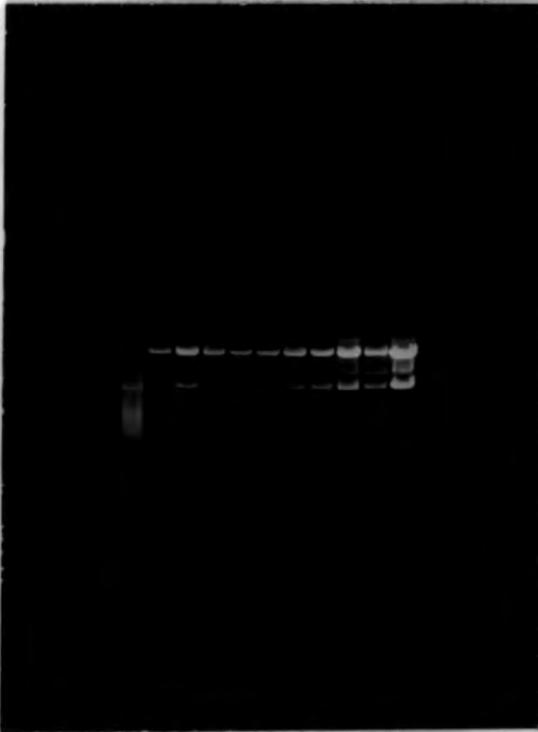
A. Screening of transformed colonies for the correct *ilvY* fragment in pGEM-Blue. The plasmid DNA from these colonies was digested with EcoRI and loaded onto a 0.7% agarose gel.

Lane 1 is *ilvY* fragment from pDUB2400. Lanes 2 to 11 show matching of cloned *ilvY* fragments with the size control.

B. Two orientations of cloned *ilvY* fragment in pGEM-Blue which differ in the position of the BglII site with respect to the T7 promoter. Lanes 2, 3, 6 & 7 represent the orientation in which the BglII site of the *ilvY* fragment is near the T7 promoter. Lanes 1, 4, 5 & 8 is the other orientation in which the BglII site is distal to the T7 promoter.

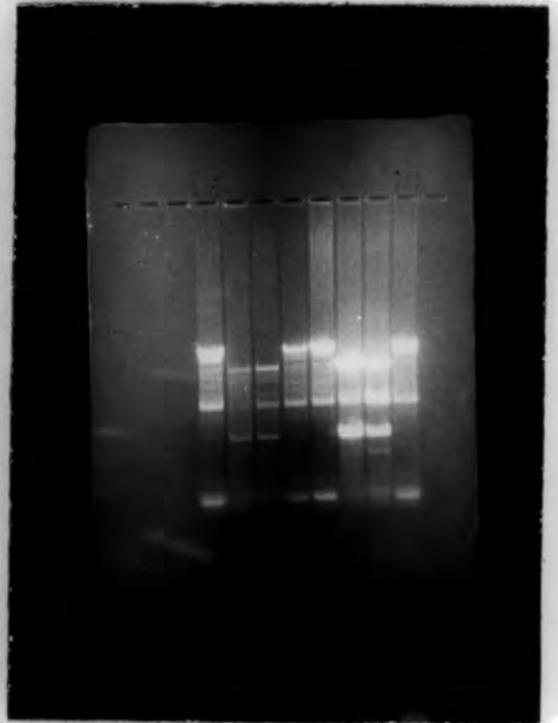
A

1 2 3 4 5 6 7 8 9 10 11



B

1 2 3 4 5 6 7 8



with respect to the T7 promoter region (see Fig.4.5). In the first orientation the BglII site of *ilvY* is near the T7 promoter and this orientation was designated as pDUB2411 (see Fig.4.5). Digestion of this plasmid with BamHI and BglII generates two DNA fragments, one representing the pGEM-Blue plasmid carrying the small fragment of *ilvY^c* while the other one is the large fragment of *ilvY^c* (see Fig.4.6B). In second orientation the BglII site is distal from the T7 promoter (see Fig.4.5). This orientation was designated as pDUB2412. Thus, on digestion of this plasmid with BglII and BamHI, two DNA fragments, one represented pGEM-Blue carrying the large fragment of *ilvY* and the second one is the small fragment of the gene (see Fig.4.6B). The two orientations were used in a T7 expression system to investigate in which direction *ilvY* is transcribed.

The two plasmids pDUB2411 and pDUB2412 were transformed separately into the strain C600 *recA⁻* which also carries pGP1-2 (see section 2.19). The cells were plated onto L-agar plates containing ampicillin, kanamycin and IPTG and incubated at 30°C for 48 hours. The transformant colonies were purified on the same L-agar plates. Thus, two strains were obtained, each has the plasmid pGP1-2 and one orientation of the cloned *ilvY* gene in pGEM-Blue. These strains were named as strain A (contains pDUB2411 and pGP1-2 plasmids) and strain B (contains pDUB2412 and pGP1-2 plasmids).

Two cultures of strains A and B were inoculated into L-broth plus ampicillin, kanamycin and IPTG and then incubated overnight at 30°C. Then, the two cultures were heat shocked at 42°C for 30 minutes, followed, by incubation at 37°C for 2-3 hours.

To determine in which direction *ilvY* is transcribed, crude extracts were prepared from each strain. This was performed by centrifuging the two cultures and resuspend the pellets in 500 μ l TE buffer, followed by sonication on ice for 6 x 30 seconds. The extract was then centrifuged to clarify the solution and remove the

cell debris. An equal volume of 2x sample buffer was added and 50 μ l from each sample was loaded onto a 10% SDS-polyacrylamide gel (see section 2.13) alongside molecular weight protein markers and then electrophoresed overnight at 40 volts. The gel was then stained with Coomassie blue, and then destained and dried.

Many protein bands appeared in the tracks from both strains and the precise location of the protein band encoded by the *ilvY* gene could not be identified (see Fig.4.7).

4.4 Preparation of radiolabelled extracts from strains A and B.

Preparation of radiolabeled extracts from strains A and B was carried out as follows:- Two 10 ml cultures of strains A and B were grown at 30°C for two complete days in M9 minimal media containing, 0.5% glucose, 10 μ g/ml thiamine, 20 μ g/ml thymine, 20 μ g/ml threonine, 20 μ g/ml leucine, IPTG, ampicillin and kanamycin. The cells were then treated by heat shock at 42°C for 30 minutes in a water bath. 1 ml from each strain was taken and 200 μ g/ml rifampicin was added to prevent *E. coli* RNA polymerase transcription of the host. Incubation was continued for a further 10 minutes at 42°C then the two samples were transferred to 37°C for 20 minutes. Next, 20 μ Ci of S³⁵-methionine was added to each sample (specific activity 1444 Ci/mmol) and incubated for a further 20 minutes at 37°C.

Crude extracts were prepared from each radioactive sample as previously described and 50 μ l from each sample was loaded onto 10% SDS-PAGE (see section 2.13) alongside molecular weight protein markers. After electrophoresis the SDS gel was stained with Coomassie stain and then destained to reveal the bands of the protein markers. The gel was then dried and exposed to X-ray film for autoradiography.

Two protein bands appeared in the gel from the extract of strain A, one of them is in the molecular weight range of 35000 daltons while the other is smaller

of molecular weight 30000 daltons (see Fig.4.8). The 35000 daltons protein band represents upsilon protein and the other band is β -lactamase protein encoded by the ampicillin resistance gene.

4.5 Discussion

The *ilvYC* genes in *E. coli* K-12 constitute a distinctive regulatory system that employs positive control, which is uncommon for biosynthetic genes of amino acid operons. In this work the location of the *ilvY* promoter was located in the small EcoRI-BglII fragment by using an *in vitro* fusion technique. Fusions vectors were constructed between the two EcoRI-BglII fragments of *ilvY* and a *lacZ* fragment which lacks its own promoter and translation initiation site. The small EcoRI-BglII is proposed by Wek and Hatfield to contain the *ilvY* promoter. Their designated reading frame of *ilvY* means that, if correct, use of the *ilvY* BglII site will allow the N-terminal end of *ilvY* to fuse correctly in frame with *lacZ*. The large fragment, if it contains the promoter, has a one in three chance of being fused in frame. After transforming these fusions into the bacterial strain MC4100 which is deleted for *lacZ*, cells were plated onto L-agar plus X-gal.

The expression of the *lacZ* gene only occurs when MC4100 cells contain pDUB2417 in which the small DNA fragment of the *ilvY* gene was fused to the *lacZ* fragment. This direction of transcription is opposite to that of *ilvC*. In addition these results imply that the fusion has occurred in the correct reading frame to the N-terminal end of *ilvY* protein. On the contrary, the appearance of white colonies on L-agar plates plus X-gal after transforming MC4100 with pDUB2418 indicates an absence of a promoter sequence in the large DNA fragment of the *ilvY* and / or out of frame fusion to a protein coding sequence.

This evidence reported here strongly supports Wek and Hatfield (1986) which states that *ilvY* and *ilvC* are transcribed in opposite directions. Furthermore the

Fig. (4.7).

10% SDS-PAGE gel of crude extracts from strains A and B after expressing the *ilvY* gene in the T7 system.

Lane 1, 2 & 3 crude extracts from strain A.

Lane 4, 5 & 6 crude extracts from strain B.

Lane 7 is standard protein markers.

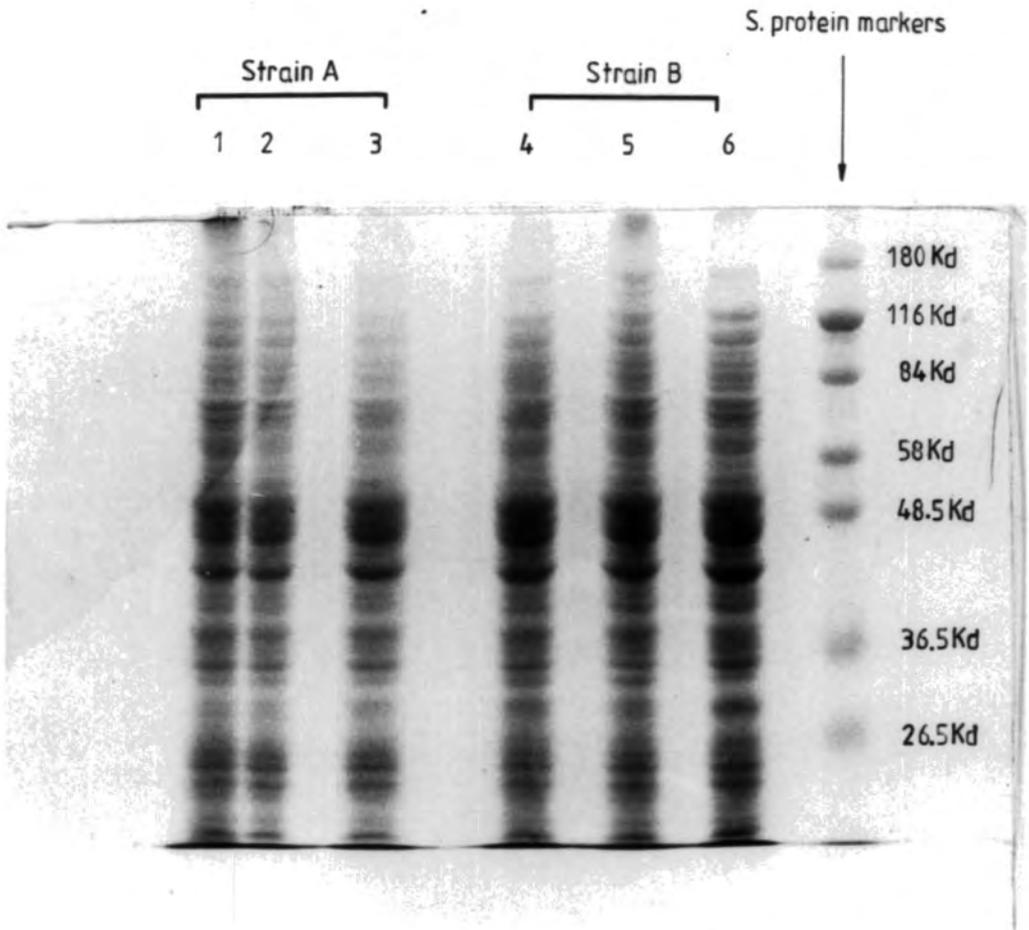
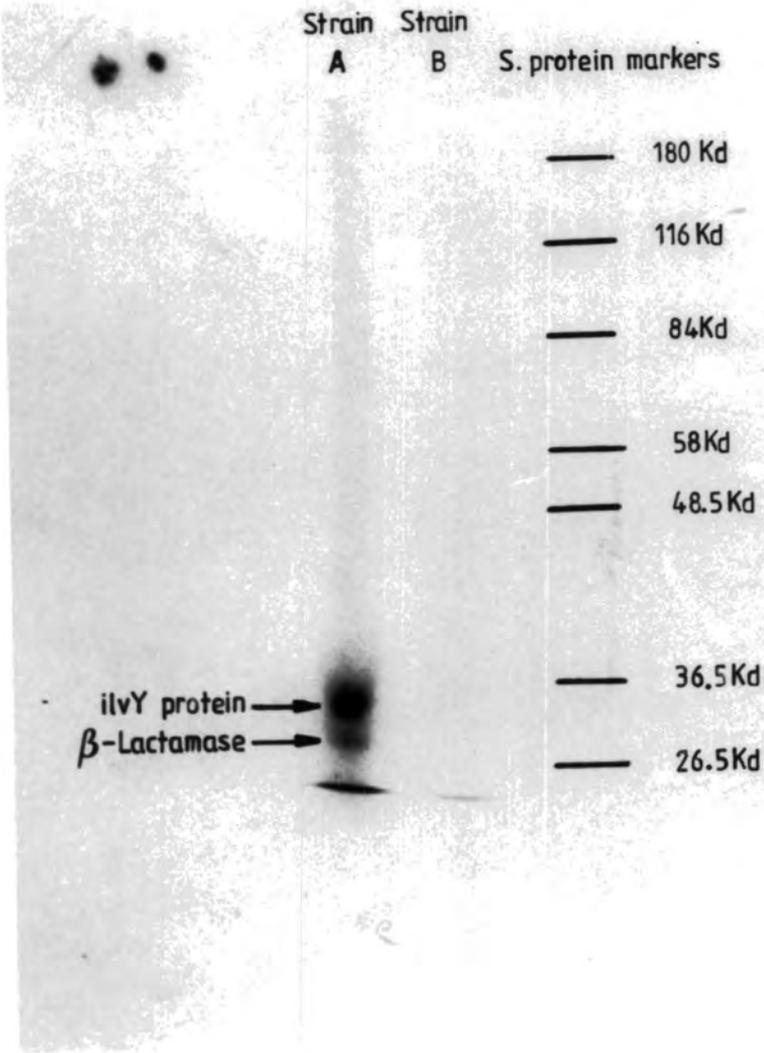


Fig. (4.8).

A. Results of electrophoresis of labelled proteins using 10% SDS-PAGE gel.

Lane 1 showing two labelled protein bands from strain A, one related to *ilvY* protein (upsilon 35 kd) while the second one is β -lactamase (30 kd).

Lane 2 no labelled protein bands have been detected from strain B.



open reading frame predicted from their DNA sequence, proposed to be the *ilvY* coding sequence, is definitely a translated sequence. Wek and Hatfield used a different strategy to verify their predictions. They determined the *in vivo* transcriptional initiation sites of *ilvY* and *ilvC* using S1 nuclease mapping. This technique involves the extraction of RNA from strains that contain plasmids carrying *ilvA*, *ilvC* and *ilvY* fragments. The extracted mRNA was hybridized separately to different labelled fragments of the *ilvYC* system, BglIII-EcoRI (-48 to 170 *ilvY*) and HincII-PvuII (-75 to 95 *ilvC*). The hybridized bands were digested with S1 nuclease and analysed on polyacrylamide/urea gels in parallel with sequence reactions of the EcoRI-BglIII and HincII-PvuII DNA fragments. They found that the two transcriptional start sites for the genes are 45 nucleotides apart therefore they concluded that the two genes are transcribed in the opposite direction.

Blazey and Burns (1980) studied the *ilvYC* genes in another enteric bacterium *S. typhimurium*. Their studies included the structural organization of the two genes and their regulation. They stated that the *ilvY* gene product mediated the induction of *ilvC* expression in a functionally similar manner between *E. coli* and *S. typhimurium*. Then in 1984 they concentrated their studies on the regulation of these two genes in *S. typhimurium*. They used a *galK* fusion vector in order to determine the direction of transcription of *ilvY* and *ilvC* by fusing separately *in vitro* DNA fragments that contain the genes into this vector. After transformation the enzyme galactokinase was assayed. They concluded that the ability to promote *galK* expression occurs only when the XhoI-SalI DNA fragment (encoding *ilvY*) was digested by BAL-31 at the SalI terminus. Thus, the *ilvY* promoter must be located at the XhoI terminus and therefore the *ilvY* gene is transcribed in the same direction as *ilvC*.

Both sets of experiments described by Wek and Hatfield (1986) and Blazey and Burns (1984) are open to criticism. S1 mapping of mRNA purified from

prokaryotes is extremely difficult, due to the difficulties in purifying intact mRNA. The fusion studies of Blazey and Burns, produced no supporting evidence.

Additional evidence on the direction of transcription was added in this work through the efforts to radiolabel the upsilon protein after expression in the T7 system. In the strains A and B the introduction of the *ilvY^c* fragment in both orientations under the control of a T7 promoter. A gene for active T7 RNA polymerase under the control of the λP_L promoter is carried on pGP1-2. The *λcI857* gene is under the control of the *lac* promoter on pGP1-2. This promoter is induced by inclusion of IPTG in the growth medium, thus allowing expression of the λ repressor. At 30°C the heat sensitive λ repressor, does not repress completely the initiation of transcription at P_L so a terminator for *E. coli* RNA polymerase was inserted into the pGP1-2, so that the expression of T7 RNA polymerase from pGP1-2 is dependent on transcriptional readthrough (Tabor and Richardson, 1985). T7 RNA polymerase was expressed in both strains A and B when the cultures were incubated at 42°C. This will inactivate the heat sensitive λ repressor which no longer binds to the P_L promoter of the T7 RNA polymerase gene. The relatively short time for induction is sufficient because even a low level of T7 RNA polymerase is enough for initiation of transcription at the T7 promoter on pGEM-Blue. Transcription initiated at the T7 promoter reads through the cloned *ilvY* fragment. The expression of upsilon protein occurs when the gene is cloned in the correct orientation with respect to the T7 promoter.

Next, labeling of the crude cell extracts from both strains was carried out by S³⁵-methionine after inactivation of the host RNA polymerase with rifampicin (Chamberlin and Ring, 1973). Since *E. coli* mRNA decays rapidly, all new mRNA in the cell is produced by T7 RNA polymerase.

From the autoradiograph in (Fig.4.8), it is clear that in strain A a protein band of molecular weight approximately 30000 was expressed in addition to the

upsilon protein band. This is probably β -lactamase which is encoded by the ampicillin resistance gene. Neither protein was identified in strain B. *In vitro* studies revealed that the transcription by T7 RNA polymerase from a T7 RNA promoter on a plasmid results in transcripts several times the plasmid length (McAllister *et al.*, 1981). T7 RNA polymerase is very efficient in transcribing the *ilvY* gene linked next to the T7 promoter in the correct orientation and a transcription termination signal seems to be absent. T7 RNA polymerase will circumscribe pGEM-Blue several times without terminating and thus transcribe the gene for ampicillin resistance as well.

ilvY is transcribed and translated first. The attachment of ribosomes at *ilvY* may allow the efficient translation of *bla* as well, as in polycistronic mRNA. When *ilvY* is inserted in the incorrect orientation no translation occurs as there is no sensible reading frame. Ribosomes may find it difficult to initiate translation of *bla*, which is distal from the 5' end of the message (analogous to polar mutations) thus accounting for the lack of the 30000 protein band in strain B. Alternatively, this non-expressing orientation of *ilvY* may fortuitously code for an efficient T7 RNA polymerase transcription terminator.

The efficient radiolabeling of the upsilon protein has shown that the T7 RNA polymerase/promoter system is a successful alternative procedure for labeling plasmid encoded proteins over the mini (Dougan and Sherratt, 1977) or maxicell (Sancar *et al.*, 1981) procedures.

The two pieces of evidence described in this work definitely show that the location of the *ilvY* promoter is located within the small EcoRI-BglII fragment and that *ilvY* transcription occurs in a direction opposite to that of *ilvC*, as originally proposed by Wek and Hatfield (1986).

CHAPTER 5

Purification and characterization of epsilon protein**5.1 Introduction.**

This chapter describes the identification and purification of the *E. coli* DNA binding protein, epsilon. Preparation of radioactive epsilon protein was accomplished by expression in a T7 based system. Additionally, the native molecular weight of the protein was determined by gel filtration.

The *in vitro* activities of epsilon protein were examined at the transcriptional level and in a cell-free coupled transcription translation system.

5.2 Purification of epsilon protein.

pDUB2400 (*ilvY^c*) was transformed into CU827, which is deleted for *ilvDAYC*, to create strain C which was used as a source of epsilon protein. Strain CU827 and CU176 (*ilvY⁺*) were used as controls. 2 litre cultures of CU827, CU176 and C were prepared separately in 2x L-broth. The cells were harvested by centrifugation at 8000 rpm and the pellets resuspended in 10 ml TE buffer. The resuspended pellets were sonicated with a medium probe at a pulse rating of 22 for 6 times 30 seconds on ice. The lysed cells were clarified by centrifugation at 15000 rpm for 20 minutes at 4°C.

The supernatants were retained and made to 1 M NaCl concentration. Then, streptomycin sulphate was added at a final concentration of 1% and stirred at 4°C overnight to remove the nucleic acids. The precipitate was removed by centrifugation for 45 minutes at 15000 rpm at 4°C.

Then, fractionation of proteins was carried out using ammonium sulphate. Various fractions were obtained at 30%, 60% and 90% saturation for each strain. To do this the supernatant volume for each strain was determined and the amount

of ammonium sulphate to be added to obtain the various fractions was calculated from Table 5.1. $(\text{NH}_4)_2\text{SO}_4$ was added slowly to the supernatants at room temperature with stirring followed by centrifugation for at least 20 minutes at 8000 rpm and 4°C. The pellets obtained were resuspended in 10 ml TE buffer.

The three ammonium sulphate fractions of the three strains were dialyzed against TE buffer. The dialysis was carried out for 48 hours with 6 to 7 changes at 4°C to remove the ammonium sulphate from the protein fractions.

5.3 Gel retardation assays of the three ammonium sulphate fractions.

Gel retardation assays were carried out to detect DNA binding activity in the three ammonium sulphate fractions. Initially, unlabelled PvuII-BglIII DNA, which was purified from pGMM201 (see Fig.3.2 & 5.2B), was used as a substrate. This encodes the overlapping region that contains the postulated promoters for *ilvY* and *ilvC*. The binding reaction was performed as follows:- To 8 eppendorff tubes, each containing 1 μg of PvuII-BglIII fragment, 1, 2, 3, 4, 5, 6, 7 and 8 μg of proteins from the 60% ammonium sulphate fraction of strain C, 5 μl of 10x binding buffer and 2 μl of Salmon sperm DNA (5 mg/ml), was added. The reaction samples were incubated at 37°C for 20 minutes and the reactions were terminated by adding stop dye. Then, the reaction samples were applied to a 1.5% agarose gel alongside untreated PvuII-BglIII fragment acting as a control.

After electrophoresis in TBE buffer the agarose gel was stained in a solution containing 0.5 $\mu\text{g}/\text{ml}$ ethidium bromide and visualized on a UV transilluminator.

The retardation of the unlabelled DNA fragment gives a good indication of the presence of DNA binding activity in the 60% ammonium sulphate fraction (see Fig.5.1A).

The above experiment was repeated, but this time proteins from the 30% and 90% ammonium sulphate fractions of strain C were used with the same amounts

Table 5.1

Percentage saturation of ammonium sulphate.

		Final concentration of ammonium sulphate, % Saturation							
		30	40	50	60	70	80	90	100
		Grams solid ammonium sulphate to be added to 1 litre of solution							
Initial concentration of ammonium sulphate	0	176	243	313	390	472	561	662	767
	30				198	273	356	449	546
% saturation	60					69	143	227	314

Fig (5.1).

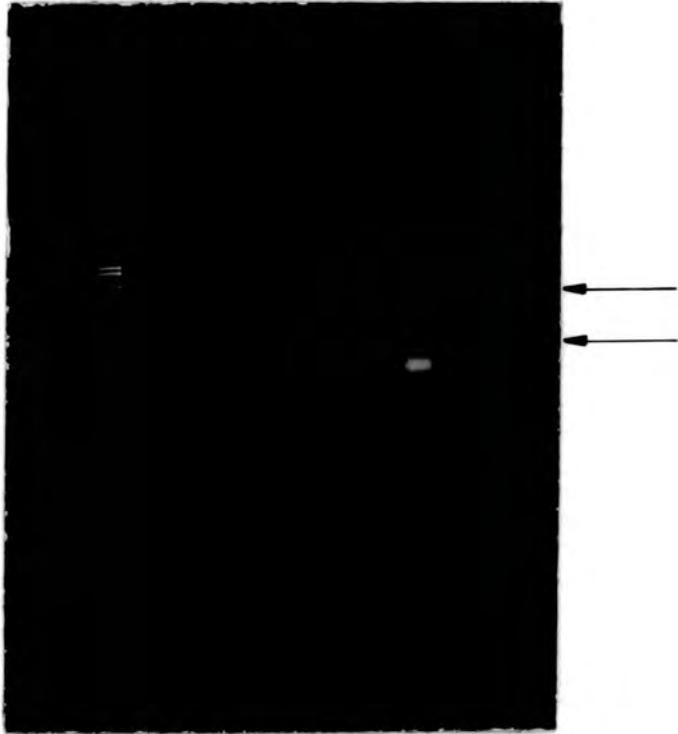
A. Retardation of the unlabelled *ilvYC* PvuII-BglIII fragment with the 60% ammonium sulphate fraction of strain C. Different amounts of proteins (8, 7, 6, 5, 4, 3, 2, and 1 μg , lanes 2-10). Lane 1 is λ DNA marker. Lane 2 to 9 shows retardation. Lane 10 no retardation occurs. Lane 11 is unlabelled fragment without protein.

B. No retardation could be detected from binding of unlabelled PvuII-BglIII fragment to proteins from 30% and 90% ammonium sulphate fractions.

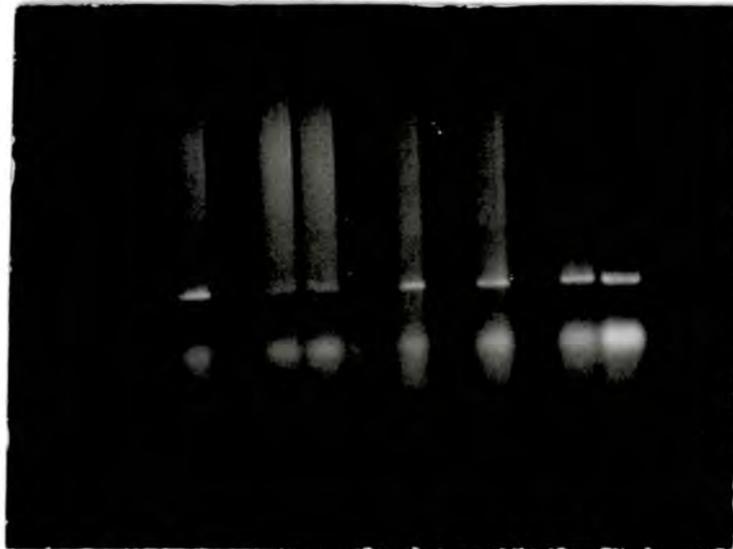
Lane 1 to 3 proteins from 30% fraction. Lane 4 to 6 proteins from 90% fraction.

Lane 7 is unlabelled PvuII-BglIII acting as a size control.

A 1 2 3 4 5 6 7 8 9 10 11



B 1 2 3 4 5 6 7



of protein as used above. No retardation was detected in either of the fractions (see Fig.5.1B).

5.4 *In vitro* transcription of *ilvC*.

To examine the activation of *ilvC* transcription by the three ammonium sulphate fractions from strain C, the following quantitative assay was carried out. This assay depends on the fact that RNA molecules synthesized *in vitro* are insoluble in trichloroacetic acid. Thus, in a reaction mixture containing double stranded DNA as a template, four nucleotide precursors (ATP, GTP, CTP and UTP) one of which was labeled and RNA polymerase, mRNA molecules are synthesized *in vitro*.

The *ilvC* PstI-BglIII DNA fragment was purified from pGMM201 using (see Fig.5.2C) using low melting point agarose (see section 2.16). The transcription reaction was carried out as follows:- 20 μ l PstI-BglIII DNA (1 μ g), 10 μ l 10x transcription buffer, 10 μ l 1 M KCl, 10 μ l of nucleotide mixture, 5 μ l ^3H -uridine (specific activity 40 Ci/mmol), 2 μ g of proteins from each ammonium sulphate fraction and 1 μ l RNA polymerase (1 unit/ μ l) in a total volume of 60 μ l. In the control sample sterile water was added instead of proteins. The samples were incubated at 37°C for 30 minutes. Then 1.5 ml of 10% trichloroacetic acid was added and kept for 10 minutes on ice. The *in vitro* synthesized RNA was precipitated on a wet GF/C filter. The filter was then washed with 10 ml 10% trichloroacetic acid followed by 5 ml absolute ethanol. After drying, the incorporation of labelled nucleotides into *in vitro* synthesized mRNA was measured by liquid scintillation counting.

The results obtained with the *in vitro* transcription of *ilvC*, show that maximum activity was obtained with proteins from the 30% ammonium sulphate fraction, whereas inhibition of transcription occurs when the 60% ammonium sul-

phate fraction was used (see Table 5.2). This activation was also observed when fractions eluted from a DEAE-cellulose column after application of the 30% ammonium sulphate fraction were assayed (see section 5.6 for method & Table 5.3).

5.5 Filter binding assays of epsilon protein.

Filter binding assays were also carried out as an additional procedure for the identification of epsilon protein. To perform this assay the PstI-BglII DNA fragment purified from pGMM201 was labelled *in vitro* by nick translation (see section 2.21) to serve as a substrate in the assay. The binding reactions were performed as follows:- To 9 tubes each containing 20 μ l of labelled DNA fragment (specific activity 10×10^6 CPM/ μ g), 30 μ l 10x binding buffer, 2 μ l of Salmon sperm DNA (5 mg/ml) and separately 2 μ g of proteins from each ammonium sulphate fraction of the strains CU827, CU176 and C. The reaction mixtures were incubated for 10 minutes at room temperature then applied to wet nitrocellulose filters, presoaked in 1x binding buffer for 30 minutes. The filters were washed with 1 ml of 1x binding buffer, then dried and the radiolabelled DNA trapped on the filters measured by liquid scintillation counting.

The filter binding assay, with the 3 ammonium sulphate fractions of strains CU827, CU176 and C confirm the result obtained with the gel retardation assay (see section 5.3). These measurements show increased retention of labelled DNA with proteins from the 60% ammonium sulphate fraction of strain C (see Table 5.4).

5.6 Purification by DEAE cellulose column chromatography.

DEAE (diethylaminoethyl) is a chemically modified cellulose introduced by Peterson and Sober (1956). It is an anion exchanger with residues linked to the hydroxyl group of the cellulose.

Table (5.2).

In vitro transcription of *ilvC* using the three ammonium sulphate fractions of strain C.

Sample number	Fraction	C.P.M.
1.	Control sample	14416
2.	30% fraction	28039
3.	60% fraction	2542
4.	90% fraction	19799

Table (5.3).

In vitro transcription of ilvC using fractions eluted from a DEAE-cellulose column after application of the 30% ammonium sulphate fraction of strain C.

Fraction	C.P.M.
1	5156
2	7172
3	7443
4	8350
5	13669
6	51755
7	82543
8	8660
9	8494
Control sample	7099

Table (5.4).

Filter-binding assay of ammonium sulphate fractions of wild-type, deleted and constitutive strains.

Strain	Ammonium sulphate fraction	Readings in CPM
CU827	30%	3105
	60%	3471
	90%	5361
CU176	30%	4419
	60%	7619
	90%	1977
Strain C	30%	8232
	60%	12154
	90%	7873

The basic idea for protein separation through DEAE cellulose columns is the formation of multiple electrostatic bonds between the charged sites on the surface of the exchanger and sites of opposite charge on the surface of the proteins. In addition, non ionic interactions are also formed between the hydroxyl groups of the cellulose and the polar residues on the surface of the protein which are mainly hydrogen bonds (Tomlinson and Tener, 1963).

DEAE-cellulose chromatography represents the second stage for upsilon protein purification. A sufficient amount of DE52 (wet type) was mixed with 20 ml TE buffer then the mixture was loaded into the column (16x1cm). After packing the column, it was equilibrated with 20 ml TE buffer containing 1.5 M NaCl then by 20 ml TE buffer. After equilibration, the ammonium sulphate fractions were applied separately to the column. The unbound proteins were washed through with 20 ml TE buffer.

In order to elute bound proteins from the column a stepwise gradient of 10 ml steps was used. The following buffers of fixed ionic strength were used:- 0.1 M NaCl in TE buffer, 0.25 M NaCl in TE buffer and 2 M NaCl in TE buffer.

The fractions eluted from this column were dialyzed against TE buffer for at least 3 changes at 4°C to remove the NaCl.

The 60% ammonium sulphate fraction from strains CU827 and C were applied separately to the column. Filter binding assays were carried out on the various eluted fractions. In this case a significant increase in retained labelled DNA was observed when fractions of strain C were compared with the corresponding fractions of CU827 (see Table 5.5).

5.7 DNA cellulose column chromatography.

DNA cellulose chromatography provides selective purification of proteins that function on DNA. The most characteristic feature of DNA binding proteins is

Table (5.5)

Filter binding assays of DEAE-cellulose column fractions.
The control sample uses sterile water in stead of
a protein fraction.

Strain	Sodium chloride fraction	Readings in CPM	
CU827	0.1 M	3365	
	0.25 M	1061	
	2 M	3987	
Control sample		207	
Strain C	0.1 M	2745	
	0.25 M	34155 39067	Duplicate sample
	2 M	3028	
Control sample		240	

their ability to recognize and bind *in vitro* to any purified DNA under certain physiological conditions.

DNA-cellulose matrix was prepared as follows:- 4g of CF11 cellulose was suspended in 12 ml DNA solution containing 8mg of highly polymerized DNA (from calf thymus) in TE buffer. The suspended matrix was spread out in a petri dish, covered with a gauze and left overnight to dry. The matrix was ground to a powder and then lyophilized to remove the remaining water. After that the powder was suspended in 20 ml TE buffer containing 0.05 M NaCl and the suspension was transferred to a column (25x1cm). After packing the column it was washed with 50 ml TE buffer containing 0.05 M NaCl.

The fractions eluted from the DEAE-Cellulose column for strain C which contain epsilon protein were mixed and applied to the column after it had been made to 0.05 M NaCl. The unbound proteins were washed out with 15 ml TE buffer containing 0.05 M NaCl. The bound proteins were eluted by a continuous gradient consisting of two solutions of 10 ml TE buffer containing 0.05 M and 0.8 M NaCl respectively. Fractions of 2 ml were collected and dialysed against TE buffer with at least 4 changes at 4°C.

The fractions collected were tested for the presence of epsilon protein by filter binding and gel retardation assays using the same procedures described above except that the amount of protein sample was 1 µg.

The gel retardation assay was also carried out using PvuII-BglII DNA from pGMM201 (see Fig.5.2B), which had been labelled by "filling in" the recessed 3' end (see section 2.22) and polyacrylamide gel rather than agarose gel electrophoresis.

The fractions eluted from the DNA-cellulose column after applying the active DEAE-cellulose fractions from strain C were examined. Particular fractions reveal binding activity with the substrate DNA fragment. The clear retardation of the

Fig. (5.2).

A. 1.5% Agarose gel electrophoresis of pDUB2416 cut with EcoRI & PvuII, and EcoRI & HincII to generate two different fragments. These are EcoRI-PvuII (100bp) and EcoRI-HincII (75 bp).

Lanes 1 & 8 are λ DNA marker digested with PstI. Lane 2 to 7 pDUB2416 digested with EcoRI & PvuII. Lane 9 to 12 pDUB2416 digested with EcoRI & HincII.

B. 0.7% Agarose gel electrophoresis illustrating:-

Lane 1 to 4 pGMM201 cleaved with PvuII & BglII.

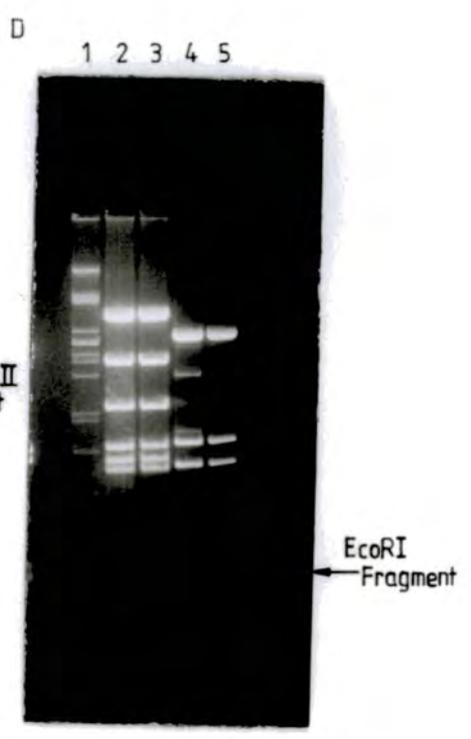
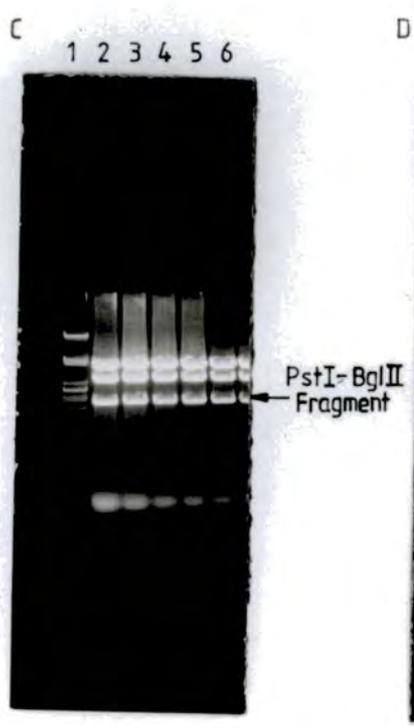
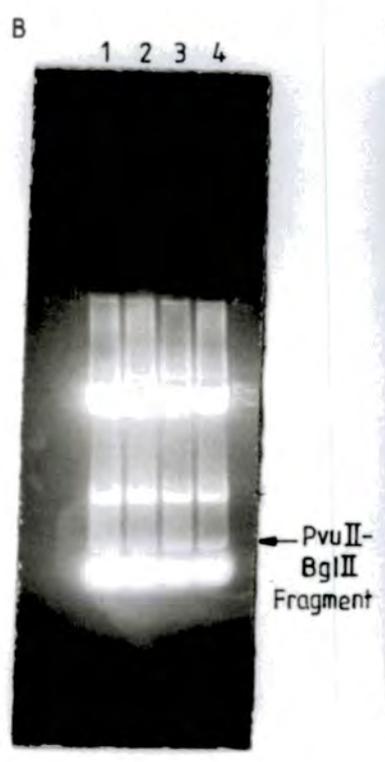
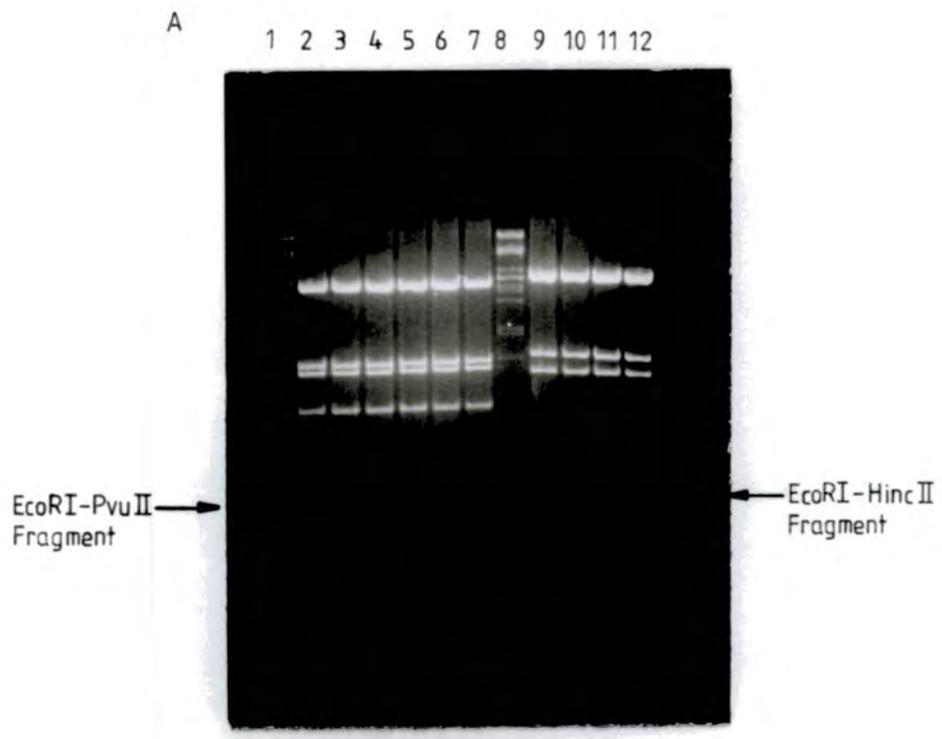
C. 0.7% agarose gel electrophoresis showing:-

Lane 1 is λ DNA marker. Lane 2 to 6 pGMM201 cleaved with PstI & BglII.

D. 1.5% Agarose gel electrophoresis showing:-

Lane 1 is DNA marker. Lanes 2 & 3 pDUB2416 restricted with PvuII & HincII.

Lanes 4 & 5 pDUB2416 restricted with EcoRI & PstI.



labelled PvuII-BglII fragment in the polyacrylamide gel (see Fig.5.3) provides convincing evidence for the presence of epsilon protein in these fractions.

The measurements of retained labelled DNA fragment in filter binding assays using the same fractions that show gel retardation give further evidence that these fractions contain epsilon protein (see Table 5.6).

5.8 Purification of radioactive epsilon protein.

The protein concentration in the fractions eluted from the DNA cellulose column using the above methods for purification of epsilon protein is low. To increase this concentration an alternative approach was used. This involves expression of the *ilvY^c* gene in a T7 system to radiolabel the accumulated epsilon protein.

Purification was carried out as follows:-.

A 1 litre culture of strain A (see section 4.3) was grown at 30°C in L-broth containing ampicillin, kanamycin and IPTG. This culture was then heat shocked at 42°C in a water bath for 45 minutes. Incubation was continued at 37°C for 3-4 hours to allow the accumulation of epsilon protein. The cells were harvested by centrifugation and then frozen at -20°C. During the preparation of the above cells a 10 ml culture of the same strain was grown at 30°C in M9 media (see section 4.4), heat shocked at 42°C and then labelled with 200 µCi of ³⁵S-methionine. The sample was incubated further for 2-4 hours at 37°C. Cells were harvested and mixed with those obtained from the 1 litre culture.

These mixed cells were lysed after suspending in 10 ml of TE buffer using sonication at a setting of 22 for 6 times 30 seconds. Next, the lysed cells were clarified by centrifugation for 30 minutes at 15000 rpm and the supernatant was removed for protein purification.

The previous steps for purification were followed with slight modifications. A continuous gradient was used to elute the bound proteins from the DEAE cellulose

Fig (5.3).

A. Autoradiograph of an 8% polyacrylamide gel for binding of labelled PvuII-BglIII DNA to fractions containing epsilon protein eluted from the DNA-cellulose column.

Lane 1 is labelled PvuII-BglIII fragment acting as a size control. Lane 6 and 7 show retardation of the this fragment. Lanes 2, 3, 4, 5, 8 & 9 no retardation occurs, these samples use various fractions from the DNA-cellulose column. Lanes 6 & 7, use protein from fraction 12 in table 5.6.

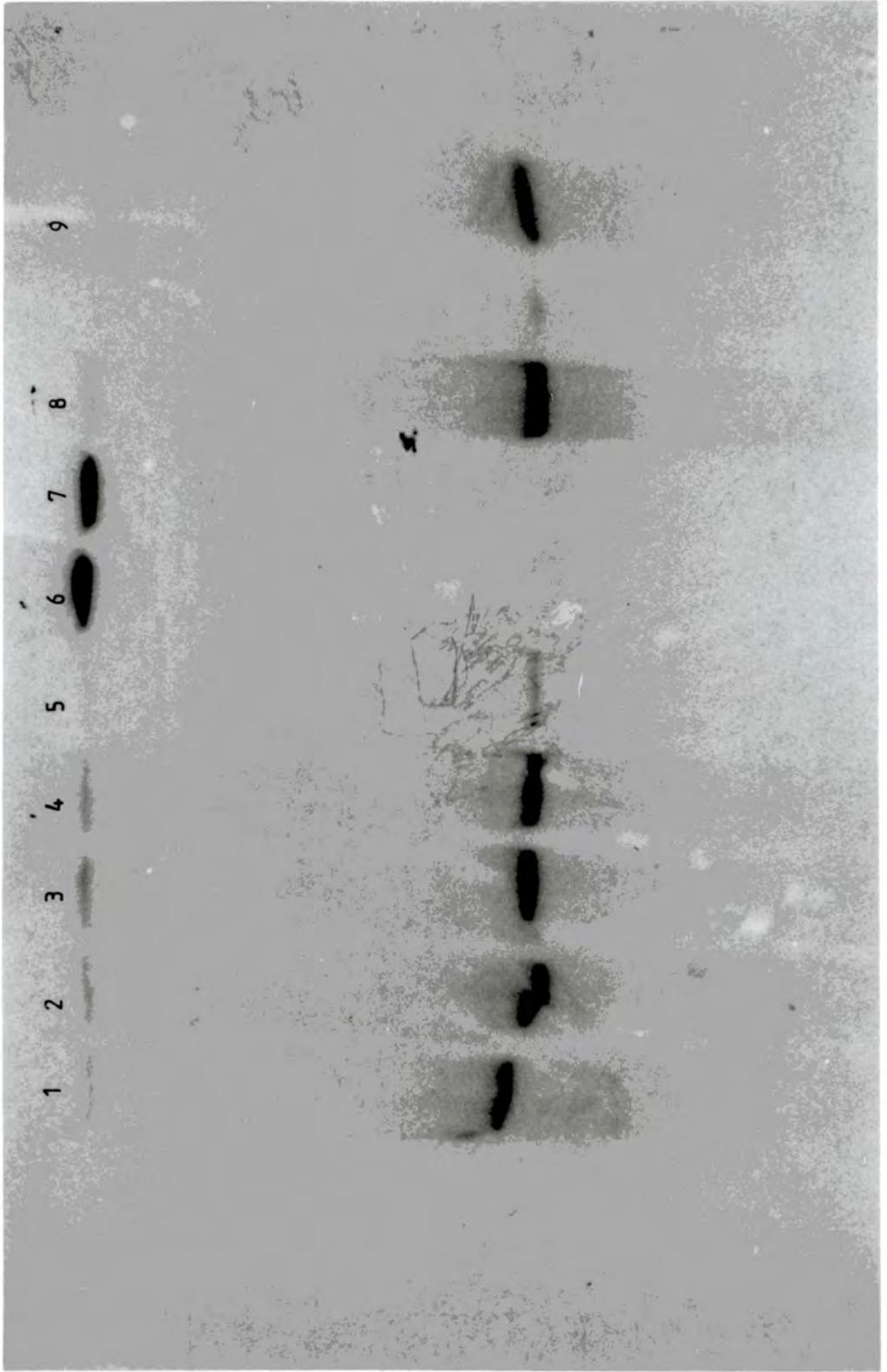


Table (5.6).

Filter binding assays of DNA-cellulose column fractions.
The control sample uses sterile water in stead of
a protein fraction.

Sample number	C.P.M.
1	51047
2	77495
3	77508
4	64442
5	81246
6	81879
7	72093
8	77098
9	52218
10	53729
11	42557
12	189749
13	90951
14	53978
15	53248
16	57167
Control sample	850

column rather than a step gradient. This gradient consisted of 2x15 ml solutions, the first is TE buffer and the other TE buffer with 1 M NaCl. The detection of epsilon protein was carried out by measuring the radioactivity of each fraction using liquid scintillation counting.

The peak fractions eluted from the DEAE-cellulose column which show radioactivity (see Table 5.7 & Fig.5.4) were mixed and dialyzed against TE buffer for at least 4 changes at 4°C. Then, the mixed fractions was applied to a DNA-cellulose column for the next step of purification.

The fractions eluted from the DNA-cellulose column were assayed for radioactivity (see Table 5.8 & Fig.5.5) and the active fractions were mixed and dialyzed against TE buffer for at least 4 changes at 4°C. This fraction was used for gel retardation assays using various labelled DNA fragments of the *ilvYC* system and to determine the native molecular weight of epsilon protein.

To examine the purity of the fractions eluted from the DNA-cellulose column fractions were subjected to SDS-PAGE and silver staining (see Fig.5.6).

5.9 Purification of epsilon protein by gel filtration.

This technique was originally introduced by Porath and Foldin (1959) and is used for the separation of protein mixtures by molecular weight.

Sephadex G-200 was used as the gel matrix due to its particular range of molecular weight fractionation. 8g of G-200 (medium grade) was suspended in 400 ml TE buffer and the gel was left in a boiling water bath for 5 hours to swell. Then, the gel suspension was cooled to room temperature and applied to the column (70x1.6cm). The packing of the column was carried out overnight using TE buffer as an eluant.

The homogeneity of the gel matrix was checked by applying 1 ml of Blue Dextran G-2000 at a concentration of 2 mg/ml. Fractions were collected every 2

Table (5.7).

DEAE-cellulose chromatography of labelled epsilon protein.

Sample number	C.P.M.	Protein concentrations in mg/ml measured at 330 nm
1	56	0.420
2	33	0.167
3	40	0.195
4	38	0.188
5	29	0.154
6	24	0.124
7	38	0.195
8	44	0.210
9	49	0.250
10	102	1.000
11	1208	2.375
12	2677	5.250
13	2225	4.750
14	2185	2.125
15	1622	1.750
16	1106	1.125
17	547	0.600
18	307	0.255
19	291	0.172
20	170	0.075

Figure(5.4) DEAE-CELLULOSE
COLUMN CHROMATOGRAPHY

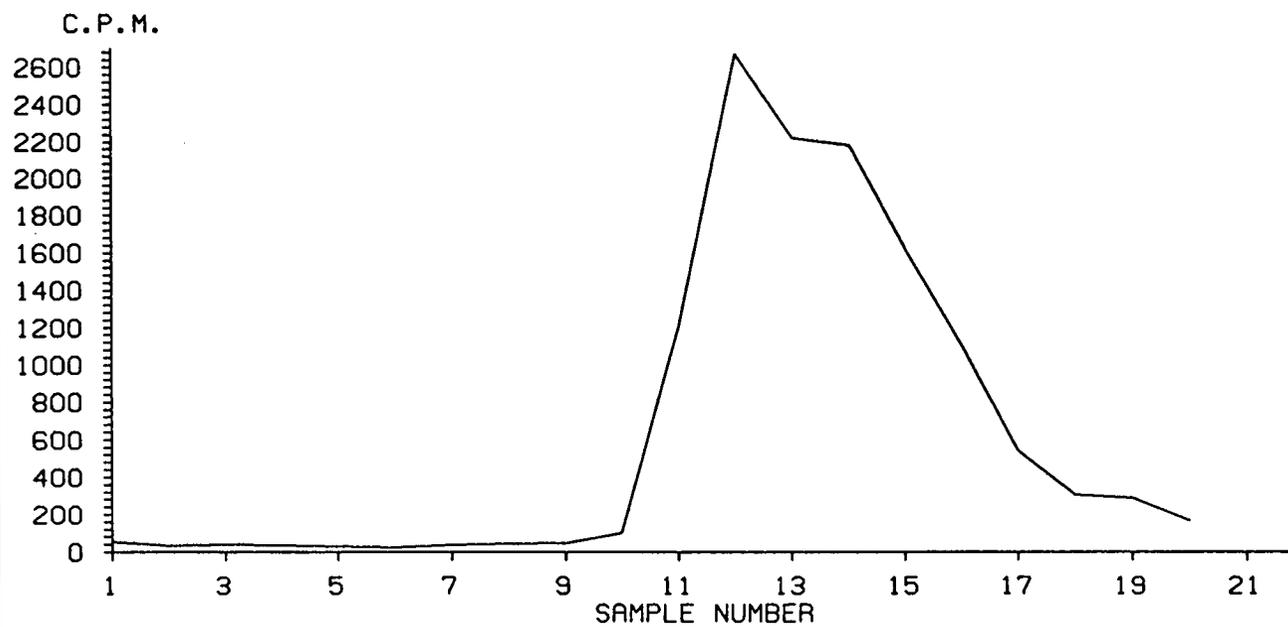


Table (5.8).

DNA-cellulose column chromatography.

Sample number	C.P.M.
1	1286
2	741
3	419
4	273
5	252
6	199
7	155
8	148
9	125
10	119
11	109
12	799
13	280
14	270
15	306
16	275
17	120

Figure(5.5) DNA CELLULOSE
CHROMATOGRAPHY

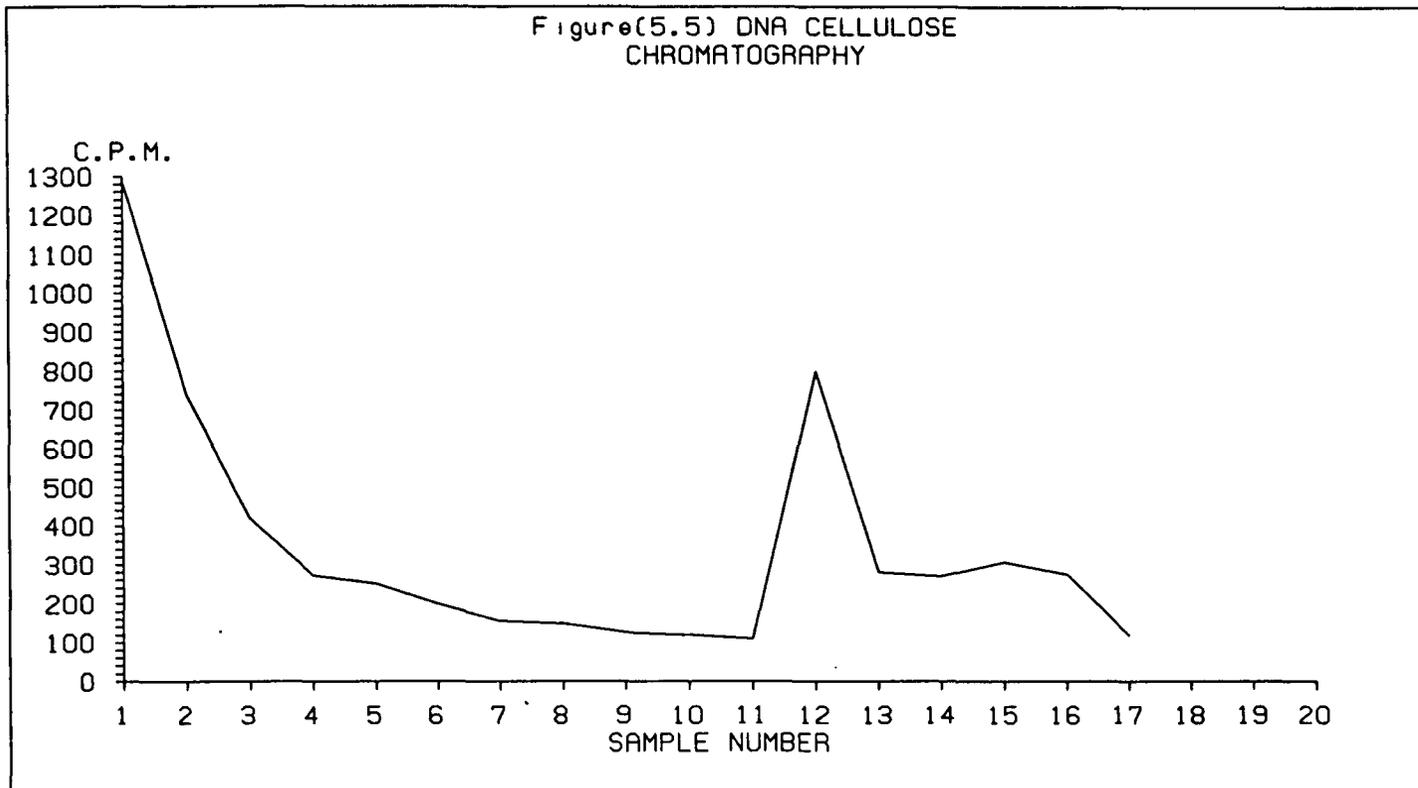
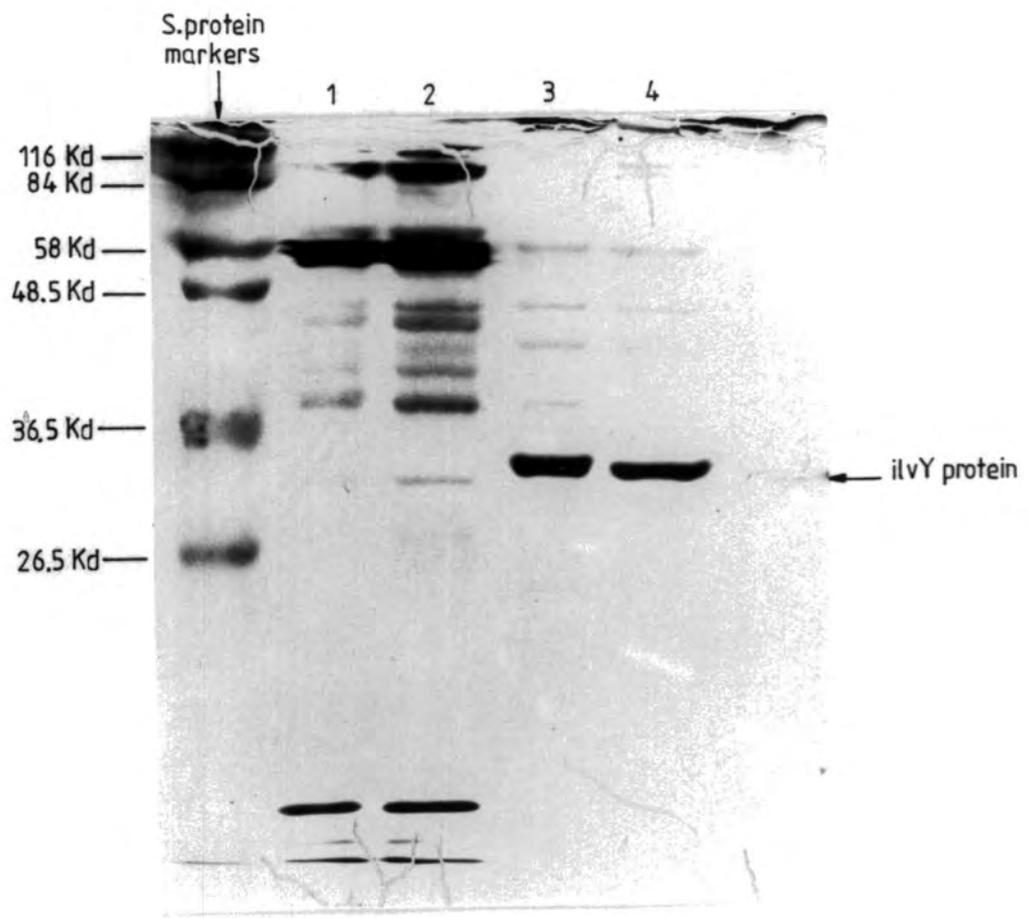


Fig. (5.6).

A. 10% SDS-PAGE gel for demonstrating the purity of of peak fractions eluted from DNA-cellulose column.

Lanes 1 & 2 are proteins from the flow through fraction. Lane 3 & 4 are proteins from fractions that represent peak activity.



minutes and the protein concentration estimated using the Bio-Rad microassay method in microtiter plates (Gibco, BRL). The absorbance were measured by a Titertek Multiskan MCC plate reader at 595 nm with filter number 8. The void volume can be calculated by multiplying the fraction number that shows peak absorbance by the fraction volume.

The next step involved the calibration of the column with particular protein markers. Thus, a mixture containing equal concentrations (2 mg/ml) of each of the following markers:- Carbonic anhydrase (29 kd), Bovine serum albumin (66 kd), Alcohol dehydrogenase (150 kd) and sweet potato β -Amylase (200 kd) was applied. The absorbance in the eluted fractions was estimated as described above.

After determining the elution volume for each protein marker (Ve1, Ve2, Ve3 and Ve4), these elution volumes were divided separately by the void volume to obtain the relative mobility for each protein marker. A graph was plotted of the log of protein molecular weight against relative mobility in order to construct a calibration curve for the column.

Following the equilibration of the gel filtration column, 1 ml of DNA-cellulose purified labelled epsilon protein was applied to the column and fractions collected every 2 minutes. After determining the fraction that contains the epsilon protein by measuring the radioactivity, the elution volume for epsilon protein was calculated (Vex) and consequently its relative mobility.

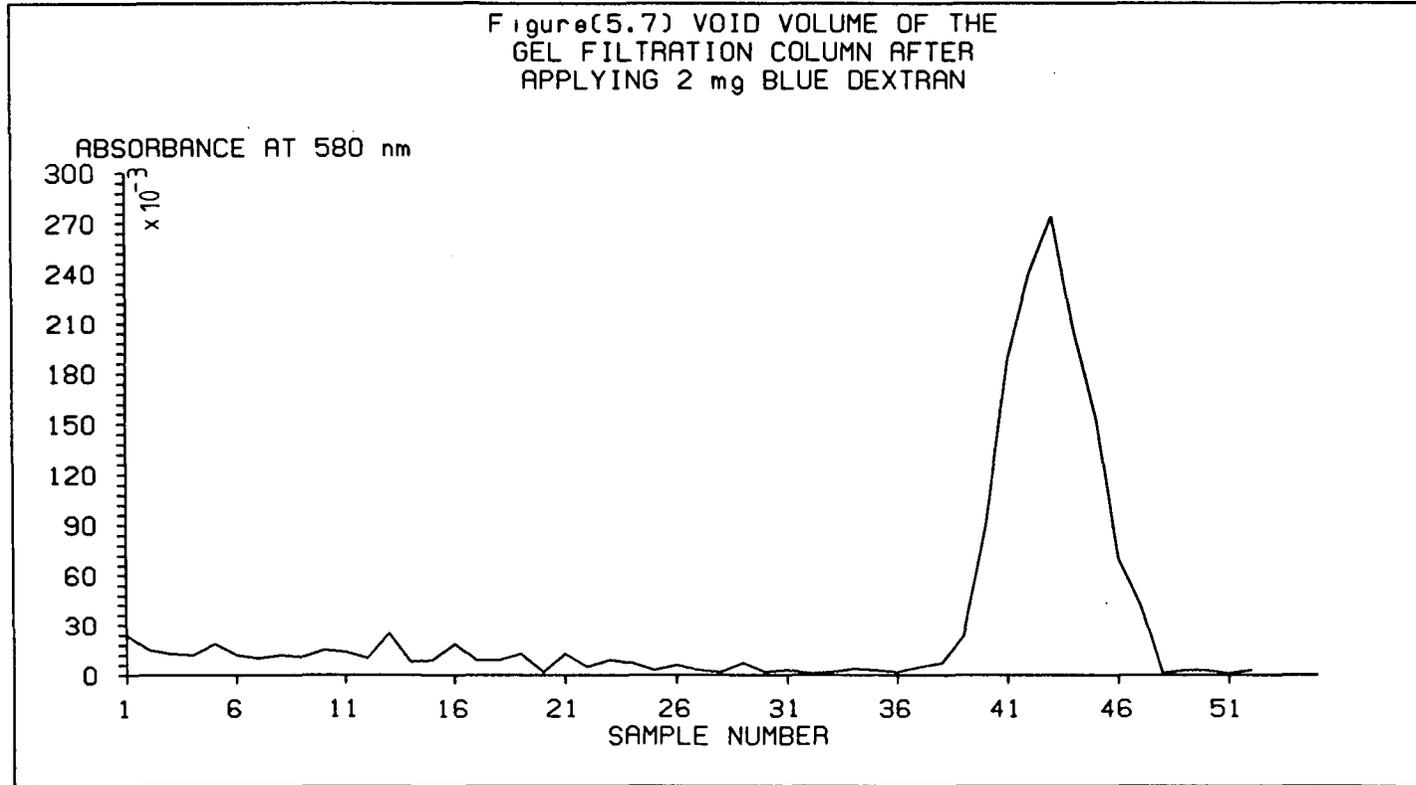
5.10 Results for gel filtration column.

After applying the Blue dextran the fractions eluted from the column had a volume of 0.7 ml. Peak of absorbance at 595 nm was obtained with fraction number 43 (see Fig.5.7).

Thus, the void volume of the column = $43 \times 0.7 = 30.1$ ml .

The absorbance of the fractions eluted from the column after applying the

Figure(5.7) VOID VOLUME OF THE
GEL FILTRATION COLUMN AFTER
APPLYING 2 mg BLUE DEXTRAN



protein markers showed 4 peaks at fraction numbers 68, 76, 106 and 119 (see Fig.5.8).

The elution volume for the 200 kd protein marker (Ve_1) = $68 \times 0.7 = 47.6$ ml.

The elution volume for the 150 kd protein marker (Ve_2) = $76 \times 0.7 = 53.2$ ml.

The elution volume for the 66 kd protein marker (Ve_3) = $106 \times 0.7 = 74.2$ ml.

The elution volume for the 29 kd protein marker (Ve_4) = $119 \times 0.7 = 83.3$ ml.

These elution volumes for the protein markers (Ve_1 , Ve_2 , Ve_3 and Ve_4) are divided by the void volume of the column to obtain the relative mobilities. These are 1.581, 1.767, 2.465 and 2.767 respectively, which were plotted against the log of the molecular weight of the protein markers to obtain the calibration curve (see Fig.5.9).

The radioactively labelled upsilon protein eluted from the column at fraction number 64 (Fig.5.10).

The elution volume for the upsilon protein (V_{ex}) = $64 \times 0.7 = 44.8$ ml.

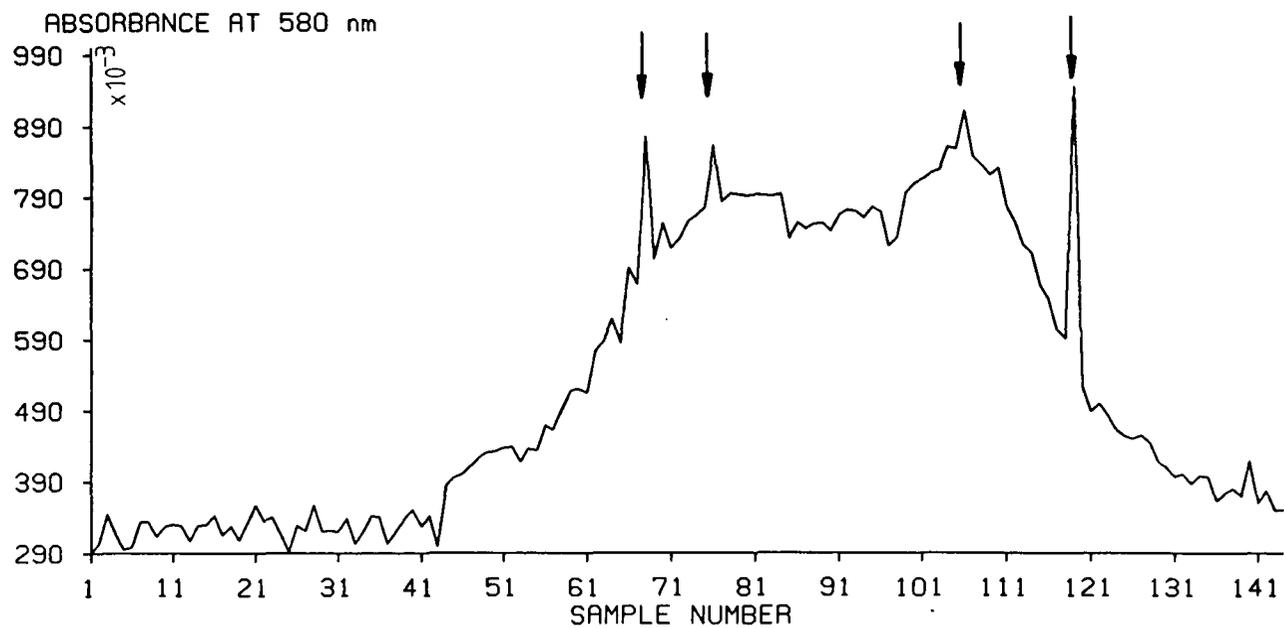
The relative mobility for upsilon protein is 1.488 and from the calibration curve, the native molecular weight was calculated to be 211.332 kd. From the autoradiograph (see Fig.4.8) it is clear that the monomer molecular weight of upsilon protein is 35 kd. Thus native upsilon protein appears to be a hexamer.

Upsilon containing fractions were applied to the SDS-PAGE gel and then silver-stained to determine the purity of the fractions (see Fig.5.11).

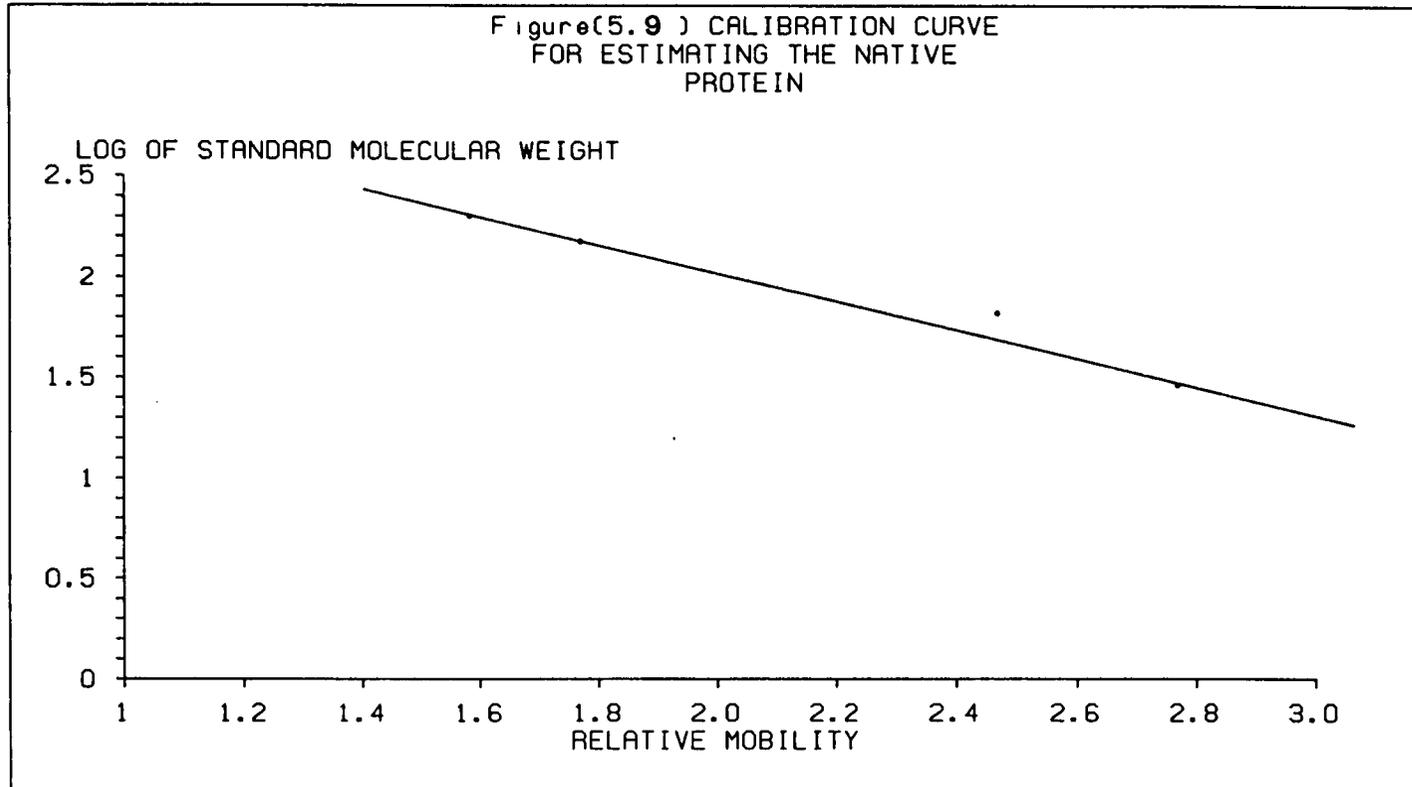
5.11 Retardation pattern of various DNA fragments of the *ilvYC* genetic system.

The DNA binding activity of the fractions eluted from gel filtration was examined with various DNA fragments of *ilvYC* system. The following DNA fragments were labelled by T4 polynucleotide kinase. EcoR1-EcoR1 (135 bp), EcoR1-PvuII (100 bp) and EcoR1-HincII (75 bp). (See Fig.5.2A & 2D). These DNA fragments

Figure(5.8) CALIBRATION OF THE
GEL FILTRATION COLUMN WITH
STANDARD PROTEIN MARKERS



Figure(5.9) CALIBRATION CURVE
FOR ESTIMATING THE NATIVE
PROTEIN



Figure(5.10) GEL FILTRATION COLUMN
CHROMATOGRAPHY OF LABELLED UPSILON
PROTEIN

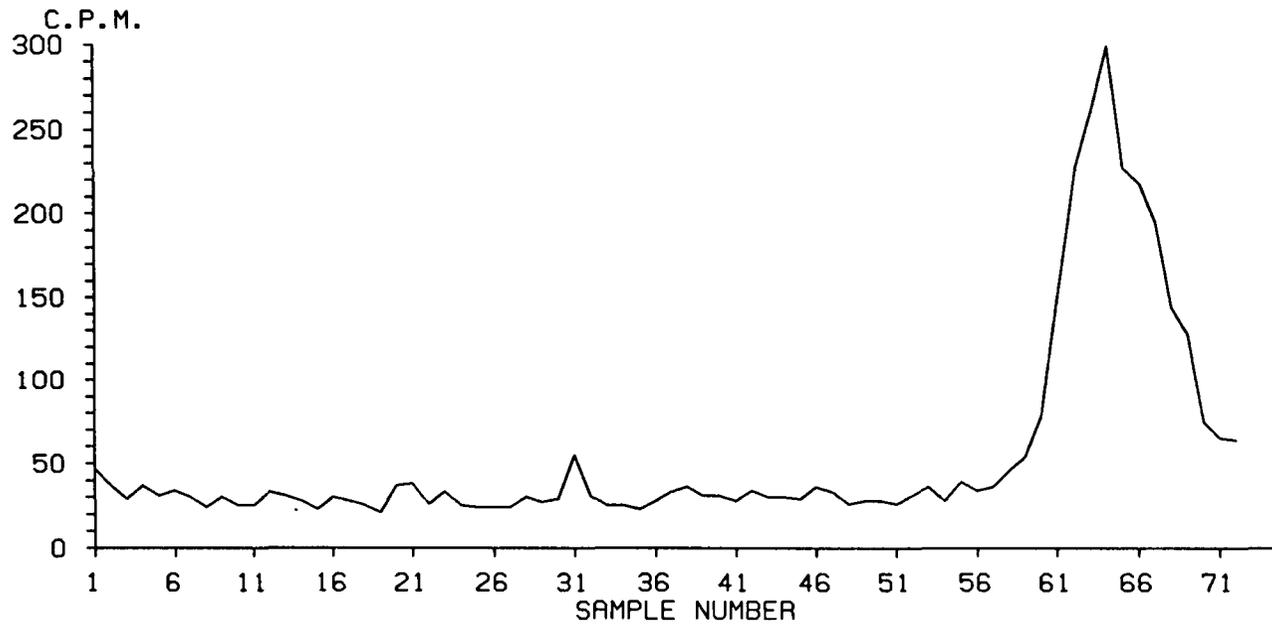


Fig (5.11).

10%SDS-PAGE gel of peak fractions eluted from gel filtration column to illustrate the purity of the upsilon protein.

Lanes 1 & 3 samples from peak fractions.

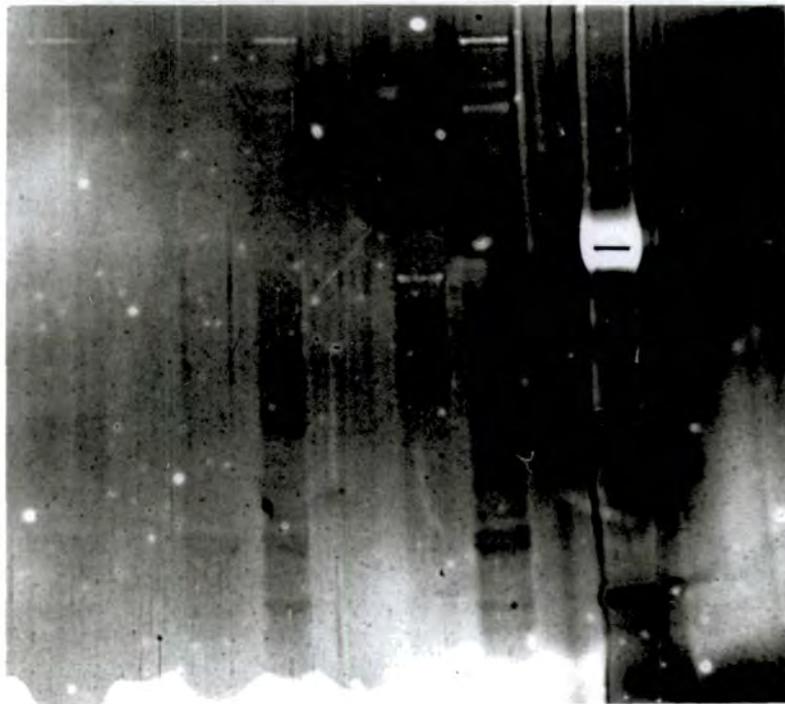
Lane 2 sample from a non-peak fraction .

1

2

3

S. protein markers



66 Kd

45 Kd

36 Kd

24 Kd

14.2 Kd

were chosen in an attempt to locate the binding site of upsilon protein. The third fragment (EcoRI-HincII) is proposed to contain the overlapping promoter region for *ilvY* and *ilvC* (Wek and Hatfield, 1986). Retardation was detected only with this DNA fragment (see Fig.5.12) and no retardation was observed with the other two fragments (see Fig.5.13A & 13B).

5.12 Examination of the *in vitro* properties of upsilon protein.

In order to examine the *in vitro* properties of upsilon protein, two experiments were carried out. Firstly, *in vitro* transcription of the *ilvC* gene.

The *in vitro* transcription of the *ilvC* gene was performed as follows:- 6 reaction samples were prepared, 3 samples used supercoiled DNA of pDUB2416 containing the entire *ilvC* gene and part of the *ilvY* gene (PstI-BglII fragment) as a template while the remaining 3 used linearized DNA (purified PstI-BglII fragment) as a template. The first reaction contains the following, 1 μ g DNA of the plasmid pDUB2416, 10 μ l 10x transcription buffer, 10 μ l 1 M KCl, 10 μ l of nucleoside triphosphate mixture, 50 μ Ci of ^3H -UTP (40 Ci/mmol) and 1 μ l RNA polymerase (1 unit/ μ l).

The second reaction was identical plus 1 μ g upsilon protein from the gel filtration column.

The third reaction was the same except that upsilon protein was added 12 minutes after RNA polymerase addition.

Reaction mixtures 4, 5 and 6 were identical to 1, 2 and 3 respectively except using the linear template.

Reactions 1, 2, 4 and 5 were incubated at 37°C for 20 minutes while the reactions 3 and 6 were incubated for 12 minutes before the addition of 1 μ g of upsilon protein and then incubated for a further 8 minutes. 1.5 ml of 10% TCA was added to each sample and placed on ice for 10 minutes. The precipitated RNA

Fig. (5.12)

Autoradiograph of an 8% polyacrylamide gel of labelled EcoRI-HincII DNA bound to upsilon protein from fractions eluted from gel filtration column.

Lane 1 is labelled EcoRI-PvuII fragment acting as a size control. Lane 2 is labelled EcoRI-HincII DNA without upsilon protein. Lanes 3, 4 & 5 are labelled EcoRI-HincII DNA bound to upsilon protein.

1

2

3

4

5

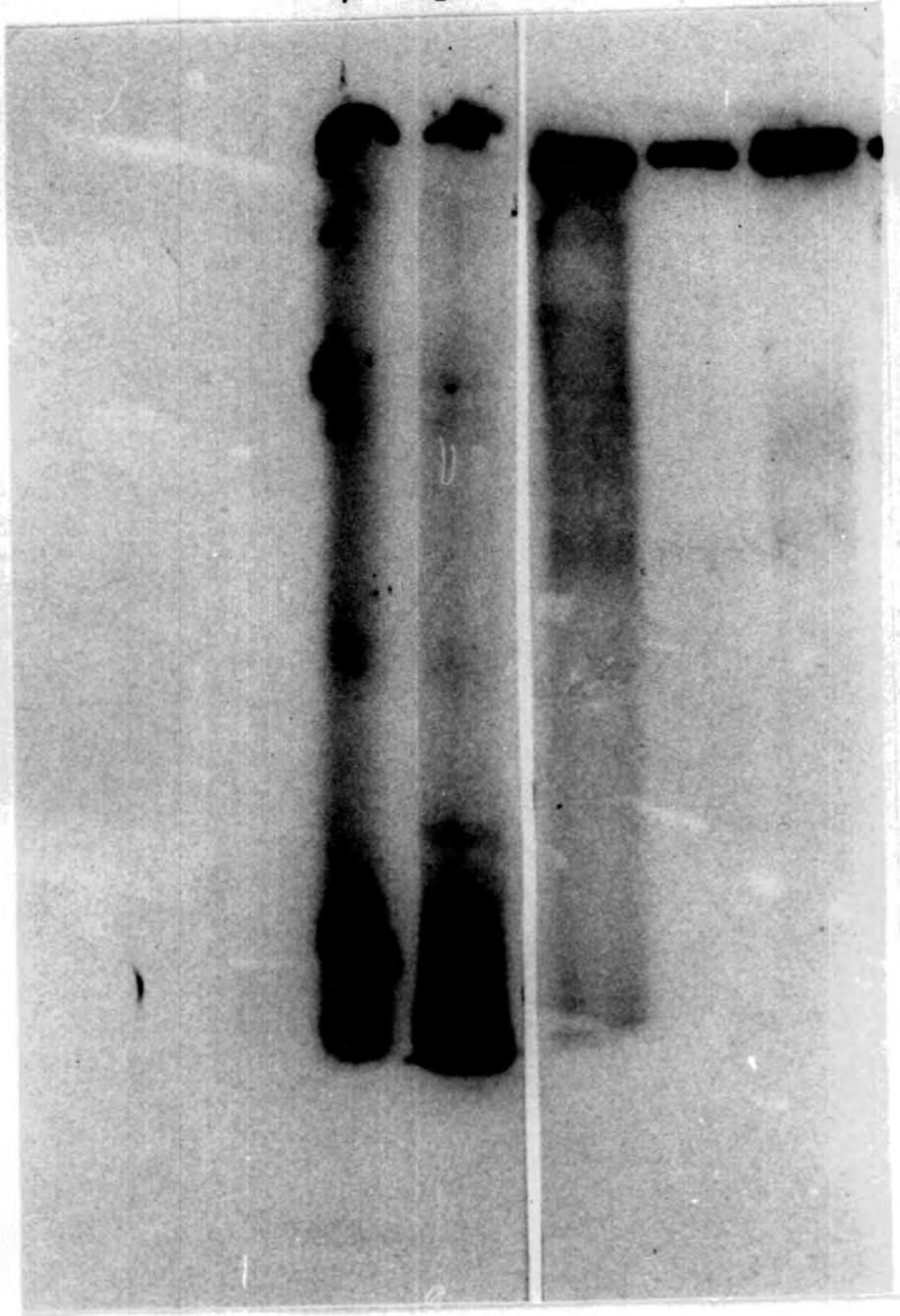


Fig. (5.13).

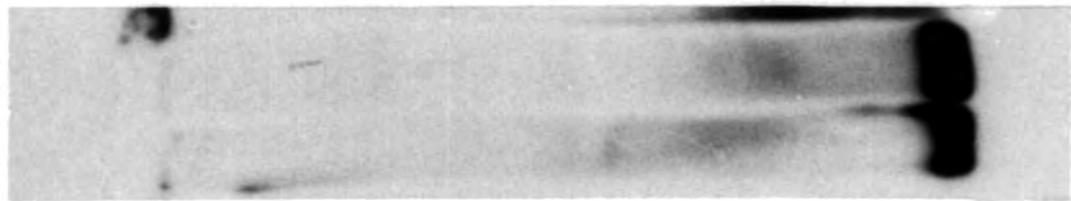
A. EcoRI-EcoRI fragment (135 bp) was 5' end labelled and bound to epsilon protein from fractions eluted from the gel filtration column, the binding reaction was loaded onto 8% polyacrylamide gel, electrophoresed, dried then autoradiographed. Lane 1 is the labelled fragment as a size control. Lane 2 shows no retardation occurs when this fragment is bound to epsilon protein.

B. Autoradiograph of binding the labelled EcoRI-PvuII DNA fragment to epsilon protein eluted from the gel filtration column.

Lane 1 represents the labelled fragment as a size control. lane 2 is the labelled fragment bound to epsilon protein, no retardation has been observed.

A

1 2



B

1 2



molecules from each sample were collected onto GF/C filters prewashed with 1 ml 10% TCA. The filters were then washed by 10 ml of 10% TCA followed by 5 ml absolute ethanol. The filters were dried and the radioactivity counted using liquid scintillation.

5.13 Results for *in vitro* transcription of *ilvC*.

From the measurement of the total number of counts per minute, which represents the incorporation of labelled nucleotides into RNA, it is clear that the addition of epsilon protein eluted from the gel filtration column shows a different pattern of activation between using supercoiled DNA and linearized DNA as a template (see Table 5.9). The addition of the activator protein at the same time as RNA polymerase causes a slight repression of transcription which may be due to the competition between the enzyme and the activator for binding to the promoter region. Additionally, the degree of activation of transcription was larger with linear template than with supercoiled DNA and only when the activator was added 12 minutes after transcription initiation.

5.14 *In vitro* transcription-translation of *ilvC*.

The first reliable method of bacterial cell-free coupled transcription-translation was described by Devries and Zubay (1967) and later modified by Zubay (1973) and Collins (1979). The cell-free system devised by Zubay involves the preparation of a crude extract from *E. coli* which contains all of the enzymes and co-factors necessary for transcription and translation. In addition the extract must be supplemented with amino acids, an energy regenerating system, tRNA and nucleotides.

An alternative method for the preparation of a cell-free coupled transcription-translation system is described by Gold and Schweiger (1971) which involves the

Table (5.9).

In vitro transcription of the plasmid pDUB2416 and the linear PstI-BglIII DNA fragment.

o

Sample number	Counts/minute of in vitro synthesized RNA molecules	
1	4211	
2	4088	
3	5202	
4	4127	
5	3940	
6	8143	Duplicate
	8764	sample

fractionation of various components of *E. coli* with subsequent reconstitution to produce an active system. This system has been used as widely as the Zubay system.

As the results of the *in vitro* transcription of *ilvC* DNA reveal that upsilon protein acts as an activator the following experiment was carried out in an attempt to show this effect in a cell-free coupled transcription translation system. The experiment was performed by using a commercial prokaryotic DNA-Directed translation kit which is a cell-free coupled transcription-translation system derived from *E. coli*. An S-30 extract was also prepared from CU827 ($\Delta ilvDAYC$) (see section 2.31). The extract from a wild-type *E. coli* strain MRE600 is already found in the kit.

Constituents of the reaction samples used in these experiments are shown in Table 5.10.

The autoradiograph of the *in vitro* synthesized proteins using supercoiled pDUB2416 as a template (see Fig.5.14), shows that the expression of *ilvC* was detected when using the S-30 extracts from wild type *E. coli*. However, the addition of upsilon containing fractions eluted from the DNA-cellulose and gel filtration columns did not show any significant increase in the level of *ilvC* expression.

The S-30 extract from CU827 showed no expression of *ilvC*. The addition of upsilon protein from the DNA-cellulose and the gel filtration columns to the above S-30 extract made no difference and no expression of the *ilvC* was detected.

Hence, no significant effect was detected on the expression of *ilvC* when upsilon protein was added to S-30 extracts from wild type *E. coli* and CU827. It is probable that the addition of upsilon protein to the wild-type S-30 system could not enhance the level of *ilvC* expression already observed. The protein concentration of the CU827 S-30 may have been too low to be effective.

Table (5.10).

Showing the constituents of various reaction samples used in in vitro expression of *ilvC*.

Sample number	DNA template	Sol. 2	Sol. 3	Sol. 5	S-35 methionine 1444Ci/mmol	S-30 extracts	Upsilon protein
1.	0	7.5	3	12	20 μ Ci	5 CU827	0
2.	+	5	3	5	20 μ Ci	5 CU827	0
3.	+	5	3	0	20 μ Ci	5 CU827	7.5 (from DNA-cellulose column).
4.	+	5	3	0	20 μ Ci	5 CU827	7.5 (from gel-filtration column).
5.	+	5	3	5	20 μ Ci	5 wild-type E. coli	0
6.	+	5	3	0	20 μ Ci	5 wild-type E. coli	7.5 (from DNA-cellulose column).
7.	+	5	3	0	20 μ Ci	5 wild-type E. coli	7.5 (from gel filtration column).
8.	0	5	3	12	20 μ Ci	5 wild-type E. coli	0

All the figures in the Table are in microlitres.

+ refers to the presence of pDUB2416 template (1 microgram).

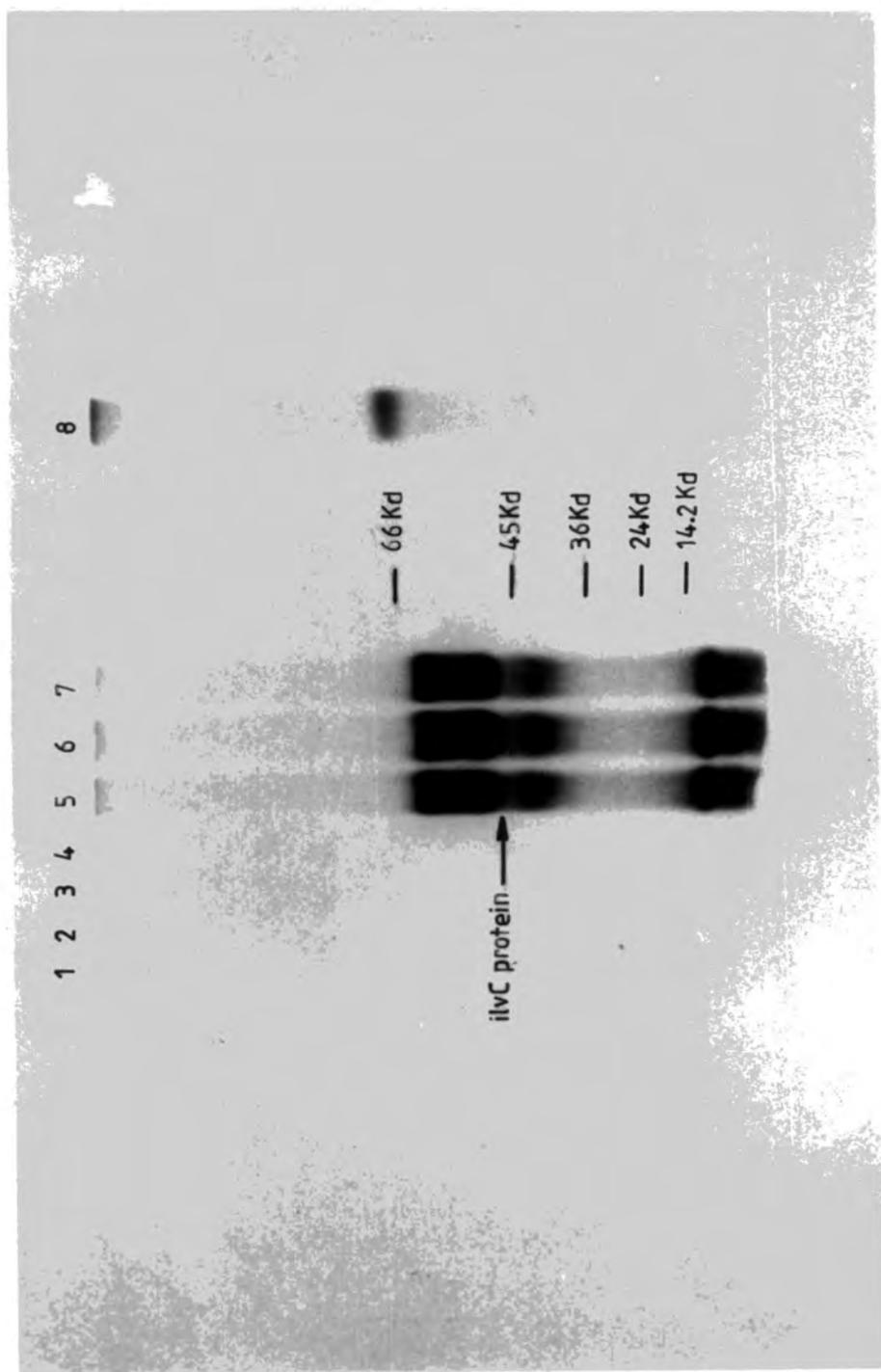
0 refers to the absence of the constituent.

Fig. (5.14).

A. Autoradiograph of an 10% SDS-PAGE gel of labelled proteins made in a cell free coupled transcription and translation system of pDUB2416 (encodes the entire *ilvC* gene).

Lane 1 represents a control *in vitro* reaction sample using S-30 extracts from CU827 without DNA. Lane 2 *in vitro* expression of pDUB2416 using S-30 extracts from CU827. Lanes 3 & 4 *in vitro* expression of pDUB2416 using S-30 extracts from CU827 and purified upsilon protein from fraction eluted from DNA-cellulose and gel filtration columns respectively.

Lane 5 *in vitro* expression of pDUB2416 using the S-30 extracts from wild-type *E. coli*. Lanes 6 & 7 represent *in vitro* expression of pDUB2416 using S-30 extracts from wild-type and purified upsilon protein from DNA-cellulose and gel filtration columns respectively. Lane 8 a control *in vitro* reaction sample using S-30 extracts from wild-type *E. coli*.



5.15 Discussion.

Retardation of the unlabeled PvuII-BglIII DNA fragment after titrating with various amounts of protein from the 60% ammonium sulphate fraction using agarose rather than polyacrylamide gels provided the first positive evidence for the presence of a DNA binding activity in this fraction (see Fig.5.1A). No retardation was detected with the same DNA fragment and the same amount of protein from other ammonium sulphate fractions (see Fig.5.1B).

To Confirm the above results *in vitro* transcription of the PstI-BglII DNA fragment was carried out to reveal any activation of transcription of *ilvC* by the various ammonium sulphate fractions. Surprisingly, proteins from the 30% ammonium sulphate fraction caused a significant apparent activation of *ilvC* transcription but inhibition of transcription was observed when using proteins from the 60% fraction (see Table 5.2). Initially, efforts were concentrated on the 30% fraction. This fraction was applied to a DEAE-cellulose column and the fractions eluted were used in examining activation of *ilvC* transcription. Again activation was detected in some fractions (see Table 5.3). At this point it was worth using another assay system for determining in which fraction *ilvC* protein occurs. This was in an attempt to resolve the discrepancy obtained with the two systems.

Thus, filter binding assays which are widely used in this field were exploited. The PstI-BglIII fragment was labelled *in vitro* by nick translation and used as a substrate. Although, the 60% ammonium sulphate fraction from strain C showed an increase in the level of retained labelled DNA (see Table 5.4) compared with fractions from other strains, the evidence was not in itself convincing. The 60% ammonium sulphate fractions from strains C and CU827 was applied separately to a DEAE-cellulose column and the eluted fractions examined for their ability to retain the labelled PstI-BglIII fragment. The proteins from the 60% fraction of strain C cause excellent retention of the labelled DNA (see Table 5.5). Hence,

two assays give a positive result for the presence of upsilon protein in the 60% ammonium sulphate fraction of strain C.

Effort was now concentrated on the 60% ammonium sulphate fraction of strain C. This fraction was applied to a DNA-cellulose column which is specific for purifying DNA binding proteins. The results obtained from the filter binding and gel retardation assays using specific fractions eluted from this column confirm that upsilon protein is present in the 60% ammonium sulphate fraction (see Table 5.6 & Fig.5.3).

The concentration of proteins in fractions eluted from the DNA-cellulose column is very low and as large amounts were needed to carry out further experiments, an alternative approach to the purification was investigated. This was to express the *ilvY^c* gene in a T7 expression system (for details see section 1.20). This type of expression increases the concentration of the protein in the DNA-cellulose fractions about tenfold so they can be used in subsequent experiments. Additionally upsilon protein was labeled so it can be detected easily.

The proposed location of the *ilvC* promoter region was confirmed to be in the overlapping region of the *ilvYC* genes. Fractions that contain upsilon protein bind only to the EcoRI-HincII fragment (see Fig.5.12). This finding supports the work of Wek and Hatfield (1986) who proposed that the promoter regions of *ilvY* and *ilvC* overlap in the small EcoRI-BglIII fragment of *ilvY*.

The *in vitro* behaviour of upsilon protein agrees with the behaviour of other purified activator proteins. Lee *et al.*, (1974) studied the *in vitro* activation of transcription of the *araBAD* operon which requires RNA polymerase, catabolite gene activator protein plus the operon specific activator *araC*. They had provided evidence that the *araC* activator acts *in vitro* at the level of transcription only.

Additionally, Arditti *et al.*, (1970) presented evidence for effects of catabolite gene activator protein (CAP) and cAMP *in vitro*. They stated that the transcrip-

tion of *lac* genes is stimulated sixfold by CAP and cAMP. This finding supports the notion that CAP and cAMP acts to promote expression of *lac* genes at the level of transcription.

Wild *et al.*, (1977) showed that the addition of isomeroreductase substrates at the time of adding rifampicin, amino acids and tRNA (at the beginning of the translation period) using an *ilvC::lacZ* fusion as a template had no effect on the expression of β -galactosidase. The identification of a specific promoter fragment to which ϵ binds, the activation of *in vitro* transcription of *ilvC* and the evidence of Wild *et al.*, clearly shows that ϵ is a true activator protein that acts at the level of transcription.

The failure of expression of *ilvC* in the presence of the S-30 extract of CU827 may be due to the lack of the *ilvY* gene in this strain, so, no ϵ protein will be present, and also the low overall protein concentration. Also this strain is RNase⁺ which causes degradation of the *in vitro* synthesized RNA molecules in the coupled transcription translation system. Although, active extracts have been prepared from several RNase⁺ strains including CU827 (Pratt *et al.*, 1981; Watson *et al.*, 1979).

In contrast, *E. coli* strain MRE600 lacks the major RNase activity of *E. coli* and its S-30 extract is often used for the identification of gene products using a plasmid or phage λ as a template. In addition it contains the wild-type *ilvY* gene and product, so the expression of the *ilvC* is possible in the presence of inducer (see Fig.5.14). However, the expression of *ilvC* in this experiment occurs in the absence of inducer.

The addition of ϵ protein to both CU827 and wild-type S-30 extracts did not stimulate *ilvC* expression. This is puzzling as the *ilvY^c* ϵ protein purified here, should act to stimulate *ilvC* expression in trans with the wild-type S-30 extracts. (See Fig. 5.14).

The appearance of a protein band in the control samples (see Fig.5.14) of the wild-type S-30 extract where no template DNA was used, may be due to the presence of residual DNA in the S-30 extracts which can serve as a template and as a result background polypeptides could be expressed.

Most DNA binding proteins are either dimers or tetramer *e.g.* CAP is a dimer (Decrombrughe *et al.*, 1984), *lac* repressor is a tetramer (Gilbert & Muller-Hill, 1966), λ repressor and *cro* protein are dimers (Gussin *et al.*, 1983). The finding that epsilon protein exists as hexamer in its native state is very unusual. This may be a function of binding to two divergent promoters simultaneously.

GENERAL DISCUSSION

General Discussion.

The *ilvY* and *ilvC* genes that encode the positive regulatory factor and the enzyme acetohydroxy acid isomeroreductase respectively, are adjacent on the *E. coli* chromosome. These genes are transcribed in opposite directions from overlapping divergent promoters (Wek and Hatfield, 1986 and this work). *ilvC* expression is under the positive control of a factor (upsilon protein) produced by the *ilvY* gene (Watson *et al.*, 1979). This positive regulatory protein binds to the *ilvC* substrates (α -acetohydroxybutyrate or acetolactate) and makes an induction complex which is essential for *ilvC* expression. Thus, in this case the upsilon protein can be considered as an auxiliary activator which helps RNA polymerase bind and initiate transcription efficiently.

Mutations have been identified in *ilvY* which make the expression of *ilvC* constitutive (Biel and Umbarger, 1980). The product of the *ilvY^c* gene is trans dominant and it is not a repressor in the absence of *ilvC* substrates. The constitutive mutant form no longer needs substrates for induction of *ilvC* expression. These characteristics of the mutant upsilon protein have already been demonstrated by Biel and Umbarger (1980). Comparing the above constitutive mutations in *ilvY* with constitutive mutations that occur in other some regulatory genes reveals similarity with *maltT^c* and *araC^c* mutations.

The positive control of gene regulation was mainly deduced from the work of Englesberg *et al.*, who concentrated their studies on the L-arabinose system of *E. coli* (Englesberg and Wilcox, 1974).

An example of positive control of gene expression is the *ara* operon. Mutants strains were isolated that express the *araBAD* operon constitutively. These mutants were classified in two classes, *araC^c* which synthesizes an *araC* product

active even in the absence of inducer (L-arabinose) (Englesberg *et al.*, 1965) and *araI^c* in which an initiator sequence *araI* is defined as the site of action of the activator. This latter mutation allows the expression of *araBAD* even in the absence of *araC* product (Englesberg *et al.*, 1969; Gielow *et al.*, 1971). The results obtained after introducing the *araC⁺* allele into strains carrying the above constitutive mutations to create merodiploid strains and growing them in medium free from arabinose, were unexpected. The constitutive expression produced by these mutations was repressed in these merodiploid strains (Englesberg *et al.*, 1969). Thus, the *araC⁺* allele was shown to repress the constitutive phenotype resulting from an *araC^c* or *araI^c* mutation. They concluded that the *araC* product is a repressor in the absence of inducer (L-arabinose) and an activator when inducer is present.

The maltose regulon in *E. coli* is believed to be positively regulated. *malT* protein is the activator for the *malPQ*, *malK* and *malEFG* operons (Schwartz, 1967a; Hofnung, 1974). Mutants have been isolated which express these three operons constitutively (Debarbouille *et al.*, 1978) by using a strain in which the *lac* operon is fused to the *malPQ* operon. Some of these mutations are located in the positive regulatory gene and called *malT^c*. In these mutants the modified *malT^c* is still active *i.e.* it has the ability to stimulate the expression of the *mal* operon even in the absence of maltose in the media. Debarbouille *et al.*, (1978) characterized this *malT^c* product by constructing merodiploid strains after introducing the *malT⁺* allele into *malT^c* strains. *malT^c* mutations are trans dominant, which is characteristic of positive control. They concluded that the *malT⁺* allele does not result in the repression of constitutive expression and as a result the *malT⁺* product is not a repressor in the absence of inducer (maltose). Thus, this constitutive mutation in *malT^c* shows similar characteristics with that of the *ilvY^c* mutation. Both of them encode products which are trans dominant and are

not repressors in the absence of inducer.

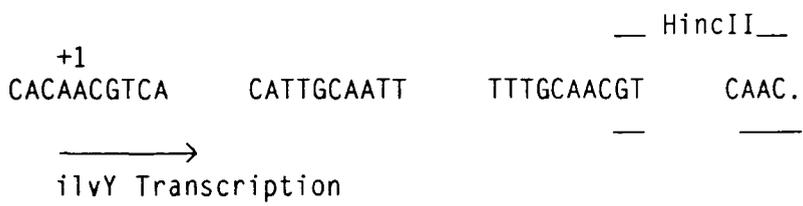
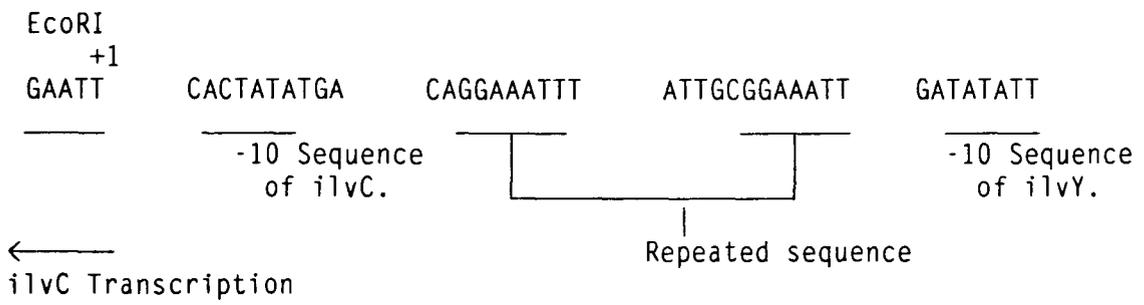
ilvYC is one of a few positive regulatory systems that have been described in *E. coli*. It is unique in being part of an amino acid biosynthesis operon. From the DNA sequence of the EcoRI-HincII region which contains the promoter regions of both *ilvY* and *ilvC* (see Fig.6.1), it is clear that the -35 regions of both promoters overlap. This creates a direct repeat sequence separated by 6 nucleotides. From the above information the mechanism of positive control via upsilon protein can be described, with some modifications, according to two models proposed to interpret the mechanism of positive control of λ repressor.

The first mechanism depends on the alteration in the intracellular upsilon protein concentration. So, at high concentration of upsilon protein, a dimer (2 monomer subunits bind together to form a dimer which comprises the DNA-binding form) will recognize the direct repeat sequence in the EcoRI-HincII DNA fragment and bind to it. This bound dimer occupies the repeated sequence and directs RNA polymerase to bind at the -10 region of the *ilvC* promoter. This direction of RNA polymerase by bound upsilon protein is enhanced not just by DNA-protein interactions but also by protein-protein interactions. In this situation the RNA polymerase recognizes the -10 region of the *ilvC* promoter more efficiently than that of the *ilvY* gene because the -10 region of the *ilvY* gene might be covered by the bound upsilon protein as it is nearer to the repeated sequence than that of *ilvC* by three base pair (see Fig.6.1). Thus, the TATA box of *ilvC* promoter is more accessible to RNA polymerase. Simultaneously, the bound upsilon protein represses transcription initiation at the TATA box of the *ilvY* gene and consequently *ilvY* expression is negatively autoregulated.

How does *ilvY* expression occur?. The answer according to this model is as follows:- In the case of a low concentration of upsilon protein there will be insufficient to form dimers and fill the direct repeat sequence. The RNA

Fig. (6.1).

B. Nucleotide sequence of the EcoRI-HincII fragment of *ilvY* gene showing the -10 regions of the *ilvYC* genes, direction of transcription and the repeated sequences which are proposed to be the binding sites for epsilon protein.



polymerase has a greater affinity for the *ilvY* promoter than the *ilvC* promoter. So, it will bind and initiate transcription of *ilvY*.

Thus, the following conclusions can be made. The transcription of the *ilvY* and *ilvC* genes occurs in a reciprocal manner depending on the intracellular concentration of upsilon protein. The upsilon protein exerts both activator (of *ilvC*) and repressor (of *ilvY*) functions when it binds to the repeated sequences.

An alternative model can be proposed to demonstrate the mechanism by which upsilon protein causes induction of *ilvC* expression. This model depends on the fact that due to the proximity of the *ilvY* and *ilvC* promoters, there will be competition for RNA polymerase by the two promoters. RNA polymerase prefers to bind to the *ilvY* promoter. This binding, which stimulates *ilvY* expression, will prevent *ilvC* expression, perhaps by steric hindrance. But in the presence of upsilon protein, it will stimulate binding of RNA polymerase to the *ilvC* promoter rather than to the *ilvY* promoter.

However, it must be emphasised that both of the preceding models of upsilon function are difficult to reconcile with a native protein that is a hexamer. It is possible that the mode of action of upsilon is much more complex, like the *ara* operon or like λ repressor protein (see below). DNA footprinting studies will be needed to firmly establish the exact location and number of upsilon binding sites.

The proposed mechanisms of upsilon protein action can be compared with the mechanisms of other regulatory proteins.

The major known regulatory protein which is an activator is catabolite gene activator protein (CAP). It plays an important role in the regulation of many operons, e.g., *lac*, *gal* and *mal*. This regulatory protein is activated by cAMP. The mechanism by which this complex activates transcription is not known exactly. Two models have been proposed, firstly the binding of this complex to DNA sequences upstream of the -35 region activate *lac* transcription by increasing the

binding constant of RNA polymerase *i.e.* RNA polymerase is attracted not just by DNA-Protein interactions but also by Protein-Protein interactions. Secondly, it could be that the cAMP-CAP complex acts at a distance by changing the structure of the DNA near the RNA polymerase binding site.

Some regulatory proteins can autoregulate their own production (Smith and Magasanik, 1971). In *S. typhimurium*, the degradation of the amino acid histidine is performed by two sets of enzymes encoded by separate operons. Each operon has its own promoter-operator. A single repressor binds to both operators and prevents transcription in the absence of histidine (Smith and Magasanik, 1971). In the presence of histidine the repressor binds histidine, dissociates from the operators and allows transcription to initiate in both operons. It is worth mentioning here that the repressor molecule synthesis is under control of histidine as well because it is encoded by a gene in one of the two operons. When histidine is scarce free repressor molecules accumulate which switch off the synthesis of more repressor. Thus, the synthesis of repressor is controlled by itself and the regulation comprise a negative autogenous regulation.

Two regulatory proteins play contrasting roles in determining the mode of growth of bacteriophage λ . These are λ repressor and *cro* which are encoded by the genes *cI* and *cro* respectively. λ repressor is required for lysogeny while *cro* protein for lytic phage growth. The ability of λ to grow in two alternate modes, lytically or lysogenically, depends upon the interaction of the regulatory proteins with three contiguous binding sites which are identified in the right operator on the λ chromosome. These are OR1, OR2 and OR3.

In lysogeny, the repressor is bound predominately to OR1 and OR2 but rarely at OR3. This binding leads to the repression of lytic genes, including *cro*, from the promoter P_R which overlaps with OR1. The transcription of *cI* from the promoter P_{RM} is stimulated by repressor bound to OR1 and OR2. These two effects of the

repressor are essential to maintain the lysogenic state. In addition, the repressor has another function, binding to OR3 represses expression from P_{RM} in order to autoregulate its synthesis. Thus, this regulatory protein can autoregulate its expression both negatively and positively.

Ptashne *et al.*, (1980) analyzed the mechanism by which repressor can repress at P_R and stimulate at P_{RM} when it bound to OR2. Sequence analysis has revealed that the two promoters P_R and P_{RM} , overlap an 82 bp region defined as the right operator. OR2 is closer to the transcription startpoint of P_R than that of P_{RM} by one base pair. They believe that this slight difference in position allows repressor to touch RNA polymerase bound to P_{RM} and to exclude it from P_R . Thus, according to this hypothesis the repressor directly stimulates transcription from P_{RM} possibly by directing protein-protein contact with RNA polymerase so stabilizing its binding or increasing its frequency of initiation of transcription.

An alternative model has been proposed to describe the mechanism of positive control by the λ repressor. This model holds that the stimulatory effect is a result of the small distance (82 bp) between the transcription startpoints of P_R and P_{RM} . The competing polymerase model holds that RNA polymerase binds preferentially to P_R and inhibits binding of RNA polymerase to P_{RM} perhaps by steric hindrance. The role of the repressor is to stimulate P_{RM} binding simply by preventing binding of polymerase to P_R (Meyer and Ptashne, 1980).

Meyer and Ptashne (1980) provided evidence to support the first model by analyzing the activity of various mutants. They showed that the ability of repressor to activate gene transcription is not a consequence of the proximity of the two λ promoters. Moreover, polymerase bound at P_R has little negative effect on the basal level of transcription at P_{RM} and that neither removal of P_R nor exclusion of RNA polymerase from P_R by another regulator (*cro*) suffices to activate P_{RM} , therefore they concluded that the ability of repressor to activate gene transcrip-

tion is not a trivial consequence of the proximity of the two λ promoters, rather, repressor activates P_{RM} directly.

In contrast to repressor the *cro* protein permits lytic growth by first binding to OR3 and so repressing synthesis from P_{RM} . As its concentration increases it will bind to OR2 and OR1 thereby turning off transcription from P_R (Johnson *et al.*, 1978; Meyer *et al.*, 1980). Thus, the *cro* protein action is less complex than that of repressor. It is strictly a negative regulator because it represses the transcription from P_{RM} and also turns off expression of its own and other lytic genes from P_R .

From the description above it is clear that the λ repressor does not require cofactors to stimulate the *cI* gene transcription, unlike CAP and ϵ protein.

An additional positive regulatory protein found in the bacteriophage λ is N protein. This regulatory protein turns on genes to the left of N including *cIII* and the recombination genes and genes to the right of *cro* including *cII* and the DNA replication genes O and Q. N protein works by enabling RNA polymerase to transcribe through DNA that would otherwise cause mRNA synthesis to terminate and is therefore called an anti-terminator. The mechanism of N protein action is still unclear. But it is known that this protein recognizes a specific sequence called *NUT* (N-utilization). There are two *nut* sites, one between P_L and the N gene and the other just to the right of the *cro* gene. The action of N protein is as follows. In the absence of N protein RNA polymerase ignores the *nut* site and dissociates from the DNA releasing mRNA when it reaches the transcription termination signals. But in the presence of N protein, which binds to the *nut* sites, as RNA polymerase passes over a *nut* it will be modified by N protein so that it will ignore certain but not all further termination signals.

The most complex positive regulator in bacteria, is *araC* protein in the *L - arabinose* operon system. This operon is under the control of two activators, the

araC and CAP proteins. The promoter region contains more than one activator binding sites. This was suggested by Englesberg and Wilcox (1974) who stated that the *araBAD* promoter includes two activator binding sites, one for *araC* protein and the other for CAP.

Later, groups of workers concentrated their studies on how these two activators function. (Ogden *et al.*, 1980; Lee *et al.*, 1981). For *araC* protein three binding sites have been identified. These are *araI*, *araO*₁ and *araO*₂, the latter is distal from the promoter regions of the *araBAD* and *araC* (Dunn *et al.*, 1984; Hahn *et al.*, 1984). *araO*₁ overlaps with the *araC* promoter. This overlapping leads to repression of the *araC* expression when the *araC* protein is accumulated in the cell. As a result of this *araC* expression is autoregulated. The role of the other binding sites *araI* and *araO*₂ is to cooperate with the binding site for CAP to mediate the activation and repression of the *araBAD* operon.

Many models have been postulated to interpret the mechanism of activation of *araC* and *BAD* promoters. One of these proposes that at low concentrations of CAP the *araC* protein binds to two binding sites *araI* and *araO*₂ and as a result of this binding, draws the DNA into a loop. The formation of this loop represents a repressing configuration so that the *araC* protein cannot activate RNA polymerase to initiate transcription at the *araBAD* promoter whether arabinose is bound to it or not, and at the same time repression of the *araC* promoter occurs.

On the other hand, the presence of cAMP-CAP at high levels will allow it to bind to its site which is located next to *araI*. This changes the *araC* protein conformation by touching it such that *araC* protein is no longer able to form the loop but will release the DNA at *araO*₂. In the meantime it activates RNA polymerase to initiate transcription at the *araBAD* promoter and also repress the *araC* promoter. The only possible way for *araC* protein to promote activation of *araBAD* expression is to bind to arabinose first. Thus, for efficient expression

of *araBAD* two molecules are required. These are cAMP-CAP and arabinose. *araC* promoter is activated when there is no *araC* protein as in this case RNA polymerase can bind to the *araC* promoter which overlaps *araO*₁ and initiate transcription.

A positive regulatory system also exists for the regulation of some genes involved in the biosynthesis of lysine. Straiger *et al.*, (1983a) identified the *lysR* gene and determined its exact location. *lysR* is closely linked to *lysA* and encodes an activator protein which is necessary for full expression of *lysA*. *lysA* is the structural gene code for the enzyme diaminopimelate decarboxylase (DAP). The two genes *lysA* and *lysR* are transcribed divergently. This organization is similar to that in *araC-araBAD* (Wilcox *et al.*, 1974) and *ilvC-ilvY* (Wek and Hatfield, 1986). Additionally, Straiger *et al.*, (1983b) presented evidence that *lysA* is under the positive control of the product of *lysR* and *lysR* protein represses expression of its own gene. Thus, the activator protein encoded by *lysR* has dual effects.

Straiger and Patte (1983) concentrated their studies on the molecular mechanism of this positive control. They examined the nucleotide sequence of *lysA-lysR*. The two ATGs for the *lysA* and *lysR* proteins are separated by 121 bp. They identified a 73 bp fragment carrying the promoter, first four codons and the information necessary for autoregulation of *lysR*. The same binding site could be involved in both effects of the *lysR* product, acting simultaneously as an operator for *lysR* expression and an initiator for *lysA*.

Models for the mechanism of activation of transcription by positive control are really much more difficult to formulate than those for negative control. Repressor proteins in negative control function primarily by preventing RNA polymerase binding at promoters. Activation by positive control could occur by stimulating either the binding or the isomerization of RNA polymerase, or by blocking competing RNA polymerase binding sites. Alternatively, the necessity for an auxiliary

protein for the activation of gene transcription might be related to the deviation of the positively controlled promoter sequences (-35 and -10 regions) from the promoter consensus sequence. So, these promoters require a protein which helps RNA polymerase binding to the promoter region. A third possibility for positive control is to function as a transcription antiterminator, like λ N protein. Finally positive control factors have also been shown to act as RNA polymerase σ (sigma) subunits which cause polymerase to recognize whole new classes of promoters.

REFERENCES

1. Adams, C. W., Rosenberg, M. & Hatfield, G. W. (1985). *J. Biol. Chem.*, **260**, 8535-8544.
2. Adelberg, E. A. (1954). *J. Amer. Chem. Soc.*, **76**, 4241.
3. Arditti, R. R., Eron, L., Zubay, G., Tocchini-Valentini, Connaway S. & Beckwith, J. R. (1970). Cold spring Harbor symp. *Quant. Biol*, **35**, 437.
4. Arfin, S. M., Ratzkin, B. & Umbarger, H. E. (1969). *Biochemical and Biophysical Research Communication*, **37**, 902-908.
5. Bachmann, B. J. (1983). *Microbiol. Rev*, **47**, 180-230.
6. Bassford, P., Beckwith, J., Berman, M., Brickman, E., Casadaban, M., Guarante, L., Saint-Girons. I., Sarthy, A., Schwartz, M., Shurman, H., & Silhavy, T. (1978). 245-261. In J. H. Miller and W. S. Renznikoff(ed). "The operon". Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
7. Bethesda Research laboratories, Inc. (1982). M13 cloning /dideoxy sequencing . BRL instruction manual.
8. Biel, A. J. & Umbarger, H. E. (1980). *J. Bacteriol.*, **146**, 718-724.
9. Birnboim, H. C. & Doly, J. (1979). *Nucl. Acid. Res.*, **7**, 1513-1523.
10. Blazey, D. L. & Burns, R. O. (1980). *J. Bacteriol.*, **142**, 1015-1018.
11. Blazey, D. L. & Burns, R. O. (1984). *J. Bacteriol.*, **159**, 951-957.
12. Busby, S. (1982). *J. Mol. Biol.*, **154**, **21**, 211-227.
13. Calhoun, D. H., Rimerman, R. A. & Hatfield, G. W. (1973). *J. Biol. Chem.*, **248**, 3511-3516.
14. Casadaban, M. J. (1976a). *J. Mol. Biol.*, **104**, 541-555.
15. Casadaban, M. J. (1976b). *J. Mol. Biol.*, **104**, 557-566.
16. Chamberlin, M. & Ring J. (1973). *J. Biol. Chem.*, **248**, 2245-2250.
17. Collins, J. (1979). *Gene*, **6**, 29-42.

REFERENCES

18. Debarbouille, M., Shuman, H. A., Silhavy, T. J. & Schwartz, M. (1978). *J. Mol. Biol.*, **124**, 359-371.
19. Decrombrugge, B., Busby, S. & Buc, H. (1984). *Science*, **224**, 831-838.
20. DeFelice, M., Guardiola, J., Esposito, B. & Iaccarino, M. (1974). *J. Bacteriol.*, **120**, 1068-1077.
21. Devries, J. K. & Zubay, G. (1967). *Proc. Natl. Acad. Sci., USA*, **57**, 1010-1012.
22. Dougan, G. & Sherratt, D. (1977). *Mol. Gen. Genet.*, **151**, 151.
23. Dunn, T. M. & Studier, F. W. (1983). *J. Mol. Biol.*, **166**, 477-535.
24. Dunn, T. M., Hahn, S., Ogden, S. & Schleif, R. (1984). *Proc. Natl. Acad. Sci., USA*, **81**, 5017-5020.
25. Englesberg, E., Irr, J. & Lee, N. (1965). *J. Bacteriol.*, **90**, 946-957.
26. Englesberg, E., Sheppard, D., Squires, C. & Meronk, F. (1969). *J. Mol. Biol.*, **43**, 281-298.
27. Englesberg, E. (1971). *In metabolic pathways: Metabolic Regulation*, ed. Vogel, H. J. (Academic, New York), **5**, 257-296.
28. Englesberg, E. & Wilcox, G. (1974). *Ann. Rev. Genet.*, **8**, 210-242.
29. Favre, R., Waiter, A., Puppo, S., Iaccarino, M., Noelle, R. & Freundlich, M. (1976). *Molec. Gen. Genet.*, **143**, 243-255.
30. Freundlich, M. (1977). *Cell*, **12**, 1121-1126.
31. Friden, P., Newman, T. & Freundlich, M. (1982). *Proc. Natl. Acad. Sci., USA*, **79**, 6156-6160.
32. Fried, M. & Crothers, D. M. (1981). *Nucl. Acid. Res.*, **9**, 6505-6525.
33. Garner, M. M. & Revzin, A. (1981). *Nucl. Acid. Res.*, **9**, 3047-3060.
34. Gielow, L., Largen, M. & Englesberg, E. (1971). *Genetics*, **69**, 289-302.
35. Gilbert, W. & Muller-Hill, B. (1966). *Proc. Natl. Acad. Sci., USA*, **56**, 1891-1898.

REFERENCES

35. Gold, L. M. & Schweiger, M. (1971). *in Methods in Enzymology*, **20**, Moldave, K. & Grossman, L. (eds), Academic press Inc., London & New York, 537.
36. Gray, J. E., Patin, D. W. & Calhoun, D. H. (1981). *Molec. Gen. Genet.*, **183**, 428-436.
37. Guardiola, J., DeFelice, M., Lamberti, A. & Iaccarino, M. (1977). *Molec. Gen. Genet.*, **156**, 17-25.
38. Gussin, G., Johnson, A., Pabo, C. & Sauer, R. (1983). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 93-121.
39. Hahn, S., Dunn, T. & Schleif, R. (1984). *J. Mol. Biol.*, **180**, 61-72.
40. Haughn, G. W., Squires, C. H., Defelice, M., Largo, C. T. & Calvo, J. M. (1985). *J. Bacteriol.*, **163**, 186-198.
41. Hauser, C. A. & Hatfield, G. W. (1983). *Nucl. Acid. Res.*, **11**, 127-139.
42. Hawley, D. K. & McClure, W. R. (1983). *Nucl. Acid. Res.*, **11**, 2237-2255.
43. Hofnung, M. (1974). *Genetics*, **76**, 169-184.
44. Jacob, F. & Monod, J. (1961). *J. Mol. Biol.*, **3**, 318-356.
45. Johnson, A., Meyer, B. J. & Ptashne, M. (1979). *Proc. Natl. Acad. Sci., USA*, **76**, 5061-5065.
46. Keller, E. B. & Calvo, J. M. (1979). *Proc. Natl. Acad. Sci., USA*, **76**, 6186-6190.
47. Lawther, R. P. & Hatfield, G. W. (1980). *Proc. Natl. Acad. Sci., USA*, **77**, 1862-1866.
48. Lawther, R. P., Calhoun, D. H., Gray, J., Adams, C. W., Hauser, C. A. & Hatfield, G. W. (1982). *J. Bacteriol.*, **149**, 294-298.
49. Lee, N. L., Wilcox, G., Gielow, W., Arnold, J., Cleary, P. & Englesberg, E. (1974). *Proc. Natl. Acad. Sci., USA*, **71**, 634-638.
50. Lee, N. (1978). *in "The operon"*, eds. Miller, J. & Reznikoff, W. Cold Spring Harbor Laboratory, Cold spring Harbor, NY, 389-409.

REFERENCES

51. Lee, N. L., Gielow, W. O. & Wallace, R. G. (1981). *Proc. Natl. Acad. Sci., USA*, **78**, 752-756.
52. Lee-Peng, F., Hermodson, M. A. & Kohlhaw, G. B. (1979). *J. Bacteriol.*, **139**, 339-345.
53. Makman, R. S. & Sutherland, E. W. (1965). *J. Biol. Chem.*, **240**, 1309.
54. Mandel, M. & Higa, A. (1970). *J. Mol. Biol.*, **53**, 154.
55. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular cloning. A laboratory Manual*. Cold spring Horbor, Laboratory press. New-York.
56. McAllister, W. T., Morris, C., Rosenberg, A. H. & Studier, F. W. (1981). *J. Mol. Biol.*, **153**, 527-544.
57. McCorkle, G. M., Leathers, T. D. & Umbarger, H. E. (1978). *Proc. Natl. Acad. Sci., USA*, **75**, 89-93.
58. Meyer, B. J., Maurer, R. & Ptashne, M. (1980). *J. Mol. Biol.*, **139**, 163-194.
59. Meyer, B. j. & Ptashne, M. (1980). *J. Mol. Biol.*, **139**, 195-205.
60. Miozzari, G. F. & Yanofsky, C. (1978). *J. Bacteriol.*, **133**, 1457-1466.
61. Musso, R. E., Dilauro, R., Adhya, S. & Decrombrugghe, B. (1977). *Cell*, **12**, 847-854.
62. Myers, J. W. & Adelberg, E. A. (1954). *Proc. Natl. Acad. Sci., USA*, **40**, 493.
63. Nirenberg, M. & Leder, P. (1964). *Science*, **145**, 1399.
64. Ogden, S., Haggerty, D., Stoner, C. M., Kolodrubetz, D. & Schleif, R. (1980). *Proc. Natl. Acad. Sci., USA*, **77**, 3343-3350.
65. Peterson, E. A. & Sober, H. A. (1956). *J. Amer. Chem. Soc.*, **8**, **7**, 756.
66. Porath, J. & Foldin, P. (1959). *Nature*, **83**, 1657-1659.
67. Pratt, J. M., Holland, I. B. & Spratt, B. G. (1981). *Nature*, **293**, 307.
68. Ptashne, M., Jeffery, A., Johnson, A. D., Maurer, R., Meyer, B. J., Pabo, C. O., Roberts, T. M. & Sauer, R. T. (1980). *Cell*, **19**, 1-11.

REFERENCES

69. Ramakrishnan, T. & Adelberg, E. (1965). *J. Bacteriol.*, **89**, 654- 664.
70. Riggs, A. D., Bourgeois, S., Newby, R. F. & Cohn, M. (1968). *J. Mol. Biol.*, **34**, 365-368.
71. Riggs, A. D., Suzuki, H. & Bourgeois, S. (1970). *J. Mol. Biol.*, **48**, 67-83.
72. Rudman, D. & Meister, A. (1953). *J. Biol. Chem.*, **200**, 591.
73. Sancar, A., Wharton, R. P., Seltzer, S., Kacinski, B. M., Clarke, N. D. & Rupp, W. D. (1981). *J. Mol. Biol.*, **148**, 45.
74. Sanger, F., Nicklen, S. & Coulson, A. R. (1977). *Proc. Natl. Acad. Sci., USA*, **74**, 5463-5467.
75. Schwartz, M. (1967a). *Ann. Inst. Pasteur*, **112**, 673-700.
76. Shapira, S. K., Chou, J., Richaud, F. V. & Casadaban, M. J. (1983). *Gene*, **25**, 71-82.
77. Smith, G. R. & Magasanik, B. (1971). *Proc. Natl. Acad. Sci., USA*, **68**, 1493-1497.
78. Smith, J. M. & Umbarger, H. E. (1977). *J. Bacteriol.*, **132**, 870-875.
79. Soberon, X. (1980). *Gene*, **9**, 287.
80. Straiger, P., Richaud, F., Borne, F. & Patte, J. C. (1983a). *J. Mol. Biol.*, **168**, 307-320.
81. Straiger, P., Danos, O. & Patte, J. C. (1983b). *J. Mol. Biol.*, **168**, 321-331.
82. Straiger, P., & Patte, J. C. (1983). *J. Mol. Biol.*, **168**, 333-349.
83. Strassman, M., Thomas, A. J. & Weinhouse, J. (1953). *J. Amer. Chem. Soc.*, **75**, 5135.
84. Strassman, M., Thomas, A. J., Locker, L. A. & Weinhouse, J. (1954). *J. Amer. Chem. Soc.*, **76**, 4241.
85. Studier, F. W. & Moffatt, B. (1986). *J. Mol. Biol.*, **189**, 113-130.
86. Subrahmanyam, C. S., McCorkle, G. M. & Umbarger, H. E. (1980). *J. Bacteriol.*, **142**, 547-555.

REFERENCES

87. Sutton, A. & Freundlich, M. (1980). *Molec. Gen. Genet.*, **178**, 179-183.
88. Tabor, S. & Richardson, C. (1985). *Proc. Natl. Acad. Sci., USA*, **32**, 1074-1078.
89. Tomlinson, R. V. & Tener, A. M. (1963). *Biochemistry*, **2**, 697.
90. Umbarger, H. E. (1956). *Science*, **123**, 848.
91. Umbarger, H. E., Brown, B. & Eyring, E. (1960). *J. Biol. Chem.*, **235**, 1425.
92. Uzan, M., Favre, R., Gallay, E. & Caro, L. (1981). *Molec. Gen. Genet.*, **182**, 462-470.
93. Vieira, J. & Messing, J. (1982). *Gene*, **19**, 259.
94. Watson, M. D., Wild, J. & Umbarger, H. E. (1979). *J. Bacteriol.*, **139**, 1014-1020.
95. Watson, M. D. (1981). *Attenuation: translational control of transcription termination. Trends in Biochemical Sciences*, (1981), **6**, 180-182.
96. Wek, R. C. & Hatfield, W. (1986). *J. Biol. Chem.*, **261**, 2441-2450.
97. Wilcox, G., Boulter, J. & Lee, N. (1974). *Proc. Natl. Acad. Sci., USA*, **71**, 3635-3639.
98. Wild, J., Smith, J. M. & Umbarger, H. E. (1977). *J. Bacteriol.*, **132**, 876-883.
99. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). *Gene*, **33**, 103.
100. Yanofsky, C., Zurawski, G., Elseviers, D. & Stauffer, G. V. (1978). *Proc. Natl. Acad. Sci., USA*, **75**, 5988-5992.
101. Yarus, M. & Berg, P. (1967). *J. Mol. Biol.*, **28**, 479.
102. Zerbib, D., Jakowee, M., Prentki, P., Galas, D. J. & Chandler, M. (1987). *EMBO J.*, **6**, 3163-3169.
103. Zubay, G. (1973). *Ann. Rev. Genet.*, **7**, 267-287.

