# STUDIES TOWARDS THE DEVELOPMENT OF A HOST:VECTOR SYSTEM FOR THE GENUS *CLOSTRIDIUM*

By

JOHN K. BREHM, B.Sc., M.Sc.

This thesis is presented for the degree of Doctor of Philosophy in the Department of Biological Sciences, University of Warwick.
April, 1989.

## DEDICATION

To my mother and father

## DECLARATION

The work contained in this thesis was the result of original research conducted by myself under the supervision of Dr. N.P. Minton and Dr. G.P.C. Salmond. All sources of information have been specifically acknowledged by means of reference.

None of the work contained in this thesis has been used in any previous application for a degree.

CONTENTS

SECTION II: DEVELOPMENT OF A HOST/VECTOR SYSTEM FOR THE SACCHAROLYTIC CLOSTRIDIA



**CHAPTER TWO MATERIALS AND METHODS**

## LIST OF FIGURES

## LIST OF TABLES

## ABBREVIATIONS

| | |
|---|---|
| A | Adenine |
| aa | Amino acid |
| Ap | Ampicillin |
| ATP | Adenosine 5'-triphosphate |
| Av | Average |
| bp | Base pairs |
| C | Cytidine |
| c. | Approximately |
| CIP | Calf intestinal phosphatase |
| ccc | Covalently closed circular |
| Cl | Clindamycin |
| cm | Centimetre |
| Cm | Chloramphenicol |
| cpm | Counts per minute |
| C23O | Catechol 2,3-dioxegenase |
| Da | Dalton |
| dATP | Deoxyadenosine 5'-triphosphate |
| dCTP | Deoxycytosine 5'-triphosphate |
| dGTP | Deoxyguanosine 5'-triphosphate |
| dTTP | Deoxythymine 5'-triphosphate |
| dNTPs | Deoxyribonucleoside 5'-triphosphates |
| DMF | Dimethylformamide |
| DNA | Deoxyribonucleic acid |
| DNAase | Deoxyribonuclease |
| DTT | DL-Dithiothreitol |
| EDTA | Ethylenediaminetetraacetic acid |
| Em | Erythromycin |
| g | Gram |
| G | Guanine |
| fmet | Formylmethionine |
| IPTG | Isopropylthio-$\beta$-galactoside |
| kb | Kilobase |
| kDa | Kilodaltons |
| Km | Kanamycin |
| l | Litre |
| mA | Milliamperes |
| MLS | Macrolide, Lincosamide, Streptogrammin B |
| Md | Megadalton |

| | |
|---|---|
| mg | Milligrams |
| min | Minute |
| ml | Millilitres |
| mM | Millimolar |
| $M_r$ | Molecular weight |
| mRNA | Messenger RNA |
| NC | Nitrocellulose |
| nt | Nucleotide |
| nt pos | Nucleotide position |
| $OD_n$ | Optical density at wavelength n |
| ORF | Open reading frame |
| PEG | Polyethylene glycol |
| po | Promoter/operator region |
| p.s.i. | Pounds per square inch |
| ($^r$) | Resistant |
| RCM | Reinforced clostridial medium |
| RNA | Ribonucleic acid |
| RF | Replicative form |
| rRNA | Ribosomal RNA |
| RNAase | Ribonuclease |
| r.p.m. | Revolutions per minute |
| ($^s$) | Sensitive |
| SDS | Sodium dodecyl sulphate |
| sec | Second |
| Sm | Streptomycin |
| Spp. | Species (pleural) |
| SSB | Single strand binding protein |
| T | Thymine |
| Tc | Tetracycline |
| TEMED | -N,N.N',N'-tetramethylethylenediamine |
| Temp | Temperature |
| U | Uracil |
| UV | Ultraviolet |
| V | Volts |
| u | Unit |
| $\mu$g | Microgram |
| $\mu$l | Microlitre |
| wt | Wild type |
| X-gal | 5-bromo-4-chloro-3-indolyl-$\beta$-D-galactopyranoside |

## ACKNOWLEDGEMENTS

I should like to thank my supervisors Dr. Nigel Minton and Dr. George Salmond for their invaluable advice and enthusiastic encouragement throughout the course of this work. In addition, thanks are also due to Dr. John Oultram, Dr. John Stephenson, and Professor Tony Atkinson for critical reading of the manuscript.

Special thanks are due to Professor Tony Atkinson for providing facilities and arranging the finance that made this study possible.

I would also like to thank Mrs Kate Oultram for her enthusiastic assistance in the preparation of figures.

Thanks are also due to Roy Hartwell for oligonucleotide synthesis.

I gratefully acknowledge receipt of the the volume "Clostridia" from Dr. Nigel Minton and Dr. David Clarke (eds.) prior to publication.

## SUMMARY

Clostridia are widely recognised as organisms of biotechnological importance. This potential, however, cannot be fully exploited until reliable methods have been developed for the transfer of genetic information into and between members of the genus. As with other Gram-positive bacteria, attempts to develop host:vector systems have focused on protoplast transformation procedures. This approach, however, has been hampered by the lack of suitable plasmid vectors. Studies have been initiated to construct such a vector.

As a potential source of a clostridial replicon, the cryptic plasmid of C. butyricum NCIB 7423 (pCB101, 6.05 kb) has been examined. The complete nucleotide sequence of pCB101 was determined and it's minimal replicon characterised together with the elucidation of some of the mechanisms involved in it's replication.

The erythromycin resistance determinant ($Em^R$) of the Gram-positive R-factor pAMβ1 was chosen as a selectable marker and the entire nucleotide sequence of this gene was determined. The Ferredoxin (Fd) gene of Clostridium pasteurianum was cloned and employed in the construction of an expression cartridge. The effectiveness of this cartridge in promoting the expression of heterologous genes was examined using a promoter-less xylE gene.

The conjugative Streptococcus faecalis plasmid pAMβ1 was used to mobilise vectors carrying the xylE gene (under the transcriptional control of the Fd promoter) into Clostridium acetobutylicum where the gene was expressed. This represents one of the first reports of directed gene expression in C. acetobutylicum.

# CHAPTER ONE

# GENERAL INTRODUCTION

# SECTION I

# THE GENUS CLOSTRIDIUM

## 1.1. INTRODUCTION TO THE GENUS

First described by Prazmowski (1880), the genus Clostridium has grown to one of the largest genera amongst procaryotes. A total of 83 species are listed in Bergey's Manual of Systemic Bacteriology (Cato et al., 1986). For inclusion in the genus, the following criteria must be met: "anaerobic or microaerophilic spore-forming rods that do not form spores in the presence of air, are usually Gram-positive, and do not carry out dissimilatory sulphate reduction" (Cato & Stackebrandt, 1989). As a result of these rather loose requirements the genus comprises a very diverse collection of Gram-positive bacteria with a wide range of mol% G+C values, 21 to 54 %, (Gottschalk et al.. 1981), which is reflected by the highly diverse mechanisms of dissimilatory metabolism encountered. This latter feature lends a probable basis for the taxonomic subdivision of the genus; the group will be presented here primarily in terms of these properties.

### 1.1.1. THE BUTYRIC ACID BACTERIA

Many clostridia perform a fermentation of soluble carbohydrates, starch, or pectin, with the formation of acetic and butyric acids, $CO_2$, and $H_2$. These "butyric acid bacteria" grow poorly (if at all) in complex media devoid of a fermentable carbohydrate. Two other characteristics are distinctive of this subgroup: they synthesise as cellular reserve material a starch-like polysaccharide, granulose; and many fix atmospheric nitrogen. Examples of this group include C. butyricum, C. butylicum. C. beijerinckii. C. pasteurianum, C. acetobutylicum. C. aurantibutyricum.

and C. saccharobutylpropylicum (Cummins & Johnson, 1971).

Other members of this group produce neutral solvents (such as butanol, acetone, isopropranol, and ethanol from sugars. The fermentative production of acetone and butanol has been exploited on an industrial level and the strains generally employed are classified as C. acetobutylicum. although numerous specific names have been applied to these organisms in the past (Beesch, 1952; McCutchan & Hickey, 1954; Ross, 1961]. In addition, a number of different species of butanol-producing clostridia are now recognised, based mainly on differences in the type and ratio of the solvents produced. C. beijerinkii (C. butylicum) produces solvents in approximately the same ratio as C. acetobutylicum. but isopropranol is produced in place of acetone, while C. aurantibutyricum produces both acetone and isopropranol in addition to butanol (George & Chen, 1983). C. tetanomorphum is a newly isolated species which produces almost equimolar amounts of butanol and ethanol but no other solvents (Gottwald et al.. 1984).

### 1.1.2. THE AMINO ACID DISSIMILATING BACTERIA

A large number of Clostridium spp. can grow well in complex media containing peptones or yeast extract, in the absence of a fermentable carbohydrate. These organisms are collectively responsible for the putrefactive decomposition of nitrogenous compounds in nature, and include the principal pathogenic clostridia (C. botulinum. C. tetani. and C. perfringens). Growth in complex media is accompanied by the formation of ammonia, $CO_2$, $H_2S$, fatty acids, and a variety of other volatile substances, often having unpleasant odours (for review, see Barker, 1961). Although a considerable number of clostridia that ferment amino acids can also ferment carbohydrates, by a typical butyric acid

fermentation, many are wholly unable to ferment carbohydrates; such organisms are exemplified by C. tetani and C. histolyticum. Similarly, although the majority of the amino acid dissimilating clostridia produce proteases, certain species are non-proteolytic. Organisms of this latter type are therefore dependent on the availability of free amino acids as growth substrates, as opposed to proteinaceous material.

### 1.1.3. THE NITROGEN-CONTAINING RING COMPOUND FERMENTING CLOSTRIDIA

Some clostridia can obtain energy by the fermentation of ring compounds including purines, pyrimidines, and nicotinic acid. The fermentation of purines (guanine, uric acid, hypoxanthine, xanthine) is carried out by C. acidiurici and C. cylindrosporum. nutritionally highly specialised species, which are unable to ferment other substrates. The fermentation products consist of acetate, glycine, formate, $CO_2$, and other precursors.

### 1.1.4. THE CARBOHYDRATE FERMENTING CLOSTRIDIA THAT DO NOT YIELD BUTYRATE

A number of clostridia utilising carbohydrates as energy sources dissimilate them by pathways other than the butyric acid pathway. These organisms include cellulose fermenting clostridia, most of which are highly specialised with respect to substrates; some species can ferment only cellulose. The products include ethanol, formate, acetate, lactate, and succinate. The species C. thermoaceticum ferments glucose and other soluble sugars with the formation of acetate as the sole end-product.

### 1.1.5. THE ETHANOL-ACETATE FERMENTING CLOSTRIDIA

The most remarkable of all the clostridial fermentations is that performed by C. kluyveri. This organism grows only at the expense of ethanol and acetate as it's energy sources. The main organic products of the fermentation are two higher fatty acids, butyrate and caproate; in addition, some $H_2$ is produced.

## 1.2. INDUSTRIAL APPLICATIONS OF CLOSTRIDIA

A number of commercial applications for Clostridium spp. have been investigated. These include hydrogen production by C. butyricum (Karube et al., 1981; Suzuki & Karube, 1980), the production of organic acids (i.e., acetic, acrylic, butyric, fumaric, propionic and succinic acid) by acetogenic and acidogenic clostridia (see Wiegel & Ljungdhal, 1986) and the potential of certain species as biocatalysts in the synthesis of chiral compounds (see Simon et al., 1985; Morris, 1989). Other clostridia are proving to be of medical value. Thus the discovery of specific oncolytic activity elaborated by some strains of C. butyricum has led to their use in tumour treatment (Schlechte et al., 1982), while the neurotoxin of C. botulinum is now finding use as a therapeutic drug in the treatment of certain aberrant muscular functions (see Hambleton et al.. 1987). However, the most successful industrial application of a clostridial-based process is still the production of solvents.

### 1.2.1. SOLVENTOGENIC FERMENTATION

The production of butanol in a microbial fermentation was first reported by Pasteur in 1861. During the latter part of the 19$^{th}$ century the production of butanol by anaerobic

bacteria was studied by a number of investigators; later, in 1905 the production of acetone was reported by Schardinger (for reviews, see: McCutchan & Hickey, 1954; Zeikus, 1980; Jones & Woods, 1986). Interest in commercialising butanol production was initiated in 1909 as a means to obtain butadiene, used in the synthesis of rubber. In 1912 Weizmann isolated an organism that was capable of fermenting starch to acetone, butanol, and ethanol. This organism was named C. acetobutylicum strain Weizmann (McCoy et al., 1926). The production of acetone was of great interest during World War I as a chemical feedstock used for the manufacture of explosives (i.e., cordite) and airplane wing dope. The demands for acetone decreased dramatically after the war, but the fermentation process continued as other needs for acetone and butanol developed. The acetone-butanol fermentation process ceased to operate in the 1960's in the U.S. because of unfavourable ecomomics caused by chemical synthesis of these products from petroleum derived feedstocks. However, until 1983 acetone and butanol were still produced from corn starch and molasses in South Africa using C. acetobutylicum. This was a consequence of an abundant feedstock supply combined with the absence of a readily available source of cheap petroleum. However, subsequent shortages of molasses caused the plant to shut down.

The potential of optimising the acetone/butanol fermentation has been recognised (Gibbs, 1983) and has played an important part since Weizmann's early strain selection and the early preference for, and hence selection of, strains of C. acetobutylicum which are both highly solvantogenic and highly solvent resistant (Jones et al., 1982; Jones & Woods, 1986; Long et al., 1984).

The acetone/butanol fermentation has received renewed interest in Japan (Ogata & Hongo, 1979). Studies with a newly isolated strain, named C. saccharoperbutylacetonicum, which produces higher levels of butanol than does C. acetobutylicum, are now underway. C. madisonii has also been employed in industrial fermentations.

### 1.2.2. FUTURE POTENTIAL

The industrial uses of clostridia cited in the preceding section are based upon existing physiological properties of the strains involved. The creation of new and unique physiological strain types by mutation or gene manipulation would be of great value to the fermentation industry. This could be performed either by mutation and selection experiments to identify improved strains for existing processes (Gibbs, 1983) or by the introduction of advantageous genes into suitable clostridial hosts by gene transfer experiments. The former strategy has been used with some success to date, particularly with C. acetobutylicum. Examples include auxotrophic mutants (Bowring & Morris, 1985; Jones et al., 1985), antibiotic resistance mutants (Bowring & Morris, 1985; Long et al., 1984), granulose mutants (Jones et al., 1982), capsule mutants (Jones et al., 1982), phage resistant mutants (Ogata & Hongo, 1979), and sporulation mutants (Jones et al., 1985; Long et al., 1984).

The alternative method employing gene transfer technology would enable the use of clostridia of interest as hosts for foreign genes, so that appropriately prepared DNA from any source could be introduced into them. Such a cloning facility would therefore be useful in strain improvement, allowing greater flexibility than is possible by mutation alone, and presenting the possibility of introducing biologically active molecules (e.g., enzymes) into a new host. Such a

cloning facility would also make it possible to introduce genes into clostridial hosts, whose products are of value in their own right, lending itself to the production of microbial enzymes and proteins.

A fact largely overlooked is that clostridia are potentially ideal fermentation organisms as the major costs of industrial fermentation (i.e., cooling and aeration) would be virtually negated by using thermophilic clostridia. Furthermore, while clostridia naturally ferment a wide variety of carbon sources, their nitrogen requirements can also be highly cost-effective as certain strains are capable of excellent growth on minimal media containing only inorganic nitrogen. An additional virtue of clostridia is that many species produce an extracellular cellulolytic complex comprised of a multitude of endoglucanases (Bisaria & Ghose, 1981) rendering them capable of hydrolysing insoluble native cellulose, in addition to soluble cellulose derivatives such as carboxymethylcellulose. This therefore gives potential for the use of low cost feedstocks such as industrial wastes which would otherwise be unsuitable substrates (Gibbs, 1983). Also, many of the cellulolytic clostridia, which are capable of converting cellulosic substrates to organic solvents, are thermophilic e.g., C. thermocellum (Zeikus, 1980). The use of these organisms as fermentation hosts would permit an elevated rate of fermentation and protein production without a requirement for cooling as is the case with mesophilic aerobic fermentations. Furthermore, such thermophilic fermentations would enhance the vaporation, and hence easy recovery, of volatile solvents (T. Atkinson, personal communication).

The clostridia are also able to secrete a variety of proteins into the external medium. Examples include the cellulolytic complexes (Bisaria & Ghose, 1981), the thermostable

beta-amylase from C. thermosulphurogenes (Hyun & Zeikus, 1985), the xylanase of C. acetobutylicum (Lee & Forsberg, 1987), the inulase of C. acetobutylicum (Efstathiou et al., 1986), and the thermostable alpha-amylase and pullulanase of C. thermohydrosulphuricum (Antranikian et al., 1987). This naturally occurring property may be exploited if the correct cloning procedures can be developed (e.g., fusion of a suitable secretory signal to the amino terminus of the protein to be secreted). In strains which produce few secreted products (or non-producing mutant strains) and which utilise low-protein or protein-free substrates, the isolation of the desired product from culture effluents should be simple. The use of a clostridial strain which does not produce external proteases would allow maximal recovery of secreted protein products during fermentation.

## 1.3. THE GENETICS OF CLOSTRIDIUM

Until very recently there were virtually no genetic studies with the clostridia. Current understanding of the genetics of Gram-positive bacteria is largely restricted to B. subtilis, although genetic studies have been initiated with other commercially significant Gram-positive bacteria such as certain Streptomyces spp., the lactic streptococci, and certain Lactobacillus spp. More recently, however, genetical studies have commenced with the clostridia; initially with the medically important members of the genus (i.e., C. perfringens, C. tetani), and more recently with the industrially important saccharolytic clostridia (i.e., C. acetobutylicum, C. thermocellum, etc.). The primary driving force behind this change has been an increase in awareness of the potential of the exploitation of these species in industrial biotechnology (for reviews, see Jones & Woods, 1986; Minton & Thompson, 1989; Rogers, 1986; Snedecor & Gomez, 1983; Walker, 1983).

### 1.3.1. CLOSTRIDIAL PLASMIDS

Screening of various Clostridium spp. for extrachromosomal DNA has revealed a ubiquitous distribution of plasmid DNA. In the vast majority of cases, however, the plasmids identified have proved to be cryptic (for reviews, see Minton and Thompson, 1989; Rogers, 1986). In certain instances, however, plasmid encoded functions have been assigned. In this respect current knowledge of the plasmids of C. perfringens is far superior than that for any other clostridia. This is probably the direct consequence of the longer research history devoted to this species and the discovery, at a very early stage, of selectable genetic markers on a number of plasmids found in this species. Even so, plasmids have been isolated from other pathogenic clostridia and in some cases linked to toxin production or some other marker.

#### 1.3.1.1. BACTERIOCINOGENIC PLASMIDS

The earliest report of a clostridial plasmid came from Ionesco and Bouanchaud (1973) who correlated the presence of a 5.7 Md (8.7 kb) plasmid in a strain of C. perfringens with the production of a UV light inducible bacteriocin (BC5). This plasmid, pIP404, was later demonstrated to also encode for immunity to BC5 and to be transferred to BC5-sensitive strains by a process requiring cell-to-cell contact (Brefort et al.. 1977). These authors were unable to ascertain whether a larger co-resident cryptic plasmid (pIP405, 32.4 Md, 49.6 kb) was responsible for the observed mobilisation of pIP404. Similar plasmids have been identified in two other bacteriocinogenic C. perfringens strains (Mihlac et al.. 1978; Li et al.. 1980). The complete nucleotide sequence of pIP404 has recently been reported (Garnier & Cole, 1988a), and functions assigned to six of the ten open reading frames identified (Garnier &

Cole, 1986, 1988a, 1988b, 1988c; Garnier et al., 1987). This represents the first report of a complete nucleotide sequence for a clostridial plasmid.

### 1.3.1.2. CONJUGAL PLASMIDS

The presence of conjugal R-factors in C. perfringens appears to be widespread, and in particular plasmids encoding resistance ($^{r}$) to tetracycline (Tc). Such a plasmid, pIP401 (54 kb), was first described by Sebald and co-workers (Sebald et al., 1975; Brefort et al., 1977). A similar plasmid, pCW3 (30.6 Md, 46.5 kb) was also isolated by Rood et al. (1978). More recent studies by this group (Rood, 1983; Abraham & Rood, 1985a) have involved the restriction enzyme analysis of a number of conjugative tetracycline resistance plasmids isolated from C.perfringens from a number of porcine origins. These workers reported that many of the isolates contained a 29 kb DNA region in common with pCW3, and proposed that many conjugative resistance plasmids of C. perfringens may contain a pCW3-like core. Abraham & Rood (1985b) further characterised pCW3 by mapping and cloning the inducible tetracycline resistance determinant of the plasmid.

The plasmid pIP401 isolated by Sebald et al. (1975) also encoded resistance to chloramphenicol (Cm), in addition to tetracycline (Tc). During in vivo conjugal transfer experiments with pIP401 ($Tc^{r}$, $Cm^{r}$) Brefort et al. (1977) observed frequent segregation of the two resistance markers with the concomitant loss of 6 kb of DNA and the $Cm^{r}$ phenotype, yielding the derivative, pIP406. It is noteworthy that the size of pIP406 was reported to be similar to that determined for pCW3 (Rood et al., 1978). The deletion in pIP401 which yielded pIP406 was later mapped to lie within a 10.55 kb EcoRI fragment during restriction endonuclease

mapping studies on pIP401 (Magot, 1984). The suggestion that this $Cm^r$ element resides on a transposable element (Magot, 1984) has now been confirmed by Abraham & Rood (1987).

Another pathogenic species shown to carry a conjugal plasmid is C. cochlearium. Pan-Hou et al. (1980) reported that a plasmid of undetermined size was responsible for organomercury resistance, more specifically for methyl mercury decomposition with the concomitant generation of hydrogen sulphide and inorganic mercury. The observed organo-mercury resistance was lost with curing of one of the two plasmids resident in the strain used. Interestingly, the cured derivative strains were capable of methylating inorganic mercury with the formation of it's methyl derivative. The plasmid was also found to be transmissible, by a conjugation like process, to a mutant derivative of the cured wild type strain. This recipient strain retained the second, cryptic, plasmid from the original strain, and hence the conjugative activity could not be unequivocally assigned to either plasmid.

### 1.3.1.3. TOXIGENIC PLASMIDS OF C. PERFRINGENS

A plasmid borne location for the genes encoding toxin production in C. perfringens has been obtained in two separate studies. An examination of 22 strains of five toxigenic types of C. perfringens A - E (Rokos et al., 1978; Duncan et al.. 1978) revealed plasmids (up to nine in a single isolate) in 18 strains. The possession of a 75 Md (114 kb) plasmid was correlated to the elaboration of beta toxin in a type C strain; a lethal dermonecrotic toxin involved in diarrheal disease, mainly in domestic animals. In another study Blaschek & Solberg (1981), demonstrated that loss of caseinase activity by C. perfringens ATCC3626

(elaboration of lambda toxin) accompanied the loss of a small 2.1 Md (3.2 kb) plasmid, pHB101. This strain, which also harbours a larger 9.4 Md (14.3 kb) cryptic plasmid, pHB102, did not appear to donate pHB101 to other strains by a conjugation-like process. These workers also reported that, concomitant with the loss of pHB101 and caseinase activity, the cells underwent a morphology change from a rod shape to a coccoid or bacillary-coccoid form (Solberg et al., 1981).

### 1.3.1.4. TOXIGENIC PLASMIDS OF C. TETANI

In C. tetani, a plasmid has been firmly implicated in the production of tetanus toxin. Early studies by Laird et al. (1980) indicated a close association between toxigenicity and the presence of a large plasmid element. In a subsequent study (Finn et al., 1984) this same group of workers demonstrated that a pool of synthetic oligonucleotides (based on the amino-terminal amino acid sequence of the tetanus toxin) specifically hybridised to plasmids derived from three toxigenic strains in the absence of hybridisation to plasmid DNA derived from non-toxigenic strains. One of the toxigenic plasmids identified (pCL1) was observed to spontaneously delete some 22 kb DNA resulting in a non-toxigenic phenotype. This derivative (pCL2), however, was still hybridisation positive suggesting that not all the toxin coding sequence had been deleted. More recently, the tetanus toxin gene has been cloned in E. coli and its complete nucleotide sequence determined (Fairweather et al., 1986; Fairweather & Lyness, 1986).

### 1.3.1.5. PLASMIDS OF C. BOTULINUM

Evidence for the presence of plasmids in C. botulinum was first presented by Scott & Duncan (1978). They reported the

presence of three 2-5 Md (3.0-7.6 kb) plasmids in E-like bacteriocinogenic (boticinogenic) strains of C. botulinum. and suggested their involvement in the production of boticin E. However, as no curing experiments were performed, plasmid-mediated production of boticin E has yet to be established. Other researchers (Strom et al.. 1984; Weickert et al.. 1986) have screened a large number of toxigenic C. botulinum strains (types A through G) and C. botulinum-like non-toxigenic strains (C. sporogenes and C. subterminale) for plasmids that might be involved in neurotoxin production. Strains of non-toxigenic C. sporogenes were focused on for comparative studies as this organism is thought to be a non-toxigenic counterpart of C. botulinum type A (P. Hambleton, personal communication). Many plasmids were identified in both toxigenic and non-toxigenic strains, ranging in size from 2.1 to 81 Md (3.2-123 kb) (Strom et al.. 1984), but no phenotypic functions could be assigned. Weickert et al. (1986) further failed to correlate toxigenicity with the presence of any one plasmid, and also demonstrated the production of toxin by a type A strain which had been cured of plasmids.

### 1.3.1.6. PLASMIDS OF C. DIFFICILE AND C. NOVYI

The apparent linkage between antibiotic-associated pseudomembranous colitis (Bartlett et al.. 1980), a severe diarrheal disease, and C. difficile prompted Muldrow et al. (1982) to screen a number of clinical isolates of C. difficile for the presence of plasmids that might be responsible for the observed toxigenicity. However, no correlation was found between plasmid content and either cytopathogenic effect or resistance to the various antimicrobial agents tested. Again, no plasmid species could be correlated with these resistances. A similar conclusion was reached in separate studies by other workers (Arai et al.. 1984;

Hayter & Dale, 1984).

Another pathogenic species, C. novyi type A, elaborates a lethal and necrotising alpha-toxin, the production of which is dependent on the presence of a specific temperate bacteriophage ($NA1^{tox+}$). Schallehn & Kramer (1981) examined the plasmid content of two toxigenic and two non-toxigenic strains of C. novyi (type A) and reported that all four strains contained a number of plasmids. The toxigenic strains appeared to contain a greater proportion of smaller plasmids (2.5-5 Md, 3.8-7.6 kb) than the non-toxigenic strains which were unique in that they contained some larger plasmid species. This observation led Schallehn & Kramer (1981) to suggest a tentative role for the smaller plasmids in toxigenicity.

### 1.3.1.7. PLASMIDS OF THE SACCHAROLYTIC CLOSTRIDIA

In more recent years, with the increasing awareness of the biotechnological importance of the non-pathogenic saccharolytic clostridia, certain laboratories have initiated studies toward the isolation and characterisation of plasmids from various species. All the plasmids isolated to date have proved to be cryptic (for reviews, see Minton & Thompson, 1989; Rogers, 1986). Truffaut and Sebald (1983) screened 21 strains of C. acetobutylicum. C. butylicum. and C. saccharoperbutylacetonicum (acetone-butanol producing strains) for the presence of extrachromosomal DNA molecules. Seven of the strains examined contained plasmids which ranged in size from 2.6 to more than 50 Md (3.9-76.0). However, no attempt was made to correlate these plasmids with any phenotypic trait.

Minton & Morris (1981) were the first to demonstrate the presence of cryptic plasmids in four out of seven strains

of C. butyricum. Three of the strains (SA1, SA2, and NCTC 6084) contained an apparently identical plasmid, designated pCB103 (4.3 Md, 6.2 kb), on the basis of restriction endonuclease analysis. The fourth strain, a bacteriocinogenic strain (C. butyricum NCIB 7423) contained two plasmids, designated pCB101 (3.9 Md, 5.6 kb) and pCB102 (5.2 Md, 7.4 kb). These were differentiated by restriction endonuclease analysis. However, butyricin (a bacteriocin) production could not be unequivocally attributed to the possession of either or both of the two plasmids. More recent studies by Luczak et al. (1985) with strains of C. butyricum have demonstrated the presence of the same plasmids reported by Minton & Morris (1981); pCBU1, 6.4 kb (pCB103), pCBU2, 6.3 kb (pCB101), and pCBU3, 8.4 kb (pCB102). However, there seems to be some confusion as to the exact plasmids ascribed to the bacteriocinogenic host C. butyricum NCIB 7423; Luczak et al. (1985) reported that this strain harboured the single plasmid species, pCBU1 (pCB103), whereas Minton & Morris (1981) reported that this strain harboured both pCB101 (pCBU2) and pCB102 (pCBU3). Additional studies by the former group (Collins et al., 1985), with pCB101 and pCB102, have identified restriction fragments of each plasmid that promote the establishment of a Gram-positive replication-deficient plasmid (pJAB1) in B. subtilis. A 3.3 kb Sau3A fragment of pCB101 conferred upon the vector the ability to transform both $Rec^+$ and $Rec^-$ strains of B. subtilis. A recombinant plasmid carrying a 2 kb Sau3A fragment of pCB102 underwent integration into the B. subtilis chromosome.

Urano et al. (1983) have reported the presence of both small and large plasmids (pSSK1, 51 Md, 77.5 kb; pSSK2, 32 Md, 48.6 kb; pSSK3, 9.4 Md, 14 kb) in the hydrogen-evolving bacterium C. butyricum IFO 3847. They further partially characterised these plasmids by restriction endonuclease

analysis and reported that the two larger plasmids (pSSK1 and pSSK2) were closely related by virtue of common restriction fragments. By performing curing experiments, they also tried to establish whether the genes involved in hydrogen evolution, or any possible resistance genes, were located on the plasmids or not. They concluded that neither hydrogenase genes nor antibiotic resistance genes are located on any of the three plasmids.

Popoff & Truffaut (1985) screened some 50 C. butyricum strains from clinical and non clinical sources, and 14 C. beijerinkii strains originating from dairy products, for plasmid content, antibiotic resistance, and bacteriocinogenic activity. The incidence of antibiotic resistance and presence of plasmid DNA was more widespread amongst the C. butyricum strains from a clinical source than among the C. beijerinkii strains. In many of the C. butyricum strains, a small 4.5 Md (6.8 kb) plasmid was encountered. However, these workers were unable to establish any relationship between plasmid pattern and antibiotic resistance, geographic localisation of the isolates, or the clinical condition of patients from which some strains originated.

The most extensive plasmid screening exercise to date is that recently reported by Lee et al. (1987). These workers reported the screening of some 150 strains of non-pathogenic clostridia. However, they only discovered 26 strains, representing 21 species, that contained at least one (maximum of five) plasmid. This was considered to be an underestimate because of difficulties with cell lysis conditions and exonuclease production, which may have allowed many plasmids to escape detection. High DNAase activity associated with many clostridial species has been previously reported (Blaschek & Klacik, 1984; Luczak et al., 1985; Urano et al.. 1983) and may be responsible for

the low frequency of plasmid isolation reported by Lee et al. (1987).

### 1.3.2. CLOSTRIDIAL BACTERIOPHAGES

Cowles (1934) was the first to isolate a bacteriophage from a member of the genus Clostridium. Since then a number of bacteriophages (and defective bacteriophages) have been described (for review, see Ogata & Hongo, 1979). The vast majority of these phages have been isolated from pathogenic species, and no bacteriophage to date has proven to possess transducing activity. Lytic phage infections, however, have presented serious problems in the acetone/butanol industry (Jones & Woods, 1986; Ogata & Hongo, 1979). It was first established by McCoy et al. (1944, cited Ogata et al.. 1981) and later confirmed by Hongo (1965, cited Ogata et al., 1981) that sluggish fermentations by C. acetobutylicum were a consequence of lytic bacteriophage infection.

Some of the more interesting bacteriophages identified have been implicated in the mediation of pathogenicity, being responsible for toxin production, amongst some of the species, particularly in the case of the species C. botulinum. A series of studies have established that toxin production by both C. botulinum types C and D required the specific participation of bacteriophages (Eklund et al., 1971, 1972). Toxin serotype conversion was also possible. C. botulinum serotypes C and D (which comprise a single group) could be interconverted by the infection of a bacteriophage-cured derivative of type C or D with the bacteriophage isolated from the complementary toxigenic strain (Eklund & Poysky, 1974). Further studies by Eklund et al. (1974) demonstrated that non-toxigenic cured cells of C. botulinum type C could also be converted to another toxigenic bacterial species, C. novyi type A, after infection

with phage NA1 from C. novyi type A. Thus three immunologically distinct toxins can be produced by a common bacterial strain following infection with specific bacteriophages.

More recently, on the basis of DNA-DNA hybridisation and other physiological parameters, Nakamura et al. (1983) studied the taxanomic relationships among C. novyi Types A and B, C. haemolyticum, and C. botulinum type C. The 17 strains studied were sorted into three groups: I, C. novyi type A; II, C. novyi type B, C. haemolyticum. and one C. botulinum type C strain; III, the remaining C. botulinum type C strains. Schallehn and Eklund (1980) demonstrated that C. novyi type D (C. haemolyticum) could be converted to the production of alpha-toxin by infection with phage from C. novyi type A. This observation together with the data used to establish the three homology groups mentioned above and the phage-mediated conversion of toxin production between C. botulinum type C and C. novyi type A, would seem to indicate that the host specificities of phages involved are not restricted to a set of genetically homologous strains and that true interspecific transfer of phage is demonstrable within the genus. Similar interspecific phage transfer has also been reported for other Gram-positive bacteria such as Bacillus (Ruhfel et al.. 1984) and Streptomyces (Chater & Carter, 1979).

### 1.3.3. GENETIC TRANSFER IN THE ABSENCE OF PLASMIDS

Within Gram-positive bacteria, and in particular the streptococci, it is becoming increasingly apparent that drug resistance genes frequently form large parts of non-homologous insertions in the host chromosome. Many of these elements are capable of conjugal transfer, in the absence of detectable extrachromosomal DNA, to suitable recipient strains, and are therefore termed "conjugal transposons"

(for reviews, see Clewell & Gawron-Burke, 1986; Minton & Thompson, 1989).

The first hint that conjugal transposons occur in clostridia came from the studies of Ionesco (1980) when the conjugal transfer of low level tetracycline resistance ($Tc^r$) was demonstrated among antibiotic-resistant strains of pathogenic C. difficile in the absence of plasmid DNA. These findings were subsequently confirmed by Smith et al. (1981), who further reported that this transfer phenomenon was insensitive to DNAase and reminiscent of a transfer mechanism in Pneumococcus (Shoemaker et al., 1980). Even though the donor strains contained two plasmids, the $Tc^r$ progeny contained no detectable plasmids and were themselves capable of acting as donors. In addition, tetracycline-sensitive ($Tc^s$) colonies, which arose spontaneously from the $Tc^r$ parents at frequencies up to three per thousand, still retained the parental plasmid profile. This $Tc^r$ loss appeared to be irreversible leading these authors to propose the hypothesis that this resistance transfer is mediated by a chromosomally located transposable element which carries the $Tc^r$ determinant.

A similar transfer phenomenon was reported for clindamycin (Cl) and erythromycin (Em) resistances, occurring between C. innocuum and a range of other clostridial species, and in particular, to C. perfringens (Magot, 1983). Linkage of $Em^r$ and $Cl^r$ in transconjugants led to the suggestion that the two resistances were resident on the same transferable element. In the case of antibiotic resistance strains of C. difficile. donors were capable of transferring resistance to Cl, Em, and streptomycin (Sm) jointly, and $Tc^r$ independently, in a similar non-plasmid mediated manner (Wust & Hardegger, 1983). These authors proposed that the Cl-Em-Sm and Tc resistance determinants are part of two

separate conjugal transposons. Interestingly Hachler et al. (1987a, 1987b) have recently demonstrated that the transferable Tc and Em (ermZ) resistances of pathogenic strains of C. difficile share homology with the corresponding resistance determinants of the Streptococcus faecalis transposon Tn916 and the Staphylococcus aureus transposon Tn551 respectively.

Most significantly, the relatively well characterised streptococcal transposons Tn916 (Franke & Clewell, 1981) and Tn1545 (Courvalin & Carlier, 1987) have now been shown to be capable of undergoing conjugal transfer from S. faecalis to C. acetobutylicum (Davies et al., 1988). Since the DNA of clostridia exhibits a very low G+C content (Cato et al., 1986) and the consensus target site for insertion is AT-rich (Caillaud and Courvalin, 1987; Caillaud et al.. 1987; Senghas et al.. 1988), these elements are likely to prove extremely useful as tools for transposon mutagenesis. Indeed, recent data has shown that they both integrate at a multitude of sites within the C. acetobutylicum genome (M. Young, personal communication). Similar results have now been obtained with C. tetani (Volk et al.. 1988). Furthermore, as both transposons excise precisely in E. coli. the extremely powerful cloning strategy suggested by Gawron-Burke & Clewell (1984) may be applied to the cloning of clostridial genes.

### 1.3.4. CLONING IN CLOSTRIDIA

Several groups have embarked on genetic analysis of the clostridia, but reliable and widely applicable procedures for promoting gene transfer have, until recently, been lacking. Recent advances have focused on both natural and artificial means of gene transfer with both pathogenic and non-pathogenic species of clostridia, namely C. perfrin-

gens, C. acetobutylicum, and C. thermohydrosulphuricum. Procedures for genetic exchange employing clostridial protoplasts are now being developed in several laboratories. Procedures for the regeneration of clostridial protoplasts back to the bacillary form has been reported for C. acetobutylicum (Allcock et al., 1982; Reysset et al., 1987), C. pasteurianum (Minton & Morris, 1983), C. saccharoperbutylacetonium (Yoshino et al., 1984), C. perfringens (Heefner et al., 1984), and C.tertium (Knowlton et al., 1984). However, protoplast transformation with exogenous DNA is limited to reports of bacteriophage transfection of C. acetobutylicum (Reid et al., 1983), and plasmid transformation of C. acetobutylicum (Lin & Blaschek, 1984) and C. perfringens (Heefner et al., 1984; Squires et al., 1984).

The protoplast mediated transformation of C. perfringens, together with the development of a family of bifunctional vectors, reported by Heefner et al. (1984) and Squires et al. (1984) is particularly interesting as it required a two stage process. Initially, heterologous DNA from E. coli was used to transform stable L-form variants of C. perfringens. DNA re-extracted from these L-forms was then used to transform autoplasts which can revert to bacillary form. The two stage procedure was necessary because (a) the efficiency of transformation of autoplasts with heterologous DNA was very poor (perhaps due to restriction?), and (b) bacillary colonies could not be obtained from the stable L-form variants. These reports firmly established the protoplast transformation of C. perfringens.

The transfection process for C. acetobutylicum reported by Reid et al. (1983) showed several unusual features, one of which was a need for prolonged incubation with transfecting DNA. These authors were also unable to detect plaques in

the primary lawn of indicator cells. This same group of workers have also demonstrated the occurrence of protoplast fusion using auxotrophic derivatives of C. acetobutylicum (Jones et al., 1985), and as is the case in B. subtilis, a great variety of exfusants appear (reviewed by Hotchkiss & Gabor, 1985).

The report of protoplast transformation of C. acetobutylicum by Lin & Blaschek (1984) was the first of its type for this organism. They reported transformation with plasmid pUB110 DNA conferring kanamycin resistance to C. acetobutylicum. The procedure required large amounts of DNA and was initially dependent upon a heat-pretreatment step to inactivate the interfering nuclease activity that is associated with the protoplasts. Attempts to extend these findings to other strains of C. acetobutylicum have, however, been unsuccessful (M. Young, personal communication).

The apparent inability to obtain reliable protoplast transformation of C. acetobutylicum led Oultram and Young (1985) to explore an alternative procedure based on "natural" gene exchange. Using the broad host range streptococcal plasmid pAMβ1 ($Em^R$), Oultram & Young (1985) demonstrated its transfer from a variety of donor species to C. acetobutylicum. These findings were later confirmed and extended by the work of Reysset & Sebald (1985) and Yu & Pierce (1986), who have established that several different species of Streptococcus can act as donors of pAMβ1 and other conjugal plasmids in filter matings with C. acetobutylicum. Transfer of pAMβ1 to C. pasteurianum and C. butyricum has also been reported (Oultram, 1986), but the plasmid appears only to confer low level resistance to Em on these hosts. Oultram et al. (1987) extended these studies by developing a plasmid transfer system where small plasmids, potential cloning vehicles, could be cointegrated with the cojugal

mobilisar pAMβ1 in a B. subtilis host and subsequently transferred to C. acetobutylicum by a filter mating. More recently, using this system, Oultram et al. (1988a) have demonstrated the transfer of biosynthetic genes from C. pasteurianum to C. acetobutylicum.

Another report of successful transformation of a clostridial host is an alkaline-Tris procedure for whole cell transformation of C. thermohydrosulphuricum with pUB110 plasmid DNA (Soutschek-Bauer et al., 1985). However, this methodology is perhaps restricted to only similar bacteria which have a paracrystalline proteinaceous surface layer as other organisms sharing this characteristic have also been similarly transformed (Fornari & Kaplan, 1982; Takahashi et al., 1983).

In conclusion, although the cointegrate conjugal mobilisation procedure described by Oultram et al. (1987) appears to be somewhat cumbersome, it probably represents the best method currently available for transferring genes into C. acetobutylicum. It has the merits of being (a) reliable, (b) applicable to other species of Clostridium. and (c) amenable to further development. For example, work is currently under way for the development of mobilisable plasmids that will replicate as multicopy elements in C. acetobutylicum and others that will act as transposon delivery vehicles (M. Young, personal communication).

The final consideration for cloning in clostridia is the existence of barriers such as restriction/modification systems, the presence of extracellular nucleases, and plasmid stability. There is ample evidence for the presence of nucleases in both C. acetobutylicum and C. perfringens (Blaschek & Klacik, 1984; Lin & Blaschek, 1984). Restriction/modification systems have been found in a variety of

clostridia (Roberts, 1987). Isochisomers of MboI have been found in C. perfringens and C. pasteurianum and a restriction enzyme from C. formoaceticum that recognises the sequence GCGC is produced commercially. The site specificity of another enzyme found in C. histolyticum has not been determined. Finally, various E. coli/C. perfringens shuttle plasmids (Squires et al.. 1984), pUB110 (Lin & Blaschek, 1984), and pAMβ1 (Oultram & Young, 1985; Reysset & Sebald, 1985; Yu & Pierce; 1986) have been demonstrated to replicate in certain clostridial species. However, whether or not they will prove to be segregationaly stable has yet to be established.

### 1.3.5. EXPRESSION OF CLOSTRIDIAL GENES IN OTHER ORGANISMS

At the commencement of this study only a handful of clostridial genes had been cloned in E. coli. These were amino acid biosynthetic genes from C. thermocellum (Cornet et al.. 1983) and C. butyricum (Ishii et al.. 1983), the hydrogenase gene from C. butyricum (Karube et al.. 1983) and various cellulase genes from C. thermocellum (Cornet et al.. 1983). Since this date numerous reports have appeared in the literature on the cloning of clostridial genes (see Table 1.1). In general, E. coli has been the choice host although some genes have been cloned and expressed in B. subtilis (Efstathiou & Truffaut, 1986; Soutschek-Bauer & Staudenbauer, 1987), B. stearothermophilus (Soutschek-Bauer & Staudenbauer, 1987) and Saccharomyces cerevisiae (Sacco et al.. 1984). The expression in E. coli of many of the cloned clostridial genes has formed the basis for phenotypic selection in the cloning strategies adopted. The complete nucleotide sequences of some of these genes, and their regulatory regions, have now been determined (see Table 1.1).

TABLE 1.1

SUMMARY OF CLONED CLOSTRIDIAL GENES

| ORGANISM | GENES CLONED | SEQUENCE | REFERENCE |
|---|---|---|---|
| *C. acetobutylicum* | Alcohol dehydrogenase | - | Youngleson *et al.*, 1988 |
| | Butyraldehyde dehydrogenase | - | Contag & Rogers, 1988 |
| | Butyrate kinase | - | Cary *et al.*, 1988 |
| | Endo-α-1,4-glucanase | + | Zappe *et al.*, 1988 |
| | α-glucosidase | - | Zappe *et al.*, 1986 |
| | Glutamine synthetase (*gluA*) | + | Usdin *et al.*, 1986 |
| | | + | Janssen *et al.*, 1988 |
| | Phosphotransbutyrylase | - | Cary *et al.*, 1988 |
| | Xylanase | - | Zappe *et al.*, 1987 |
| *C. acidiurici* | Formyltetrahydrofolate synthetase | - | Whitehead & Rabinowitz., 1986 |
| *C. butyricum* | Hydrogenase | - | Karube *et al.*, 1983 |
| | α-isopropylmalate dehydrogenase | - | Ishii *et al.*, 1983 |
| | 2 chloramphenicol acetyltransferases | - | Dubbert *et al.*, 1988 |
| *C. cellulolyticum* | 2 endo-α-1,4-glucanases | - | Faure *et al.*, 1988 |
| *C. difficile* | Chloramphenicol acetyltransferase | - | Wren *et al.*, 1988 |
| | Enterotoxin A | - | Wren *et al.*, 1987 |
| | Tetracycline resistance determinant | - | Hachler *et al.*, 1987b |

| ORGANISM | GENES CLONED | SEQUENCE | REFERENCE |
|---|---|---|---|
| *C. pasteurianum* | Ferredoxin | + | Graves *et al.*, 1985 |
| | Galactokinase | p | Daldal & Applebaum, 1985 |
| | Mo-pterin binding protein (mop) | + | Hinton & Freyer, 1986 |
| | Nitrogenase Fe protein (*nifH*) | + | Wang *et al.*, 1988 |
| | Nitrogenase MoFe protein (*nifDK*) | p | Wang *et al.*, 1987 |
| *C. perfringens* | Chloramphenicol acetyltransferase | - | Abraham & Rood, 1987 |
| | Sialidase | + | Roggentin *et al.*, 1988 |
| | Tetracycline resistance determinant | - | Abraham *et al.*, 1988 |
| | Bacteriocin (pIP404, bcn) | + | Garnier & Cole, 1986 |
| | Bacteriocin immunity/ | + | Garnier & Cole, 1988a |
| | secretion (pIP404, *uviAB*) | + | Garnier *et al.*, 1987 |
| | Recombinase (pIP404 *res*) | | |
| *C. stercorarium* | Endo-α-1,4-glucanase | + | Schwarz *et al.*, 1988c |
| | α-glucosidase | - | Schwarz *et al.*, unpublished |
| | 2 xylanases | - | Schwarz *et al.*, unpublished |
| | α-xylosidase | - | Schwarz *et al.*, unpublished |
| *C. tetani* | Tetanus toxin | + | Eisel *et al.*, 1986<br>Fairweather *et al.*, 1986 |

| ORGANISM | GENES CLONED | SEQUENCE | REFERENCE |
|---|---|---|---|
| C. thermoaceticum | Formyltetrahydrofolate synthetase | - | Lovell et al., 1988 |
| | Leucine dehydrogenase | - | Shimoi et al., 1987 |
| C. thermocellum | Endo-α-1,4-glucanase (celA) | + | Beguin et al., 1985<br>Schwarz et al., 1986 |
| | Endo-α-1,4-glucanase (celB) | + | Grepinet & Beguin, 1986 |
| | Endo-α-1,4-glucanase (celC) | + | Petre et al., 1986<br>Schwarz et al., 1988a |
| | Endo-α-1,4-glucanase (celD) | + | Joliff et al., 1986 |
| | Endo-α-1,4-glucanase (celE) | + | Hall et al., 1988 |
| | Endo-α-1,3-glucanase (licA) | - | Schwarz et al., 1988b |
| | Endo-α-1,3-1,4-glucanase (licB) | - | Schwarz et al., 1985 |
| | α-glucosidase (bglA) | + | Grabnitz & Staudenbauer, 1988 |
| | α-glucosidase (bglB) | | Romaniec et al., 1987<br>Grabnitz and Staudenbauer, 1988<br>Kadam et al., 1988 |
| | Xylanase (xynZ) | + | Haslewood et al., 1988 |

Symbols: +, complete sequence; p, partial sequence; - no sequence.

## SECTION II

## DEVELOPMENT OF A HOST/VECTOR SYSTEM FOR THE SACCHAROLYTIC CLOSTRIDIA

### 1.4. INTRODUCTION

There are two basic requirements of a host/vector system: (1) a suitable cloning vector and; (2) a means of introducing the vector and its chimaeric constructions into the intracellular environment of the host.

The vector should preferably be an identifiably marked plasmid, capable of autonomous replication and stable maintenance in the intended host, and should possess unique restriction endonuclease sites to facilitate the cloning of foreign DNA restriction fragments into a non-essential region of the vector backbone. Although not absolutely essential, it is desirable to provide suitable transcriptional control signals that may be employed to elicit the expression of cloned DNA fragments. Such signals should be juxtaposed to the unique cloning sites. The most direct means of introducing such a vector into the host cell is by transformation of the naked DNA. Alternatively, the vector may be transferred via an intermediary donor cell by a conjugal mechanism.

### 1.5. VECTOR DELIVERY SYSTEMS

#### 1.5.1. TRANSFORMATION

All transformation procedures are dependent on the preparation of cells "competent" for the uptake of naked DNA. Both competence and DNA uptake may be induced in a variety of ways. These are considered below.

### 1.5.1.1. TRANSFORMATION OF NATURALLY COMPETENT CELLS

In contrast to the situation in E. coli where competence is normally an artificially induced condition, several species of Bacillus in addition to some streptococcal spp. (i.e., S. sanguis, S. mutans. and S. pneumoniae), and isolated members of other diverse genera, develop a natural physiological state in which they are capable of absorbing DNA from the surrounding medium (for review, see Stewart & Carlson, 1986).

Of the organisms studied, the mechanism by which transformation occurs has been studied in most detail for B. subtilis (for review, see Dubnau, 1976). However, all the bacterial transformation systems studied appear to follow a common sequence of events: (1) development of competence to import DNA, (2) binding of DNA, (3) entry of DNA, and (4) intracellular processing of DNA to result in integration, recombination, or the establishment of the autonomous plasmid state.

### 1.5.1.2. PEG-DEPENDENT, WHOLE CELL TRANSFORMATION

PEG-dependent, non-protoplast transformation procedures have been reported for a few Gram-positive organisms, namely Bacillus brevis (Takahashi et al.. 1983), Clostridium thermohydrosulphuricum (Soutschek-Bauer et al.. 1985), and Streptococcus lactis (Sanders & Nicholson, 1987). The specifics of the these protocols vary, but in all cases cited, preconditioning of the cells in certain buffers was essential for transformation. Also, each procedure exhibited an absolute requirement for PEG treatment, which distinguishes these procedures from other whole cell procedures which require treatment with alkaline cations, but not with

PEG, and from protoplast procedures which require cell wall digestion.

### 1.5.1.3. PROTOPLAST TRANSFORMATION

The first successful fusions of bacterial protoplasts with PEG were performed with polyauxotrophic mutants of B. subtilis (Schaeffer et al.. 1976) and B. megaterium (Fodor & Alfoldi, 1976). The conditions for efficient preparation and regeneration of B. subtilis protoplasts were originally reported by Wyrick and Rogers (1973) and later modified by Chang & Cohen (1979). These workers reported that up to 80% of a protoplasted population could be transformed with plasmid DNA with an efficiency of greater than $10^7$ transformants $\mu g^{-1}$ of DNA, making the method suitable even for the introduction of phenotypically cryptic plasmids. Similar procedures have been developed for Streptomyces (Bibb et al.. 1978) and at low efficiency for Streptococcus lactis (Kondo & Maksy, 1982). Since then several protocols have been adapted and optimised to allow efficient protoplast transformation ($10^4$-$10^6$ transformants $\mu g^{-1}$) of a limited number of Gram-positive species (see Table 1.2). However, the published procedures show a remarkable strain dependence.

TABLE 1.2

PROTOPLAST FORMATION, REGENERATION, AND TRANSFORMATION IN GRAM-POSITIVE BACTERIA

| ORGANISM | TRANSFORMATION | REFERENCE |
|---|---|---|
| *Bacillus subtilis* | + | Chang & Cohen, 1979 |
| *Bacillus megaterium* | + | Brown & Carlton, 1980 |
| *Bacillus licheniformis* | + | Imanaka *et al.*, 1981 |
| *Bacillus thuringensis* | + | Martin *et al.*, 1981 |
| *Bacillus stearothermophilus* | + | Imanaka *et al.*, 1982 |
| *Streptococcus lactis* | + | Kondo & McKay, 1982 |
| *Streptococcus lactis* subsp. *diacetyllactis* | + | Geis, 1982 |
| *Streptococcus faecalis* | + | Smith, 1985 |
| *Streptococcus thermophilus* | + | Mercenier *et al.*, 1988 |
| *Streptococcus cremoris* | + | Simon *et al.*, 1985 |
| *Staphylococcus aureus* | + | Gotz *et al.*, 1981 |
| *Streptomyces* spp. | + | Bibb *et al.*, 1978 |
| *Lactobacillus acidopholus* | + | Lin & Savage, 1986 |
| *Lactobacillus reuteri* | + | Morelli *et al.*, 1987 |
| *Lactobacillus plantarum* | + | Posno *et al.*, 1988 |
| *Lactobacillus casei* | - | Lee-Wickner & Chassy, 1984 |
| *Clostridium perfringens* | + | Haefner *et al.*, 1984 |
| *Clostridium acetobutylicum* | + | Lin & Blaschek, 1984 |
| | - | Allcock *et al.*, 1982 |
| | - | Reysset *et al.*, 1987 |
| *Clostridium pasteurianum* | - | Minton & Morris, 1983 |
| *Clostridium tertium* | - | Knowlton *et al.*, 1984 |
| *Clostridium saccharoperbutyl-acetonicum* | - | Yoshino *et al.*, 1984 |

Symbols: +, successful tranformation; -, formation and regeneration only.

### 1.5.1.4. ELECTROPORATION

It has been demonstrated that various types of cells can take up DNA under the influence of an electric field pulse (Fromm et al.. 1985; Hashimoto et al.. 1985; Potter et al.. 1984). This DNA uptake results from the transient permeabilisation of the cell membrane via field induced pore formation resulting from a high voltage electric discharge through a suspension of cells (Sowers et al.. 1986). This approach (electropermeabilisation or electroporation) has recently been applied to the transformation of both Gram-negative organisms such as E. coli and Erwinia caratovora (Minton, personal communication; Kazutoshi et al., 1988) and Gram-positive organisms such as Bacillus cereus (Shivarova, 1983, cited Chassy and Flickinger, 1987), Streptococcus lactis (Harlander, 1987; Powell et al.. 1988), and Lactobacillus casei (Chassy & Flickinger, 1987).

Transformation frequencies reported for electroporation range from $10^4/10^5$ transformants $\mu g^{-1}$ DNA for Gram-positive bacteria (Chassy and Flickinger, 1987; Powell et al.. 1988) to $10^6/10^9$ transformants $\mu g^{-1}$ DNA for Gram-negative bacteria (Minton, personal communication; Kazutoshi et al.. 1988). Although, to date, the factors which determine whether a particular bacterial strain is readily transformable by electroporation are not defined, the cell size, chain length, and degree of cell aggregation (Knight and Scrutton, 1986); the structure of the cell wall and glycocalyx; the presence of DNA restriction systems; the presence of extracellular exonucleases; and strain-specific variables affecting plasmid establishment and maintenance are probably crucial.

### 1.5.1.5. TRANSFORMATION OF SACCHAROLYTIC CLOSTRIDIA

During the 1970's a number of laboratories screened numerous saccharolytic clostridia, in particular C. pasteurianum and C. acetobutylicum. for evidence of a natural competence mechanism. No transformable strains were found (J.G. Morris, D.T. Woods, and M. Sebald, personal communication). The publication of protoplast methodology for B. subtilis (Chang & Cohen, 1979) and Streptomyces (Bibb et al., 1978) signaled a change in emphasis in the search for a transformation protocol. Accordingly, several publications appeared in the literature (see Allcock et al., 1982; Jones et al., 1985; Minton & Morris, 1983; Reysset et al. 1987; Yoshino et al. 1984) describing the preparation and regeneration of clostridial protoplasts. However, subsequent transformation of protoplasts appears to be the exception rather than the rule.

At the initiation of this study only two reports had appeared on the transformation of saccharolytic clostridial protoplasts. The host in both studies was C. acetobutylicum. The study of Reid et al. (1983) utilised phage CA1 DNA as the transfecting molecule. The efficiency of the transformation obtained could not be estimated, however, as the number of transfected cells (plaques) could not be directly quantified. In the second study the S. aureus plasmid pUB110 was employed (Lin and Blaschek, 1984). This plasmid encodes resistance to kanamycin (Km). All saccharolytic clostridia are inherently resistant to this antibiotic, and high numbers (equivalent to the reported transformation frequency) of spontaneous colonies are known to arise resistant to the concentration of Km used in this study (M. Young, personal communication). More importantly, repeated attempts by other laboratories to transform C.

acetobutylicum protoplasts with pUB110 have been unsuccessful (M.Young, D. T. Woods, M. Sebald, personal communication).

It is therefore apparent that at this time, no suitable plasmid vector was available which could be used to develop and optimise a protoplast transformation procedure. It therefore seemed more appropriate to construct such a vector by inserting a gene conferring resistance to an antibiotic, to which the clostridial host is particularly susceptible, into a plasmid known to replicate in a clostridial host. i.e., one of the many small cryptic plasmids isolated from the saccharolytic species.

### 1.5.2. CONJUGAL TRANSFER

#### 1.5.2.1. COINTEGRATE CONJUGAL PLASMID TRANSFER

This method is based on the previously observed broad host range of some of the large Streptococcal R-factors such as pAMβ1 (Clewell et al.. 1974; Leblanc & Lee, 1984), pIP501, and a conjugal mutant plasmid of pIP501, pVA797 (Evans & Macrina, 1983). The plasmid that has been studied most intensely to date is pAMβ1 (26.5 kb, $MLS^r$), which was originally isolated from Streptococcus faecalis (Clewell et al., 1974), and is conjugally transferable between a number of other Streptococcus species (Gibson et al.. 1979; Leblanc et al.. 1978; Gasson & Davies, 1980), various Lactobacillus species (Gibson et al.. 1979; Shrago et al.. 1986; Vescco et al.. 1983; Tannock, 1987), Staphylococcus aureus (Engel et al.. 1980; Schaberg et al.. 1982), Clostridium acetobutylicun (Oultram & Young, 1985; Yu & Pearce, 1986), Clostridium butyricum and Clostridium pasteurianum (Oultram, personal communication), and various species of Bacillus (Lereclus et al.. 1983; Orzech & Burke, 1984),

including B. subtilis (Landman & Pepin, 1982; Oultram & Young, 1985). The general strategy employed has been to conjugally transfer suitably marked non-conjugative plasmids to the intended Gram-positive host as a cointegrate assemblage with the conjugal mobiliser pAMβ1. Resolution of the cointegrate structure thus releases the "passenger" (shuttle) plasmid in the intracellular environment of the intended host. Subsequent segregation and plasmid stability are enhanced when both the shuttle and mobiliser plasmid are mutually incompatible by nature of their replication regions (Romero et al., 1987). If the "passenger" plasmid is not capable of autonomous replication in the new host, resolution will lead to it's loss. However, in such circumstances, the "passenger" plasmid may be maintained in the cointegrate form with the conjugal mobiliser if appropriate selection is maintained. It is therefore possible to introduce heterologous genes into new Gram-positive backgrounds as an integral part of a much larger cointegrate structure, but this is not desirable if significant levels of gene expression are required as a result of the low gene dosage dictated by the mobiliser moiety.

### 1.5.2.2. MOBILISATION

It has been reported that pAMβ1 can mobilise small non-conjugative plasmids such as pAMα1 (9.0 kb, $Tc^r$) and pAM610 (9.5 kb, $km^r$, $Tc^r$) in filter matings between different strains of S. faecalis (Clewell et al.. 1974; Schaberg et al.. 1982). In addition apparent mobilisation of pAM610 from S.faecalis to C. acetobutylicum has also been reported (Yu & Pearce, 1986), but this has not been confirmed in other laboratories (M. Young, personal communication). Neither were plasmids pUB110, pBC16, pHV33, nor pTV1 mobilised, when they were co-resident with pAMβ1 in B. subtilis donors (Oultram et al.. 1987). However, it must be born in

mind that the frequencies of transfer of pAMβ1 alone are poor and were any of these plasmids naturally mobilised at a low frequency, they might have escaped detection.

## 1.6. VECTOR COMPONENTS

The absence of indigenous small plasmids within the saccharolytic clostridia which encode selectable phenotypic traits means that such a plasmid must be constructed *in vitro* by the insertion of selectable markers into one of the ubiquitous cryptic plasmids previously isolated from this group. However, it should be noted that the insertion of foreign DNA into a plasmid can cause inactivation of essential replication functions. At the inception of this study, the available published methodology for introducing plasmids into saccharolytic clostridia was limited to extremely inefficient protoplast transformation procedures which were proving difficult to corroborate in other laboratories (M. Young, personal communication). It follows that should any transformation experiments prove unsuccessful, then it would not be possible to distinguish between the inability of the constructed chimaeric plasmid to replicate, or failure of the transformation procedure itself. One potential way around this paradoxical situation would be to identify the plasmid replicon, and thus ensure that this region remains intact during construction of chimaeras. As little is known of clostridial plasmid replication regions, it is instructive to review current knowledge of other Gram-positive plasmids.

### 1.6.1. REPLICATION OF GRAM-POSITIVE PLASMIDS

The majority of plasmids whose mode of replication has been extensively studied have been isolated from *E. coli* and this is well documented elsewhere (Scott, 1984, Novick,

1987). In comparison, plasmids isolated from Gram-positive bacteria have received scant attention. During the course of this study, however, considerable advances have been made with regard to our understanding of the replication strategy adopted by many of the commonly used Gram-positive cloning vectors. These additions to established knowledge are briefly reviewed below.

#### 1.6.1.1 MODE OF REPLICATION

Many of the multi-copy staphylococcal and streptococcal plasmids exhibit a broad host range being capable of autonomous replication in E. coli and B. subtilis in addition to their native hosts suggesting the existence of special features in their replication machinery (Ehrlich, 1977; Goze & Ehrlich, 1980; Lacks et al.. 1986; Espinosa et al.. 1982; Kok et al.. 1984). The recent observations that plasmids pT181 (Koepsel & Khan, 1987), pLS1 (Puyet et al., 1988), pC221, pC223, pE194, pT127, pUB110, pBC16, and pC194 (te Riele et al.. 1986a; Gros et al.. 1987) were present in the host cells as single-stranded forms in addition to double-stranded forms led these and other authors (Viret & Alonso, 1987) to suggest that these plasmids might replicate asymmetrically via a rolling circle mechanism analogous to that of the isometric (e.g., ØX174) and filamentous (e.g., M13) bacteriophages of E. coli, which are known to generate intracellular single-stranded DNA during their replication cycle (Koths & Dressler, 1978; Schaller, 1978). The details of the replication mechanisms of these bacteriophages is reviewed elsewhere (Baas, 1985).

The first evidence that the replication proteins of Gram-positive plasmids behaved as trans-active single-stranded endonucleases with a sequence specific topoisomerase-like activity (nicking-ligating), analogous to the bacterio-

phages of E. coli. was provided by Koepsel et al. (1985; Koepsel & Khan 1987). These authors reported that the RepC protein of pT181, which was also demonstrated to be rate-limiting for replication in an autogenously regulated manner (Manch-citron et al. 1986; Novick et al.. 1984a), cleaved pT181 single-stranded or double-stranded DNA at a unique "nick-site" in the "plus" origin of replication (Koepsel & Khan, 1987). The cleavage site was determined to lie immediately 5' to the AA dinucleotide. Comparison of the pT181 primary nick site with that of the single-stranded bacteriophages revealed a striking homology with the nick site of the filamentous bacteriophage M13 (Meyer et al.. 1979; Gros et al.. 1987). Additional similarities with the single-stranded bacteriophages include the unidirectional mode of replication (Khan et al.. 1982), the observation that the 5'-end at the RepC cleavage site is blocked, suggesting a covalent attachment (Koepsel et al., 1985), and that only the leading strand of DNA near the origin is replicated in vitro in the presence of high concentrations of dideoxynucleotides (Koepsel et al., 1986). It has also been proposed that plasmids exhibiting considerable mutual homology with pT181, such as pC221 and pS194 (Iordanescu et al.. 1978; Projan et al.. 1985), and the almost identical plasmid pNS1 (Noguchi et al., 1986) replicate in the same way. Examination of the published sequences of pC221 (Brenner et al., 1985), pS194 (Projan et al. 1988), and pNS1 (Shishido, 1988) revealed the presence of similar "nick-sites" showing extensive homology with that of pT181 (this thesis). It therefore appears that pT181 represents a prototype for a family of related plasmids which possess M13-type "nick-sites" in their "plus" strand origins which are located within the N-terminal region of the replication protein coding sequences.

Chronologically parallaled studies by Michel and Ehrlich

(1986a) provided evidence indicating that pC194 was "nicked" in a similar fashion to pT181 within the "plus" replication origin by its own essential replication protein. The putative "nick-site" was mapped to lie within the shorter palindrome and exhibited a striking homology with the cleavage site of the isometric bacteriophage, ØX174 (Langeveld et al., 1978; Gros et al., 1987). Gros et al. (1987) also revealed the presence of a very similar sequence in the plasmid pUB110 (McKenzie et al., 1986, 1987) and further demonstrated that pC194/pUB110 hybrids containing both intact origins resulted in initiation at one site and termination at another, a typical characteristic of a rolling circle replication mechanism (Dotto et al., 1982; Peeters et al., 1987). These observations and the reported presence of intracellular single-stranded DNA (te Riele et al., 1986a) prompted these authors to propose an asymmetric rolling circle mechanism of replication similar to pT181. Close examination of the published sequences of pFTB14 (Murai et al., 1987) and pRBH1 (Muller et al., 1986) also revealed the presence of this highly conserved consensus "nick-site" lying within the mapped "plus" origins of replication, again external and upstream of the replication protein coding sequences (this thesis). It would thus appear that pC194 might be prototypical for a second family of related plasmids which possess ØX174 type "nick-sites" in their "plus" strand origins of replication that are located external to and upstream to the replication protein encoding sequences.

Although plasmids pE194 and pLS1 are also reported to replicate via a rolling circle type mechanism (Villafane et al., 1987; Puyat et al., 1988), close examination of their sequences did not reveal the presence of either a ØX174 or a M13 type "nick-site" (this thesis). Some degree of structural resemblance between the "plus" strand origin of pLS1

and the equivalent region of the bacteriophage f1, however, has been noted (Puyet et al.. 1988). These two plasmids may in fact form a third family of Gram-positive plasmids. This contention is supported by two factors. Firstly, considerable DNA sequence homology exists between the "plus" strand origin regions of both plasmids (Lacks et al.. 1986; Villafane et al.. 1987). Secondly, comparison of the amino acid sequences of repB of pLS1 and RepF of pE194 has indicated considerable mutual homology in the N-terminal regions (Minton, et al.. 1988). This is of particular interest as it has been reported that the C-terminal portion of RepF is not absolutely essential to replication (Villafane et al., 1987).

### 1.6.1.2. INITIATION OF LAGGING STRAND SYNTHESIS

During the replicative form (RF) to RF rolling circle replication of the isometric and filamentous bacteriophages, the displaced leading ("plus") strand remains single-stranded until a palindromic element is exposed, forming a hairpin that serves as the lagging strand origin (Sims & Dressler, 1978; Shlomai & Kornberg, 1980; Baas, 1985). Similar palindromic elements comprising a major axis of hyphenated dyad symmetry (pal A) have been identified in the S. aureus plasmids, pT181, pC221, pE194, pC194, pS194, pIM13, and pSN2 (Gruss et al.. 1987), and in the streptococcal plasmid pLS1 (del Solar et al., 1987). However, no such sequence has been identified for pUB110.

Experimental studies using palA plasmids, have shown that the pal A defective genotype is accompanied by a marked reduction in copy number, increased plasmid instability, and the accumulation of large quantities of strand-specific circular single-stranded DNA (del Solar et al.. 1987; Gruss et al.. 1987). In addition, the orientation of pal A is of

vital importance, being required in the same orientation as the replication gene; consistent with the asymmetric rolling circle mechanism of replication in which the initiation of replication of the displaced strand occurs de novo. only after the palindromic lagging strand initiation site is exposed. Additional work by Gruss et al. (1987) has demonstrated that the position of pal A on the plasmid is not of vital importance. However, deletion or inactivation of pal A is not lethal to the plasmid, suggesting the existence of a less efficient alternative lagging strand initiation site(s). Such an alternative lagging strand origin has been proposed for pLS1 (del Solar et al.. 1987).

The efficiency by which pal A is recognised appears to vary with the host. The highest efficiencies observed for the conversion of single to double-stranded DNA appears to be in the native hosts; i.e., Streptococcus pnuemoniae for pLS1 (del Solar et al.. 1987) and Staphylococcus aureus for the staphylococcal plasmids. PalA derivatives in B. subtilis of the staphylococcal plasmids have copy numbers equivalent to their pal $A^+$ parents. These same pal A derivatives exhibit a much reduced copy number on transformation into S. aureus (Gruss et al., 1987).

1.6.1.3. CONTROL OF REPLICATION

Regulation of replication of these Gram-positive plasmids is indirectly controlled by the "inhibitor-target" type mechanism in which the inhibitor (usually a pair of 5' countertranscript RNA molecules) blocks the synthesis of the trans-active replication proteins (for review, see Gruss & Ehrlich, 1989; Novick, 1987).

### 1.6.2. PROMOTER SYSTEM

A desirable feature of a *Clostridium* cloning vector would be the provision of a strong promoter which could be utilised to overexpress inserted genes. At the initiation of this project, no clostridial promoter had been characterised. However, during the course of these studies the nucleotide sequences of various genes and their regulatory regions have been determined (Table 1.1). In the case of the *C. pasteurianum* ferredoxin (Fd) gene (Graves *et al*., 1985; Graves & Rabinowitz, 1986), the *C. thermocellum celA* gene (Beguin *et al*., 1986), and the *C. tetani* tetanus toxin gene (Fairweather *et al*., 1986), the sequences important in transcriptional initiation have been mapped. The early availability of the Fd sequence allowed the construction of a promoter cartridge based on this gene (see Chapter 4). As this cartridge was initially tested in *B. subtilis*, and because of the lack of available information on clostridial gene expression, it would be beneficial to briefly review gene expression in other Gram-positives, and in particular *B. subtilis*.

#### 1.6.2.1. PROMOTERS

Promoters are binding sites on DNA at which RNA polymerase binds and initiates transcription of DNA into RNA. Although binding sites for RNA polymerase generally extend from about 45 base-pairs upstream from the start point of transcription to about 20 base-pairs down stream, promoter recognition in bacteria is principally governed by two limited hexanucleotide sequences at two regions of the promoter (for reviews see Rosenberg & Court, 1979; Hawley & McClure, 1983). These are the recognition or "-35" region and the Pribnow box or "-10" region, which are centered at positions approximately 35 and 10 base-pairs, respectively,

preceding the transcription start point. RNA polymerase is believed to interact directly with bases in both regions because: (1) the hexanucleotide sequences T-T-G-A-C-A and T-A-T-A-A-T are approximately conserved at the -35 and -10 regions, respectively, among a large number of promoters whose sequences have been determined (principally in E. coli but also in many genes from Gram-positive organisms); (2) mutations within either hexanucleotide sequence severely impair promoter function; and (3) bases within or near both regions are in close contact with RNA polymerase.

It follows that transcriptional specificity in procaryotic organisms resides in the interaction between the ribonucleic acid (RNA) polymerase holoenzyme and the promoter site of genes and operons. The RNA polymerase subunit which controls the specificity of this interaction has been shown to be the sigma ($\sigma$) factor, since the RNA polymerase core itself does not recognise promoter sites (Burgess et al., 1969). This is not to imply, however, that the other subunits of the core do not have any role in the specificity of the interaction once the holoenzyme has been formed by the addition of $\sigma$ to the core enzyme.

In E. coli as well as Gram-positive organisms, initiation of transcription takes place preferentially at a purine in position 7 or 8 downstream from the last T in the -10 sequence (Murray & Rabinowitz, 1982; Aoyama & Takanami, 1985). It has been indicated that certain positions in the consensus sequences are more highly conserved and important for efficient utilisation in B. subtilis than is the case in E. coli. For example, the "invariant G" in the -35 region. However at least one strong B. subtilis promoter has been characterised which does not have the invariant G in this region (Nakahama et al.. 1986).

A compilation analysis of 29 promoter regions derived from Gram-positive organisms (mostly B. subtilis) by Graves & Rabinowitz (1986) demonstrated additional conserved regions outside the classic -10 and -35 regions. These authors pointed out an "A" cluster at positions 41 to 45 with greater than 50% conservation rate; this feature being absent in E. coli promoters with the exception of a weakly conserved "A" at position -45. They also reported two other conserved areas bordering the -10 region, and in effect, broaden the limits of that promoter element. This led these authors to propose the existence of an "extended" consensus sequence for Gram-positive vegetative promoters. This theory is supported by the observations that the E. coli lacUV5 and tac promoters, which do not conform to the additional conserved regions, are poorly utilised by B. subtilis vegetative RNA polymerase when compared to B. subtilis promoters (Lee et al., 1980; Moran et al., 1982).

### 1.6.2.2. THE "GRAM-BARRIER"

Gene transplantation experiments have shown that the well studied Gram-negative organism, E. coli, is promiscuous in its ability to recognise transcription and translation signals from a wide variety of microorganisms, especially those derived from Gram-positive organisms. Early examples included the expression in E. coli of drug resistance genes from Gram-positive bacteria (Ehrlich, 1978). More recently, many genes from Gram-positive organisms have been cloned in E. coli based on their expression in this organism. Examples, among others, include genes from Bacillus (Makaroff et al., 1983), Staphylococcus (McLaughlin et al., 1981a; Shuttleworth et al., 1987), Corynebacterium (Kaczorek et al., 1985), and clostridia (Cornet et al., 1983; Hinton & Freyer, 1986; Kadam et al., 1988; Karube et al., 1983; Millet et al., 1985; Schwarz et al., 1985; Zappe et

al., 1986; 1987). However, in general, the reverse does not appear to be the case. Studies with B. subtilis have shown that, in contrast to E. coli, it is very limited in its ability to express genes from other genera (Kreft et al., 1978); clear examples of foreign genes utilised efficiently in B. subtilis are drug resistance genes from other Gram-positive bacteria such as Staphylococcus (Ehrlich, 1978; Kreft et al., 1978) and Streptococcus (Yagi et al., 1978). The observations that B. subtilis fails to efficiently transcribe from lacUV5 (Lee et al., 1980) and coliphage T4 (Shorenstein & Losick, 1973) promoters and to form stable, rapidly starting complexes with coliphage T7 promoters (Wiggs et al., 1979) and that B. subtilis ribosomes fail to translate effectively from E. coli mRNAs (Stallcup & Rabinowitz, 1973; Legault-Demare & Chambliss, 1975) has suggested that heterologous gene expression is probably restricted at both the transcriptional and translational levels. Whether or not the observed difficulties with heterologous gene expression in B. subtilis and perhaps other Gram-positives are attributable to a generalised "Gram-barrier" or a species-specific expression barrier has yet to be fully determined.

### 1.6.2.3. OTHER PROMOTER SPECIFICITIES IN B. SUBTILIS

The recent concept of the transcriptional machinery of the procaryotic cell has been influenced by the earlier work done with E. coli in which one RNA polymerase holoenzyme was identified that appeared to control the expression of all of the genes in the cell. The E. coli RNA polymerase holoenzyme consists of a core polymerase containing the four subunits ($\alpha_2\beta\beta'$) (Burgess, 1969) and the $\sigma^{70}$(the superscript stands for the molecular weight of the $\sigma$ factor X $10^{-3}$) (Burton et al., 1981). Thus it appeared that $\sigma^{70}$ allowed the holoenzyme to recognise all promoters of E.

coli. sometimes with aid of positive factors (Raibaud & Schwartz, 1984).

More recently, studies with phage SP01-infected and uninfected B. subtilis cells has provided evidence for the existence of several other RNA polymerase forms besides the major RNA polymerase holoenzyme ($E\sigma^{43}$) and is fully reviewed elsewhere (Losick & Pero, 1981; Doi & Wang, 1986).

### 1.6.2.4. TRANSCRIPTIONAL CONTROL BY SPECIFIC $\sigma$ FACTORS: A BARRIER TO HETEROLOGOUS GENE EXPRESSION?

It would appear that the expression of heterologous genes amongst the procaryotes may be affected to great extents at the level of transcription, particularly if the the intended host is a Gram-positive organism such as B. subtilis where it seems that "extended" promoter regions are required for efficient transcription. This observation does pose limitations on the flexibility of Gram-positive hosts such as B. subtilis which are in their own right of great industrial potential by virtue of their effective protein secretory systems. However, with careful planning, it should be possible to overcome problems of poor transcription of heterologous genes in Gram-positive hosts by the provision of appropriate promoter sequences. This is supported by the observations by Peschke and co-workers (1985) that certain promoters from Gram-negative organisms are efficiently utilised by $E\sigma^{43}$ of B. subtilis effecting high expression of adjacent genetic material. These promoters, the T7A1 promoter and T5 promoters, all possess A+T rich regions between positions -35 and -50 relative to the initiation sites. This feature is absent in the tac, lacUV5, trp, and T7A2 promoters which are poorly recognised by $E\sigma^{43}$. The importance of this A+T rich region has been proposed by these workers and by Banner et al. (1983).

In addition. These authors further pointed out that the "spacer" region between the -10 and -35 sequences is not so critical varying from 16 to 18 nucleotides.

In conclusion, there does appear to be a natural barrier, of varying stringency, to heterologous gene expression at the level of transcription in B. subtilis. Whether this phenomenon is particular to Bacillus spp. or more generalised amongst Gram-positive organisms is at present largely unknown, but seems likely. This hypothesis is supported by the "extended" promoter consensus sequence for Gram-positive genes proposed by Graves & Rabinowitz (1986).

### 1.6.3. TRANSLATIONAL CONTROL OF GENE EXPRESSION

It can be envisaged that translational control of gene expression might be effected primarily at two levels: (1) at the level of the ribosome/mRNA interaction thus determining the rate of translation initiation, and (2) at the level of elongation rate which might be regulated by codon choice (i.e., tRNA availability) and codon context (the nature of adjacent bases). Apart from the transcriptional differences already discussed, it is at the translational level that considerable restriction of heterologous gene expression takes place in Bacillus subtilis.

#### 1.6.3.1. TRANSLATION INITIATION

The initiation of protein synthesis by procaryotic ribosomes involves the selection of an appropriate site ("ribosome binding site") on the mRNA by the 30 S subunit of the ribosome (Gold et al.. 1981; Gran, 1984). The ribosome binding site consists of an initiation codon, a polypurine Shine-Dalgarno (SD) sequence which is complementary to varying degrees to the 3'- end of the 16 S rRNA,

and an appropriate spacing ("window") between the two, ranging in size, from 3 to 11 base pairs (Shine & Dalgarno, 1974; Gold et al., 1981).

Mclaughlin and co-workers (1981b) have suggested that the Shine-Dalgarno complementarity required by B. subtilis and other Gram-positive ribosomes is significantly greater than that required by E. coli ribosomes, and is required for species-specific translation. Analysis of sequence information derived from over 40 Gram-positive translation initiation sites has supported this hypothesis as they all show "strong" Shine-Dalgarno sequences (Hager & Rabinowitz, 1985a).

The spacer "window" region has been shown to be important in affecting the initiation of translation, at least in Gram-negatives. It has been demonstrated that both the sequence and length of the window are important. The number of nucleotides between the last base of the Shine-Dalgarno complementarity and the first nucleotide of the initiation codon reduces translation efficiency if less than about 5 or greater than about 9 nucleotides (Thummel et al., 1981; Singer et al., 1981).

It has also been reported that many mRNAs are characterised by the potential to form stem loops in the ribosome binding region; typically the SD sequence is often found in a non-helical "loop" section with the initiation codon comprising part of the "stem" (Selker & Yanofsky, 1979). These mRNA secondary structures are believed to restrict translation by sequestering initiation sites within the stem of the stem-loop structures until the stem is disrupted by the formation of alternate structures or by translation read-through (Kastelein et al., 1983). Secondary structure has also been reported to be important in sequences upstream of

the ribosome binding site; alterations in these sequences having dramatic effects on translation efficiency (Hall et al., 1982; Iserentant & Fiers, 1980; Kastelein et al.. 1983; Stanssens et al.. 1985).

Amongst the Gram-positive genes sequenced to date are many cases of alternative initiation codons (such as GUG and UUG) which are only rarely encountered in Gram-negative mRNAs (29% versus 9%; Hager & Rabinowitz, 1985b). Although not the "normal" codons for directing the binding of fMet-tRNA in the translation initiation reaction, both UUG (a leucine codon) and GUG (a valine codon) have been shown to effectively bind fMet-tRNA (Murray & Rabinowitz, 1982).

A noteworthy characteristic of Gram-positive mRNAs is the apparent reduced dependence for initiation factors (IF's) in the translation reaction by systems derived from either Gram-positive or Gram-negative sources (McLaughlin et al., 1981c). This is in marked contrast to the absolute requirement for initiation factors for translation of mRNA derived from Gram-negative sources. Addition of E. coli ribosomal protein S1 was not found to effect the translational activity of the system derived from B. subtilis.

1.6.3.4. TRANSLATION ELONGATION

It is well established that there is preferential codon usage in E. coli with the concomitant presence of varying amounts of iso-accepting tRNA species (Ikemura, 1981). For example, among six codons for Leu, CUG is preferentially used (about 70%) and among six codons for Arg, CGU or CGC are used more than 90% when 25 genes of E. coli were analysed (Konisberg and Godson, 1983). On the other hand, Leu, Ile and Arg codons such as CUA, AUA, CGA, CGG, AGA, AGG are rarely used. It has been proposed that highly expressed

genes are made up of codons corresponding to abundant tRNA species and that weakly expressed genes are correspondingly comprised of less abundant tRNA species (Grosjean & Fiers, 1982; Gouy & Gautier, 1982). This postulate therefore implies that codons corresponding to rare tRNAs are translated more slowly than the common ones. Because of the high correlation between codon usage, tRNA abundance and level of gene expression, these authors suggested that codon usage is a modulator of gene expression. However, more recent studies have suggested that codon choice has more to do with the energetics of proof-reading than with elongation rates per se as far as abundant proteins are concerned and that codons corresponding to rare tRNAs influence elongation rates only if their frequencies on messengers is elevated from those occurring naturally in E. coli (Holm, 1986).

In contrast to E. coli. in B. subtilis there does not appear to be such a marked codon bias even though differing tRNA species profiles have been found both in vegetative cells and spores (Vold, 1973). For example, the two amino acids previously cited that each have six possible codons (Leu and Arg) have a strong bias toward one of the six codons in their genes in E. coli: in B. subtilis the six codons are more evenly distributed among the genes so far sequenced (Piggot & Hoch, 1985). Analysis of 21 genes from B. subtilis has supported this observation of reduced codon bias although it has also been pointed out that rare codons are present in some B. subtilis genes (Ogasawara, 1985). It has also been suggested that the unusually unbiased codon usage in B. subtilis may be related to the unique organisation of tRNA genes in this bacterium (Ogasawara, 1985) where all these genes appear to be organised into only six putative transcriptional units (Vold, 1985). Further analysis of the sequences of the mature tRNAs and their relative

abundance in the cell is necessary to reveal the biological significance of the unusual codon usage of B. subtilis genes. Yet, so far as the codon usage is concerned, B. subtilis and perhaps any other similar Gram-positive organism exhibiting relatively unbiased codon usage, should be a good host for the production of foreign gene products.

A final consideration which is reported to affect elongation rate of mRNAs in E. coli is the "codon context" (Shpaer, 1986). Shpaer reports that the choice of synonomous codons in highly expressed genes depends on the nucleotides adjacent to the codon; the efficiency of translation of some codons being modulated by their context. Shpaer further suggests that consideration of codon context can enable improved expression of "synthetic" genes in E. coli. Whether or not a similar codon context phenomenon exists in Gram-positive organisms has yet to be established.

### 1.6.4. ANTIBIOTIC RESISTANCE GENES.

Of the wide range of heterologous Gram-positive resistance genes available only a few have been demonstrated to function efficiently in a clostridial host (C. acetobutylicum): these being the $Em^r$ gene of pAMβ1, and to a lesser extent the $Cm^r$ gene of pC194 (Oultram & Young, 1985; Oultram et al., 1987). It is worth drawing attention to the fact that the $Em^r$ marker of pAMβ1 is likely to be of most use in C. acetobutylicum as spontaneous $Em^r$ bacteria have never been encountered (M. Young, personal communication), and that most strains appear to be partially tolerant of chloramphenicol arising to their ability to undertake the ferredoxin-dependent reduction of acryl nitro groups on the molecule (O'Brien & Morris, 1971). However, enhanced tolerance to the drug has been demonstrated when the $Cm^r$ deter-

minant of pC194 has been introduced into C. acetobutylicum (Oultram et al.. 1987). It is noteworthy that $Km^r$ and $Tc^r$ individuals arise at appreciable frequencies in C. acetobutylicum (M. Young, personal communication).

## 1.7. AIMS OF THE THESIS

In view of the increased awareness of the biotechnological importance of the saccharolytic clostridia it was elected to initiate studies towards the development of a host/vector system for this important genus. Such a system could therefore be exploited to facilitate strain improvement of those commercially important species; for example, increased substrate utilisation.

As no suitably "marked" small plasmids of clostridial origin had been described it was decided to initially characterise a suitable clostridial plasmid so that it could be used in the *in vitro* construction of an identifiable *E. coli/Clostridium* shuttle vector. Once developed, it was anticipated that this vector be used to develop a reliable transformation procedure for the saccharolytic clostridia.

In addition, it was intended to provide a means for gene expression in the intended clostridial host. In the absence of any knowledge of the transcriptional control mechanisms in clostridia, it was decided to exploit the transcriptional control signals of a highly expressed clostridial gene.

# CHAPTER TWO

# MATERIALS AND METHODS

## 2.1. MATERIALS

### 2.1.1. CHEMICALS AND MATERIALS

Except where stated below, all chemicals were purchased from BDH Chemicals Ltd., Poole, Dorset, and were of analytical reagent grade (AnalaR) wherever possible.

Sigma Chemical Company (UK) Ltd., Poole, Dorset,England

Ethidium bromide, ampicillin, tetracycline, chloramphenicol, bromophenol blue, xylene cyanol, orange G, Ficoll, adenosine 5'-triphosphate (ATP), DL-dithiothreitol (DTT), 5-bromo-4-chloro-3-indolyl-B-D-galactopyranoside (X-gal), isoamylalcohol, dimethylformamide, polyethylene glycol (PEG M wt 6000), Brij 58, sodium deoxycholate, polyvinylpyrrolidone, bovine serum albumin (BSA pentax fraction V), salmon sperm DNA, thiamine, proline, Triton X-100, Coomassie blue.

Bethesda Research Laboratories, Paisley, Scotland

Urea, caesium chloride, agarose , ammonium persulphate, formamide, isopropylthio-B-galactoside (IPTG), phenol, -N,N.N',N'-tetramethylethylenediamine (TEMED) (all ultrapure), DNA kilobase ladder, Lambda DNA/HindIII fragments.

Biorad Laboratories, Caxton way, Watford, Herts

Acrylamide, methylene-bis acrylamide, "zeta probe" nylon membrane, 10 cm econo-columns for sephadex chromatography.

Difco Laboratories, West Molesley, Surrey, England.

Tryptone, yeast extract, casaminoacids (vitamin assay), agar.

Oxoid Ltd., Wade Rd. Basingstoke, Hants.

Reinforced Clostridial Medium (RCM)

Pharmacia (GB) Ltd. Middlesex, England.

Sephadex G-50, Sephadex G-25 prepacked columns (PD-10)

Anderman & Co., East Molesley, Surrey ,England.

Nitrocellulose sheets, nitrocellulose filter discs, manufactured by Schleicher & Schull.

Amersham International, Bucks, England

[alpha-$^{32}$P] dATP (aqueous), [gamma-$^{32}$P] dATP (50% ethanol), Amplify, protein molecular weight size markers, in vitro transcription/translation kit (code N.380).

### 2.1.2. ENZYMES

All restriction endonucleases, T4 polynucleotide ligase, and T4 polynucleotide kinase, were obtained from Bethesda Research Laboratories (Gibco BRL), Paisley, Scotland. Calf intestinal phosphatase, T4 DNA polymerase, DNA polymerase I, Klenow polymerase, DNAase I, RNAase T1, and Proteinase K were obtained from Boehringer Mannheim (BCL), Lewes, Sussex, England. Lysozyme (Grade I) was obtained from Sigma Chemical Company (UK Ltd).

### 2.1.3. MEDIA

All media listed below were prepared in double distilled, deionised, water and autoclaved at 15 p.s.i. (121$^{o}$C) for 15 min unless otherwise stated.

### 2.1.3.1. ROUTINE GROWTH MEDIA

Luria (L) Broth

This medium was routinely used for the cultivation of *E. coli* and *B. subtilis*.

| | g $l^{-1}$ |
|---|---|
| Tryptone | 10.0 |
| Yeast Extract | 5.0 |
| NaCl | 5.0 |

The pH was adjusted to 7.4 using 1 M NaOH. For solid media (L-agar) 2% (w/v) purified agar was added.

2 X YT Broth

This medium was routinely used for the cultivation of *E. coli* JM101.

| | g $l^{-1}$ |
|---|---|
| Bacto-tryptone | 16.0 |
| Bacto-yeast extract | 10.0 |
| NaCl | 10.0 |

The pH was adjusted to 7.4 using 1 M NaOH. For solid media (2 X YT agar), 2% (w/v) purified agar was added.

H-Top Agar

This medium was used in the molten state (45°C) for the soft agar overlays in the cultivation of M13 phage plaques.

| | g $l^{-1}$ |
|---|---|
| NaCl | 8.0 |
| Agar | 8.0 |

The pH was adjusted to pH 7.4 with 1 M NaOH.

M9 Medium (10 X stock)

This defined minimal medium was used for the cultivation of *E. coli* strains.

| | g $l^{-1}$ |
|---|---|
| $Na_2HPO_4$ | 60.0 |
| $KH_2PO_4$ | 30.0 |
| NaCl | 5.0 |
| $NH_4Cl$ | 10.0 |

The pH was adjusted to 7.4 with 10 M NaOH, and the mixture autoclaved. An appropriate volume of this 10 X stock was added to molten (50°C) 2% agar/water (w/v) prior to the addition of the following constituents (per 200 ml):

2 ml 1 M $MgSO_4$ (autoclaved seperately)
0.1 ml 0.1 M $CaCl_2$ " "
10 ml 20% Glucose (filter sterilised)

M9/Folate agar

This solid medium was used for the selection of clones carrying a functional Carboxypeptidase $G_2$ gene. M9 medium was prepared as already described, with the incorporation of 2 ml folate solution per 200 ml which was added with the later constituents.

Folate solution was prepared by the slow addition of 350 g of folic acid to 1.5 l distilled water. The pH was adjusted to 7.5 with 5 M NaOH. When the solution cleared, becoming deep orange in colour, the volume was adjusted to 3.5 l.

Tryptone, Yeast extract, and Glucose broth (TYG broth)

This medium was used for the routine cultivation of clostridia.

| | $g\ l^{-1}$ |
|---|---|
| Tryptone | 30.0 |
| Yeast extract | 20.0 |
| Glucose | 5.0 |
| Sodium thioglycollate | 1.0 |

Reinforced Clostridial Medium (RCM)

This was used for the routine cultivation of clostridia and was obtained from Oxoid Ltd. 52.5 g was added per litre prior to autoclaving. For solid media, purified agar was added to 2% (w/v).

### 2.1.3.1. BACILLUS TRANSFORMATION MEDIA

The following media were used exclusively for the transformation of B. subtilis.

Spizizen's minimal salts medium (SMM)

| | g $l^{-1}$ |
|---|---|
| $K_2HPO_4$ | 14.0 |
| $KH_2PO_4$ | 6.0 |
| Na citrate.$2H_2O$ | 1.0 |
| $MgSO_4$ | 0.2 |

The pH was adjusted to 7.0 prior to use.

Competence medium

| | |
|---|---|
| SMM | 20.0 ml |
| Glucose (20% w/v) | 0.5 ml |
| $CaCl_2$ (5mM) | 0.2 ml |
| $MgSO_4$ (0.5 M) | 0.1 ml |
| $MnSO_4$ (0.5 mM) | 0.02 ml |
| Casamino acids (10% w/v) | 0.4 ml |

Transformation medium

| | |
|---|---|
| SMM | 20.0 ml |
| Glucose (20% w/v) | 0.5 ml |
| $MgSO_4$ (0.5 M) | 0.2 ml |
| Casamino acids (10% w/v) | 0.02 ml |

Sporulation salts A

| | g $l^{-1}$ |
|---|---|
| $FeCl_2.6H_2O$ | 0.98 |
| $MgCl_2$ | 8.30 |
| $MnCl_2$ | 19.79 |

Sporulation salts B

| | g $l^{-1}$ |
|---|---|
| $NH_4Cl$ | 53.5 |
| $Na_2SO_4$ | 10.6 |
| $KH_2PO_4$ | 6.8 |
| $NH_4NO_3$ | 9.7 |

2.1.3.3. ANTIBIOTICS

Antibiotics were incorporated in both solid and liquid media for the selection of resistant bacterial clones.

Ampicillin

Stock solution: 25 mg $ml^{-1}$ of the sodium salt of ampicillin in distilled water. This was sterilised by filtration (0.22 $\mu$m Millipore disposable filter) and stored at -20°C.

Working concentration: 50-100 $\mu$g $ml^{-1}$.

Tetracycline

Stock solution: 12.5 mg $ml^{-1}$ tetracycline hydrochloride in 50% ethanol (v/v). This was stored at -20°C in the dark.

Working concentration: 12.5-15 $\mu$g $ml^{-1}$.

Chloramphenicol

Stock solution: 34 mg $ml^{-1}$ in 50% ethanol (v/v). This was stored at -20°C.

Working concentration: For amplification of plasmids, 170 $\mu$g $ml^{-1}$; for selection of resistant bacteria, 5-10 $\mu$g $ml^{-1}$ (*Bacillus subtilis* and *Clostridium acetobutylicum*) and 30 $\mu$g $ml^{-1}$ (*Escherichia coli*).

Erythromycin

Stock solution: 10 mg $ml^{-1}$ in 100% ethanol. This was stored at -20°C.

Working concentration: 10 $\mu g$ $ml^{-1}$.

### 2.1.4. BUFFERS AND SOLUTIONS

TE Buffer (1 X)

| | |
|---|---|
| Tris-HCl | 10 mM |
| $Na_2EDTA$ | 0.1 mM |

pH 8.0

SSC Buffer (20 X)

| | |
|---|---|
| NaCl | 3.0 M |
| $Na_3Citrate.2H_2O$ | 0.3 M |

pH 7.0

TBE Buffer (10 X)

| | |
|---|---|
| Tris-borate | 0.9 M |
| Boric acid | 0.9 M |
| $Na_2EDTA$ | 0.03 M |

pH 8.3

Acetate-EDTA

| | |
|---|---|
| NaAcetate | 3.0 M |
| $Na_2EDTA$ | 0.001 M |

pH 5.5

Saline-EDTA

| | |
|---|---|
| NaCl | 0.15 M |
| $Na_2EDTA$ | 0.1 M |

pH 8.0

TM buffer

| | |
|---|---|
| Tris-HCl | 0.1 M |
| $MgCl_2$ | 0.05 M |

pH 8.5

Oligonucleotide storage buffer

| | |
|---|---|
| Tris-HCl | 0.01 M |
| NaCl | 0.005 M |
| $Na_2EDTA$ | 0.001 M |

pH 8.0

Oligonucleotide elution buffer

| | |
|---|---|
| $NH_4$ acetate | 0.5 M |
| Mg acetate | 0.01 M |
| SDS | 0.1 % |
| $Na_2EDTA$ | 0.0001 M |

pH 8.0

Ligation buffer (10 X)

| | |
|---|---|
| Tris-HCl | 0.3 M |
| NaCl | 0.3 M |
| $MgCl_2$ | 0.075 M |
| Spermidine | 0.01 M |
| ATP | 0.0025 M |
| DTT | 0.02 M |
| $Na_2EDTA$ | 0.002 M |

pH 7.5

Kinase buffer (10 X)

| | |
|---|---|
| Tris-HCl | 0.5 M |
| $MgCl_2$ | 0.1 M |
| DTT | 0.05 M |

pH 7.4

Nick-translation buffer (10 X)

| | |
|---|---|
| Tris-HCl | 0.2 M |
| $MgCl_2$ | 0.1 M |

pH 7.5

T4 polymerase buffer (10 X)

| | |
|---|---|
| Tris-acetate (pH 7.9) | 0.33 M |
| Potassium acetate | 0.66 M |
| Magnesium acetate | 0.10 M |
| DTT | 0.005 M |
| BSA (pentax Frac V) | 1 mg $ml^{-1}$ |

Calf intestinal phosphatase buffer (10 X)

| | |
|---|---|
| Tris-HCl | 0.5 M |
| $MgCl_2$ | 0.01 M |
| $ZnCl_2$ | 0.001 M |
| Spermidine | 0.01 M |

pH 9.0

Nuclease S1 buffer (X10)

| | |
|---|---|
| Na acetate | 0.5 M |
| NaCl | 3 M |
| $ZnSO_4$ | 0.05 M |

pH 4.7

B. subtilis washing buffer

| | |
|---|---|
| Tris-Cl | 0.05 M |
| $Na_2EDTA$ | 0.005 M |
| NaCl | 0.1 M |

pH8.0

B. subtilis lysis buffer

| | |
|---|---|
| $Na_2EDTA$ | 0.01 M |
| NaCl | 0.15 M |
| lysozyme | 10 mg $ml^{-1}$ |

S.T.E.T buffer

| | |
|---|---|
| Sucrose | 8 % (w/v) |
| Triton X-100 | 5 % (v/v) |
| $Na_2EDTA$ | 0.005 M |
| Tris-HCl | 0.05 M |

pH 8.0

Denhardt's reagent (50 X)

| | g $l^{-1}$ |
|---|---|
| Polyvinylpyrrolidone | 10 |
| BSA (Pentax fraction V) | 10 |
| Ficoll | 10 |

Filter sterilised and stored at -20°C.

40% Acrylamide stock

| | g $l^{-1}$ |
|---|---|
| Acrylamide | 380 |
| Bis-acrylamide | 20 |

Made up to 1 l, deionised with Amberlite MB-1 resin, and filtered.

0.5 X TBE gel mix

| | |
|---|---|
| 40% acrylamide stock | 150 ml |
| 10X TBE | 50 ml |
| Urea (ultra-pure) | 460 g |

Made up to 1 l.

5.0 X TBE gel mix

| | |
|---|---|
| 40% acrylamide | 150 ml |
| 10 X TBE | 500 ml |
| Urea (ultra-pure) | 460 g |
| Bromophenol blue | 0.05 g |

Made up to 1 l.

Formamide gel loading dye

| | |
|---|---|
| Deionised formamide | 100 ml |
| Xylene cyanol FF | 0.1 g |
| Bromophenol blue | 0.1 g |
| $Na_2EDTA$ (0.5 M) | 2 ml |

SDS-PAGE running gel buffer (lower Tris) (2X)

| | |
|---|---|
| Tris-HCl | 0.75 M |
| SDS | 0.2% w/v |

pH 8.8

SDS-PAGE stacking gel buffer (upper Tris) (2X)

| | |
|---|---|
| Tris-HCl | 0.25 M |
| SDS | 0.2% w/v |

pH 6.8

SDS-PAGE electrophoresis buffer

| | $g\ l^{-1}$ |
|---|---|
| Tris-HCl | 3.03 |
| Glycine | 14.4 |
| SDS | 1.0 |

pH 8.3

SDS-PAGE sample loading buffer

| | |
|---|---|
| Tris-HCl | 0.125 M |
| Glycerol | 20% (v/v) |
| Beta-mercaptoethanol | 10% (v/v) |
| SDS | 4% (v/v) |
| Bromophenol blue | 0.0015% (w/v) |

SDS-PAGE gel fixing solution

| | |
|---|---|
| isopropranol | 25% (v/v) |
| glacial acetic acid | 10% (v/v) |

DNA sequencing reaction mixes

dNTP's were stored at 50 mM in 1 X TE buffer and diluted to a working concentration of 0.5mM. ddNTP's were at 10 mM in 1 X TE buffer.

REACTION MIXES

| dNTP stocks | T | C | G | A |
|---|---|---|---|---|
| 0.5 mM dTTP | 25 | 500 | 500 | 500 |
| 0.5 mM dCTP | 500 | 25 | 500 | 500 |
| 0.5 mM dGTP | 500 | 500 | 25 | 500 |
| 10 mM ddTTP | 50 | | | |
| 10 mM ddCTP | | 8 | | |
| 10 mM ddGTP | | | 16 | |
| 10 mM ddATP | | | | 6 |
| 1 X TE buffer | 1000 | 1000 | 1000 | 1000 |

The above table shows the volume of each solution (in $\mu$l) added.

### 2.1.5. BACTERIAL STRAINS PLASMIDS AND PHAGES

The bacterial strains, plasmids, and phages used in this study are listed in Tables 2.1 and 2.2.

TABLE 2.1

BACTERIAL STRAINS

| STRAIN | GENOTYPE | SOURCE |
|---|---|---|
| *Escherichia coli* | | |
| JM83 | K12, *ara*, (*lac-pro*) *rpsL*, *thi*, (80d*lacI*$^{q}$Z M15). | N.P. Minton |
| JM101 | *supE*, *thi*, (*lac-proAB*), [F', *traD*36, *proAB*, *lacI*$^{q}$Z M15]. | N.P. Minton |
| W5445 | *pro*, *leu*, *thi*, *supE*, *lacY*, *tonA*, *hsdR*, *hsdM*, *rpsL*. | N.P. Minton |
| TG1 | K12, (*lac-pro*), *supE*, *thi*, *hsdD5*/F', *traD*36, *pro* A$^{+}$ B$^{+}$, *lacI*$^{q}$Z M15. | D.A. Barstow |
| TG2 | K12, (*lac-pro*), *supE*, *lacY*, *tonA*, *hsdD5*/F', *traD*36, *pro* A$^{+}$ B$^{+}$, *lacI*$^{q}$Z M15. | D.A. Barstow |
| BMH71-18 | K12, *lac-pro*, *supE*, *thi*/F', *pro* A$^{+}$ B$^{+}$, *lacI*$^{q}$Z M15. | D.A Barstow |
| DB11 | *thy*, *thi*, r$^{-}$, m$^{-}$, Em$^{s}$. | N.P. Minton |
| *Bacillus subtilis* 168 wt | | M. Young |
| *Clostridium pasteurianum* ATCC 6013 | | N.P. Minton |
| *Clostridium butyricum* NCIB 7423 | | N.P. Minton |
| *Clostridium acetobutylicum* NCIB 8052 | | M. Young |

## TABLE 2.2

PLASMIDS AND PHAGES

| PLASMIDS | MARKER | REFERENCE |
|---|---|---|
| pUC7 | $Ap^r$ | Vieira & Messing, 1982 |
| pUC8 | " | " " " |
| pUC9 | " | " " " |
| pMTL20 | " | Chambers *et al*., 1988 |
| pMTL21 | " | " " " |
| pAT153 | $Ap^r$, $Tc^r$ | Twigg & Sherratt, 1980 |
| pBD64 | $Cm^r$, $Km^r$ | Gryczan *et al*., 1980 |
| pR29 | $Ap^r$, $Em^r$ | Vasseghi & Claverias, 1983 |
| pAMB1 | $Em^r$ | Leblanc & Lee, 1884 |
| **PHAGES** | | |
| M13mp7 | | Messing & Vieira, 1982 |
| M13mp8 | | " " " |
| M13mp9 | | " " " |
| M13mp18 | | Yanisch-Perron *et al*., 1985 |
| M13mp19 | | " " " |

## 2.2. METHODS

### 2.2.1. STORAGE AND GROWTH OF BACTERIAL STRAINS

#### 2.2.1.1. E. COLI

*E. coli* strains were routinely grown aerobically, at 37°C, either in L-broth or 2 X YT broth. These strains were also cultivated on agar solidified (2% w/v) media of these broths. Cultures were stored in the short term at 4°C on L-agar plates. Longer term storage was at -70°C, as broth cultures supplemented with 10% glycerol.

#### 2.2.1.2. B. SUBTILIS

Strains of *B. subtilis* were grown aerobically at 37°C, either in L-broth or on L-agar plates supplemented with sporulation salts. Short term storage was on solidified L-agar at room temperature or at 4°C. Long term storage was at -70°C, as broth cultures supplemented with 10% glycerol.

#### 2.2.1.3. CLOSTRIDIA

*Clostridium* strains were grown anaerobically, at 37°C, either in reinforced clostridial medium (RCM) or tryptone, yeast extract, glucose (TYG) broth or on the solidified agar plates (2% w/v) of these broths. Agar plate cultures required the maintenance of a low oxygen tension. This was achieved by incubation in a sealed gas jar by using a commercial anaerobic gas generating kit (Gas Generating System, Oxoid Ltd, Basingstoke, UK). Broth cultures, however, required only the minimisation of head space air. In some instances pre-reduction of the media by autoclaving was performed.

Short term storage of clostridial strains was as sporulated colonies on RCM (Oxoid Ltd) agar plates in an oxygen-free atmosphere. Longer term storage was in the form of deep frozen slopes (RCM agar) which were covered with 80% (v/v) glycerol following overnight anaerobic growth. These slopes

were stored at -20°C.

### 2.2.2. TRANSFORMATION OF BACTERIAL STRAINS

#### 2.2.2.1. TRANSFORMATION AND TRANSFECTION OF COMPETENT E.COLI CELLS

Strains of E. coli were transformed and transfected (with plasmid and M13 RF DNA respectively) essentially as described by Cohen et al. (1972).

An overnight culture of the desired E. coli strain, grown in L-broth at 37°C, was used to inoculate 50 to 100 ml of prewarmed L-broth at a dilution of 1:50 in a 250 ml conical flask. The culture was incubated at 37°C with vigorous aeration and the $OD_{450}$ of the culture monitored at 15 to 30 min intervals. When the $OD_{450}$ reached 0.6 the flask was placed on ice for 20 min. 50 ml of the exponentially growing culture was centrifuged (Sorval RC5B, SS34, 5,000 r.p.m., 10 min) and the bacterial pellet washed in 25 ml of chilled (4°C) 0.1 M $MgCl_2$. Following centrifugation, the pellet was resuspended in 2.5 ml chilled (4°C) 0.1 M $CaCl_2$ and maintained on ice for a minimum of 2 hour prior to use.

Transformation was achieved by incubation of the transforming DNA (typically 50 to 500 ng of DNA in a maximum volume of 10 µl) with 100 µl of the competent cells on ice for 30 min, a 2 min heat shock at 42°C, followed by an additional incubation on ice for 30 min. Transformed cells were brought to room temperature, diluted in L-broth, and used to inoculate appropriate L-agar plates in the case of plasmid transformation, and soft agar overlays in the case of M13 RF DNA transfections.

#### 2.2.2.2. TRANSFORMATION OF NATURALLY COMPETENT B. SUBTILIS CELLS

Cells of B. subtilis 168 were transformed using the method of Anagnostopoulos and Spizizen (1961).

B. subtilis 168 cells were grown overnight on L-agar sporulation salts medium. A heavy inoculum of this fresh overnight growth was used to inoculate 20 ml of "competence" medium in a 250 ml conical flask. The culture was incubated with vigorous aeration, at 37°C, and the $OD_{600}$ monitored hourly until stationary phase was reached (an $OD_{600}$ of 3.0 or more). Later samples were diluted in Spizizens minimal medium + glucose prior to measurement of their absorbance. Maximal competence for transformation was displayed at the onset of stationary phase and was maintained at this high level for at least 60 min.

Cells were transformed by taking the early stationary phase culture and diluting it tenfold with transformation medium, again ensuring adequate aeration of the suspension by vigorous shaking at 37°C. The transforming plasmid DNA (typically 10 µl volume containing up to 5 µg DNA) was added to the suspension of cells and incubation continued for 90 min to allow phenotypic expression. Following this period the transformed cells were plated onto selective and non-selective media.

### 2.2.3. PLASMID AND M13 RF DNA ISOLATIONS

#### 2.2.3.1. LARGE SCALE ISOLATION OF PLASMID DNA FROM E. COLI

Plasmid DNA was purified from E. coli strains by the method of Clewell and Helsinki (1969).

One litre L-broth supplemented with appropriate antibiotics was inoculated with 10 ml of an overnight culture and incubated overnight at 37°C with moderate aeration (250 r.p.m. on a shaker platform). The cells were harvested by centrifugation (Sorval RC5B, GSA rotor, 6,000 r.p.m., 10 min) and resuspended in 10 ml of 25% sucrose; 50 mM Tris-HCl (pH 8.0). lysozyme, 1 ml of a 10 mg $ml^{-1}$ solution in 25 mM Tris-HCl (pH 8.0), was added and the mixture left on ice for 2 min, after which time 2 ml of 250 mM $Na_2EDTA$ was added. After a further 2 min on ice, 16 ml of lysing solution (1% (w/v) Brij 58; 0.4% sodium deoxycholate; 50

mM Tris-HCl (pH 8.0); 25 mM $Na_2EDTA$) was added and the solution mixed gently. After 5 min on ice or after lysis had occurred, the mixture was centrifuged (Sorval RC5B, SS34 rotor, 17,000 r.p.m., 30 min, 4°C) and the cleared lysate recovered by aspiration. The cleared lysate was then subjected to isopycnic centrifugation (2.2.4.).

2.2.3.2. SMALL SCALE ISOLATION OF PLASMID DNA FROM E. COLI

All small scale plasmid isolations were carried out using a modification of the rapid boiling method of Holmes and Quigley (1981).

A 1.5 ml sample of an overnight L-broth culture was centrifuged in an Eppendorf microcentrifuge (MSE Microcentaur, 1300 r.p.m., 2 min). The bacterial pellet was resuspended in 200 µl S.T.E.T. buffer by vortex mixing and 16 µl of a freshly prepared solution of lysozyme in water (10 mg $ml^{-1}$) was added. After brief vortex mixing the suspension was left for 5 min at room temperature and then placed in a boiling water bath for 40 sec. The suspension was immediately centrifuged (MSE Microcentaur, 13,000 r.p.m., 10 min) and the aqueous supernatant carefully removed from the gelatinous pellet and transferred to a fresh tube. An equal volume of isopropranol (100 to 200 µl) was then added, the tube contents briefly vortex mixed, and placed at -20°C for 1 hour. The plasmid DNA present was recovered from the precipitate by centrifugation (MSE Microcentaur, 13,000 r.p.m., 10 min, 4°C), the supernatant carefully decanted away, and the DNA pellet immediately washed with 200 µl 70% (v/v) ethanol. The washed DNA was again recovered by centrifugation (MSE Microcentaur, 13,000 r.p.m., 10 min, 4°C) and dried under vacuum. The dried DNA pellet was resuspended in 20 to 30 µl 1 X TE buffer; 10 µl of which was sufficient for visualisation by agarose gel electrophoresis following restriction endonuclease digestion (2.2.10.1.).

### 2.2.3.3. SMALL SCALE ISOLATION OF PLASMID DNA FROM B. SUBTILIS

All the Bacillus plasmid isolations performed were on a small scale; the method employed being the rapid alkaline extraction method of Birmboim and Doly (1979).

Samples (1.5 ml) of overnight broth cultures, grown in the presence of appropriate antibiotic, were centrifuged (MSE Microcentaur, 13000 r.p.m., 2 min) in a 1.5 ml Eppendorf tube. The supernatant was removed by aspiration and the bacterial pellet resuspended in 100 $\mu$l of solution I (2 mg $ml^{-1}$, 50 mM glucose, 10 mM $Na_2EDTA$, 25 mM Tris-HCl (pH 8.0). Following a 30 min incubation on ice, 200 $\mu$l of solution II (0.2 M NaOH, 1% (w/v) sodium dodecyl sulphate [SDS]) was added and the tube contents vortex mixed and incubated on ice for a further 5 min. To this mixture, 150 $\mu$l of solution III (3 M sodium acetate, pH 4.8) was added and the tube contents gently mixed by inversion. After incubation on ice for 60 min, the preparation was centrifuged (MSE Microcentaur, 13,000 r.p.m., 5 min) and the supernatant removed to a fresh tube. The plasmid DNA present was precipitated by the addition of 1 ml chilled absolute ethanol (-20$^{o}$C) followed by incubation at -70$^{o}$C for 15 min or at -20$^{o}$C for 1 hour. The DNA was recovered by centrifugation (MSE Microcentaur, 13,000 r.p.m., 10 min, 4$^{o}$C), briefly washed with 200 $\mu$l 70% ethanol, and pelleted again. The DNA pellet was dried in a vacuum and resuspended in 50 $\mu$l 1 X TE buffer.

### 2.2.3.4. LARGE SCALE ISOLATION OF PLASMID DNA FROM C. BUTYRICUM

A 2 litre batch culture of the organism was grown anaerobically in TYG medium and the cells harvested by centrifugation (Sorval RC5B, GSA rotor, 10,000 r.p.m., 10 min, 4$^{o}$C). The cells were washed with 200 ml 1 X TES buffer and resuspended in 40 ml 50 mM Tris-HCl (pH 8.0) containing 25% (w/v) sucrose. The suspension was placed on ice for 15 min and then 4 ml lysozyme (10 mg $ml^{-1}$, in 25 mM Tris-HCl, pH

8.0) was added. After vortex mixing, the suspension was held on ice for 15 min and 16 ml 250 mM $Na_2EDTA$, pH 8.0 was added. The suspension was gently mixed and held on ice for a further 15 min. Lysis was effected by adding 40 ml 2.5% (w/v) SDS in 50 mM Tris-HCl; 25 mM $Na_2EDTA$, pH 8.0. After gentle mixing the suspension was held on ice for a further 30 min and 25 ml 5 M NaCl was added. The lysate was then centrifuged (Sorval RC5B, SS34 rotor, 17,000 r.p.m., 30 min, 4°C) and the cleared lysate carefully aspirated and transferred to a fresh tube. The cleared lysate was then subjected to isopycnic centrifugation (2.2.4.).

2.2.3.5. LARGE SCALE ISOLATION OF M13 RF DNA

Single stranded M13 template DNA was used (1 μl) to transfect competent E. coli JM101. Immediately following the 42°C heat shock the transformed cells were used to inoculate 10 ml 2 X YT broth which was incubated overnight at 37°C with vigorous aeration. The culture was centrifuged (Sorval RC5B, SS34 rotor, 6,000 r.p.m., 10 min) and the supernatant decanted into a fresh tube. This "phage inoculum" was temporarily stored at 4°C. Meanwhile, 25 ml of an overnight culture of E. coli was used to inoculate 1 litre of prewarmed 2 X YT broth. This culture was incubated at 37°C with vigorous aeration until the $OD_{600}$ reached 0.6. The 10 ml "phage inoculum" was then added and the culture incubated for a further 4 hour at 37°C with vigorous aeration. The cells were harvested by centrifugation (Sorval RC5B, GSA rotor, 6,000 r.p.m., 10 min) and a cleared lysate prepared as described in section 2.2.3.1. The cleared lysate was then subjected to isopycnic centrifugation (2.2.4.)

**2.2.4. ISOPYCNIC CENTRIFUGATION OF PLASMID AND M13 RF DNA**

Plasmid DNA was separated from chromosomal DNA by centrifugation to equilibrium in caesium chloride-ethidium bromide density gradients (isopycnic centrifugation) as follows.

One gram of solid caesium chloride and 0.1 ml of ethidium

bromide solution (10 mg $ml^{-1}$ in distilled water) were added to each 1 ml volume of cleared lysate. After gentle mixing by inversion, the contents were transferred to a Du Pont/Sorval "quick-seal" crimp tube and the tube completely filled and sealed by "crimping" with a Du Pont/Sorval tube crimping system. If insufficient gradient mix was available the "quick-seal" tubes were topped up with paraffin oil and balanced appropriately. The tubes were centrifuged either in a Sorval Ti50 fixed angle rotor at 40,000 r.p.m. (10 ml tubes) for 36 to 48 hour (20°C), or in a Sorval TV 865 vertical rotor at 48,000 r.p.m. (5 ml tubes) for 12 to 18 hour (20°C). After centrifugation, the plasmid or M13 RF (lower) and chromosomal (upper) DNA bands were visualised with a UV long wave transilluminator (Blak-ray, Ultraviolet products Inc., San Gabriel, Ca, U.S.A.). The lower band was removed by piercing the side of the tube with an 18-gauge hypodermic needle 0.5 cm beneath the band and withdrawing it into a 2 ml syringe; an additional needle was used to pierce the top of the tube thereby facilitating a smooth removal of the desired band. Ethidium bromide was removed by extracting the sample 3 to 4 times with isoamyl alcohol which had been equilibrated with CsCl. The sample was then transferred to prepared 1/4" dialysis tubing and the CsCl removed by dialysis against three changes of 1 X TE buffer at 4°C. At least 1,000 volumes of dialysis buffer were used and the buffer changed after 1, 2, and 4 hour of dialysis.

The degree of protein contamination in the sample and the concentration of the purified plasmid DNA was determined by measuring the optical density at 260 and 280 nm using quartz cuvettes. Protein contamination was considered to be negligible if the 260/280 nm ratio was 1.7 or greater. When necessary contaminating protein was removed by extraction with phenol (2.2.16). An optical density of 1.0 at 260nm was taken to be equivalent to a DNA concentration of 50 $\mu g$ $ml^{-1}$ (Maniatis <u>et al</u>., 1982).

#### 2.2.4.1. PREPARATION OF DIALYSIS TUBING

The dialysis tubing was cut into pieces of convenient length (10-20 cm) and boiled for 10 min in a large volume of 2% (w/v) sodium bicarbonate and 1 mM $Na_2EDTA$. After brief rinsing in distilled water the tubing was boiled in distilled water for a further 10 min. The tubing was allowed to cool and stored submersed at 4°C. Immediately prior to use the tubing was washed with sterile distilled water.

### 2.2.5. CHROMOSOMAL DNA ISOLATIONS

The isolation of genomic DNA from both E. coli JM83 and C. pasteurianum was performed by a slight variation of the method of Marmur (1961).

One litre overnight cultures of E. coli JM83 and C. pasteurianum were grown in L-broth (aerobically) and TYG broth (anaerobically) respectively. The cells were harvested by centrifugation (Sorval RC5B, GSA rotor, 6,000r.p.m., 10 min), washed in saline-EDTA, recentrifuged, and the wet weight of each harvest determined. The cells were resuspended in saline-EDTA at a density equivalent to 10 ml saline-EDTA per gram of cells (wet weight). Lysozyme solution (10 mg $ml^{-1}$ in distilled water) was added to the cell suspensions at a dilution of 1ml per 25 ml of cells and the mixtures incubated at 37°C for 30 min and 3 hour for E. coli and C. pasteurianum respectively. Hot SDS (60°C, 25% w/v) was added to a final concentration of 2% and the contents, in a conical flask, incubated at 60°C for 15 min with intermittent mixing. The mixture was then cooled to ambient temperature and 5 M sodium perchlorate added to a final concentration of 1 M. An equal volume of chloroform:isoamylalcohol (24:1 v/v) was added and the emulsion gently shaken in a 10°C water bath for 30 min. The emulsion was separated by centrifugation (Sorval RC5B, GSA rotor, 10,000 r.p.m., 5 min, 10°C). The aqueous phase (top layer) was carefully removed with a "cut-off" (i.e., wide bore) 10 ml disposable pipette and transferred to a

chilled (-20°C) glass measuring cylinder. Two volumes of chilled (-20°C) absolute ethanol were gently added and the DNA, precipitating at the ethanol/aqueous interface, spooled onto a sterile glass rod, set aside and left to drain. The DNA was then resuspended in a volume of 0.1 X SSC equivalent to 75 % of the original volume and the SSC concentration adjusted to 1 X SSC with 10 X SSC. An equal volume of chloroform:isoamylalcohol (24 : 1 v/v) was again added, the solution extracted, the DNA precipitated and spooled onto a glass rod, exactly as before. The DNA was again resuspended in 0.1 X SSC, this time in a volume equivalent to 50% of the original volume, and the SSC concentration adjusted to 1 X SSC with 10 X SSC. DNAase-free RNAase (2.2.6.6.) was then added to a final concentration of 50 $\mu g\ ml^{-1}$ and the solution incubated for 1 hour at 37°C. SDS (25% w/v) was added to a final concentration of 0.5%, followed by the addition of proteinase K (20 mg $ml^{-1}$ in distilled water) to a final concentration of 1 mg $ml^{-1}$. The solution was incubated for a further 60 min at 37°C and then extracted three times with an equal volume of chloroform:isoamylalcohol (24 : 1 v/v). The DNA solution was then subjected to extraction with a half volume of phenol (2.2.16.) by gentle shaking at ambient temperature for 10 min. After separation of the phenol and aqueous phases by centrifugation (Sorval RC5B, SS34 rotor, 10,000 r.p.m., 5 min) the remaining traces of phenol were removed by 3 extractions with equal volumes of water saturated diethyl ether. The aqueous phase was recovered and the DNA again precipitated with two volumes of chilled (-20°C) absolute ethanol, the DNA spooled onto a glass rod and the DNA eventually resuspended in a volume of 1X SSC equivalent to 25% of the original volume. The DNA was finally precipitated by the addition of $1/10^{th}$ volume of acetate-EDTA (3.0 M sodium acetate; 0.001 M $Na_2EDTA$; pH 7.0) and the gradual addition of a 0.54 volume of isopropranol whilst simultaneously spooling the DNA onto a sterile glass rod. The spooled DNA was then progressively washed in 70%, 80%, 90%, and finally 100% ethanol, dried under vacuum, and resuspended in a minimum volume of 1 X TE buffer. The DNA concentration was then determined spectrophotometrically by

reading the optical density of the solution at 260 nm and the DNA solutions stored at -20°C.

### 2.2.6. TREATMENT OF DNA WITH ENZYMES

#### 2.2.6.1. RESTRICTION ENDONUCLEASES

Restriction endonucleases were purchased from Boehringer Mannheim (BCL) or Bethesda Research Laboratories (BRL). The reaction conditions were in accordance with manufacturers' instructions, except that a two to tenfold excess of enzyme was used routinely. The digestion products were analysed directly by agarose gel electrophoresis. If the restricted DNA was to be ligated the remaining restriction enzyme activity was destroyed by either heat inactivation at 70°C for 10 min (for those heat labile enzymes), or by phenol extraction (2.2.16.).

#### 2.2.6.2. T4 POLYNUCLEOTIDE LIGASE

DNA fragments were typically ligated at a final concentration of 30-50 $\mu g\ ml^{-1}$ in 1 X "optimised Ligation buffer" (Lathe et al.. 1984) which was added as a tenfold concentrate to the reaction mixture. T4 DNA ligase was added at a final concentration of 0.1 to 1.0 units $\mu g^{-1}$ of DNA and the reaction mixture, in a 0.5 ml Eppendorf tube, incubated overnight at 15°C.

In some instances, aliquots of the ligation samples were analysed on agarose mini-gels for the presence of new molecular species which was an indication that in vitro ligation of the DNA fragments had occurred.

#### 2.2.6.3. T4 DNA POLYMERASE

T4 DNA polymerase was used to convert "sticky-ended" DNA to "blunt-ended" DNA, particularly with DNA fragments possessing recessed 5' termini, by the method of O'Farrell (1981).

The DNA sample was diluted in an appropriate volume (20 to 50 $\mu$l) of 1 X T4 polymerase buffer supplemented with a one-tenth volume of dNTPs (0.25 mM each) and 1 $\mu$l T4 DNA polymerase. The reaction mixture was incubated at 37$^{o}$C for 30 min and the enzyme subsequently inactivated by heat treatment (70$^{o}$C/10 min). Alternatively, the treated DNA was then purified by phenol extraction and ethanol precipitation (2.2.16.).

### 2.2.6.4. DNA POLYMERASE, LARGE FRAGMENT (KLENOW FRAGMENT)

Klenow polymerase was used in reactions where sequential base additions to recessed 3' termini were required. These included the conversion of "sticky-ended" DNA to "blunt-ended" DNA, the end-labelling of HindIII digested Lambda DNA, and in the chain elongation reaction of DNA sequencing (2.2.17.5.). All such reactions were carried out in 1 X TM buffer supplemented with a one-tenth volume of 0.25 M dNTPs, and usually 1 to 2 $\mu$l of Klenow polymerase. Reactions were typically carried out for 30 min at room temperature. When necessary, remaining polymerase activity was inactivated by heat treatment (70$^{o}$C/10 min).

### 2.2.6.5. CALF INTESTINAL PHOSPHATASE (CIP)

Calf intestinal phosphatase was used to dephosphorylate the 5'-terminus of DNA. Following digestion of the DNA with the appropriate restriction endonuclease, CIP (BCL) was generally added directly to the reaction tube, at a rate of 1 to 5 units of CIP $\mu g^{-1}$ of DNA, and the tube incubated for a further 30 to 60 min at 37$^{o}$C. Alternatively, the digested DNA was ethanol precipitated and resuspended in an appropriate volume of 1 X CIP buffer supplemented with CIP at a concentration of 1 unit $\mu g^{-1}$ DNA. The reaction was then incubated for 60 min at 37$^{o}$C. Remaining CIP was removed by phenol extraction and subsequent ethanol precipitation (2.2.16.).

2.2.6.6. DNAase-FREE RNAase

DNAase-free RNAase was used for the selective removal of RNA in DNA samples. Stock solutions of RNAase A (10 mg $ml^{-1}$ in 10 mM tris-HCl, pH 7.5) were boiled (100°C) for 15 min prior to use or storage at -20°C.

DNA samples were typically treated with RNAase at a final concentration of 25-50 $\mu$g $ml^{-1}$ by incubation at 37°C for 1 hour. When necessary, remaining enzyme was removed by phenol extraction (2.2.16.).

## 2.2.7. ESTABLISHMENT OF CONDITIONS FOR PARTIAL DIGESTION OF CHROMOSOME DNA WITH RESTRICTION ENDONUCLEASES

This method was used to obtain a partial digestion of C. pasteurianum genomic DNA with MboI.

A reaction mixture was prepared which contained 50 $\mu$g C. pasteurianum genomic DNA in the appropriate restriction enzyme buffer in a final volume of 150 $\mu$l. This mixture was aliquoted into Eppendorf tubes as follows: 30 $\mu$l into tube 1, 15 $\mu$l into tubes 2 to 8, and the remainder into tube 9. All the tubes were chilled on ice for 10 min and then 20 units of restriction enzyme was added to tube 1 and gently mixed. This gave an enzyme concentration of 2 units $\mu g^{-1}$ DNA. Half of this (15$\mu$l) was transferred to tube 2, and so on serially diluted down the line to tube 8. Nothing was added to tube 9. Tubes 1 to 8 were then incubated at 37°C for 60 min. The reactions were stopped by chilling on ice and the addition of 1 ul of 0.5 M $Na_2EDTA$. Tracking dye (3 $\mu$l) was added to all nine tubes and the samples electrophoresed through 0.8% agarose (2.2.10.1.). Subsequent photography of the gel revealed the appropriate restriction enzyme concentration required to achieve partial digestion of the DNA.

### 2.2.8. RADIO-LABELLING OF DNA

#### 2.2.8.1. NICK-TRANSLATION OF DOUBLE-STRANDED DNA

Purified double stranded DNA was labelled to a high specific activity ($5.10^8$ cpm $\mu g^{-1}$ DNA) by the nick-translation reaction described by Rigby et al.. (1977).

Routinely, between 100 ng and 1 $\mu g$ of DNA was added to 10 $\mu l$ [alpha-$^{32}P$] dATP (~3,000 Ci m $mol^{-1}$) in an Eppendorf tube on ice. The reaction volume was made up to 48.5 $\mu l$ with distilled water following the addition of 5 $\mu l$ 10 X NT buffer and 5 $\mu l$ of dNTP mix (1 mM dTTP, 1 mM dGTP, 1mM dCTP, and 30 $\mu M$ dATP). Prior to commencing the reaction, 5 $\mu l$ DNAase I (1 mg $ml^{-1}$ in 10 mM HCl) was diluted 10-fold and incubated on ice for 2 hour. The reaction was started by the simultaneous addition of 1 $\mu l$ of the preincubated DNAase I and 2 $\mu l$ of Klenow polymerase. The reaction mixture was incubated for 2 to 3 hour at 15°C, after which time it was stopped by the addition of 2 $\mu l$ 0.5 M $Na_2EDTA$. Unincorporated $^{32}P$ dATP was removed from the sample by gravitational passage down a 0.8cm X 10cm Sephadex G-50 column equilibrated with 1 X TE buffer. A drop of a 10% solution (w/v) of Orange G dye was added to the sample prior to loading on the column. After allowing the sample to completely enter the column matrix the top reservoir was filled with 1 X TE buffer. The progress of the labelled DNA was monitored with a hand held Geiger counter and the relevant fraction, the fastest migrating species, collected in a 1.5 ml Eppendorf tube.

#### 2.2.8.2. 5'-END LABELLING OF OLIGONUCLEOTIDES

Oligonucleotide probes were end-labelled by the addition of [gamma-$^{32}P$] dATP to the 5'-hydroxyl terminus with T4 polynucleotide kinase (Maxam & Gilbert, 1977).

Typically, 50 $\mu l$ [gamma-$^{32}P$] dATP (3000Ci m $mol^{-1}$; 50% ethanol) was completely evaporated in a 1.5 ml Eppendorf

tube under vacuum. To this, 100ng oligonucleotide probe, 2$\mu$l of 10 X kinase reaction buffer, and 1 $\mu$l 100 mM DTT was added and the reaction mixture made up to a total volume of 20 ul with distilled water. The mixture was incubated at 37$^{o}$C for 30 to 60 min. The reaction was stopped by the addition of 1 $\mu$l 0.5 M $Na_2EDTA$ and kept on ice or frozen at -20$^{o}$C prior to use. It was not usually considered necessary to remove excess label by chromatography through Sephadex G-50 as the oligonucleotide probe was usually in a 1.5 molar excess over dATP.

### 2.2.8.3. "REVERSE PRIMER" LABELLING OF M13 SINGLE STRANDED TEMPLATE DNA

The method employed was based on that described by Hu and Messing (1982). The method involves the use of a synthetic 13 base oligonucleotide (5'-GAAATTGTTATCC-3') which is complementary to the 5'-side of the polylinker region of M13mp (+) strand DNA. The primer is used to direct the synthesis of the complementary strand of the viral DNA from any M13mp vector containing an insert. The synthesised strand is labelled and used as a hybridisation probe for the inserted sequence.

The oligonucleotide primer (2 $\mu$l of a 2 ng $\mu l^{-1}$ solution) was mixed with 5 $\mu$l of the M13 template DNA, in a total volume of 20 $\mu$l of 1 x TM buffer. The mixture was annealed exactly as described for the annealing of sequence universal primer (2.2.17.2.). 4 $\mu$l of A$^{o}$ (dTTP, dCTP, dGTP; 0.25 mM each), 5 $\mu$l [alpha-$^{32}$P] dATP, 2$\mu$l Klenow polymerase, and 13 $\mu$l distilled water was added and the reaction carried out at room temperature for 30 to 60 min. The reaction was then terminated by the addition of 1 $\mu$l of 0.5 M $Na_2EDTA$. The probe was kept on ice prior to use.

### 2.2.8.4. END-LABELLING OF LAMBDA/HINDIII DNA SIZE MARKERS

Lambda DNA which had previously been digested with HindIII was end-labelled in the following manner. Approximately 10

ug of DNA was incubated (room temp/30 min) with 2 $\mu$l Klenow polymerase in the presence of 0.25 mM dNTPs (dTTP, dGTP, dCTP) and 10 $\mu$Ci (1 $\mu$l) alpha-$^{32}$P dATP, in a total volume of 100 $\mu$l 1 X TM buffer. Aliquots (3-10 $\mu$l) were used directly for electrophoresis (2.2.10.1.).

### 2.2.9. SCINTILLATION COUNTING OF LABELLED DNA

Typically 1% of the labelled DNA probe, which had been desalted by Sephadex G-25 or G-50 chromatography, was transferred to a glass scintillation vial. Using a Rackbeta scintillation counter, Cerenkov radiation was counted on a tritium program. This enabled the calculation of the specific activity of the probe (cpm $\mu g^{-1}$ DNA).

### 2.2.10. ELECTROPHORESIS

#### 2.2.10.1. AGAROSE GEL ELECTROPHORESIS

This was performed with either vertical or horizontal slab gel electrophoresis tank systems. The horizontal gel systems employed were either of the Schaffner type (Maniatis et al., 1982), in which the gel is poured into a portable gel casting tray (20 X 20 cm gels), or of the "mini-gel" type where the gel is cast directly into the electrophoresis apparatus (5 X 8 cm). The apparatus was obtained from BRL (H4 system) and Cambridge Life Sciences respectively. The vertical gel system used was an "in-house" design where the gel was cast between two glass plates sealed with 3 or 5 mm perspex spacers. The gel routinely used consisted of agarose (BRL ultra-pure) at concentrations between 0.8 and 1.5% containing ethidium bromide at 0.5 $\mu$g ml$^{-1}$ in 1 X TBE buffer. The gel was formed by boiling together (in a microwave oven) appropriate amounts of agarose, 10 X TBE buffer, distilled water, and ethidium bromide solution (10 mg ml$^{-1}$ in distilled water). The molten gel was allowed to cool to 60$^{o}$C before pouring; 35 ml, 250 ml, and 100 ml was routinely used for the horizontal mini gel, the analytical/preparative (20 X 20 cm) horizontal gel, and the vertical gel respectively. Perspex combs of varying teeth

size were inserted into the molten gels, prior to setting, to create the loading wells. Samples to be electrophoresed were mixed with a one-fifth volume of tracking dye (50% ficoll, 0.1% bromophenol blue) and loaded into the wells of the gel submersed in 1 X TBE gel running buffer. The gel was then subjected to electrophoresis to give a voltage across the gel of between 1 and 5 V $cm^{-1}$. After electrophoresis the DNA was visualised by medium wave UV illumination and photographed with Polaroid 667 black and white film (ASA 3000) using a red filter. Exposure times of between 1 and 4 sec at an aperture setting of F8 were usually sufficient.

### 2.2.10.2. DENATURING POLYACRYLAMIDE GEL ELECTROPHORESIS

Denaturing polyacrylamide electrophoresis was routinely used for the resolution of DNA sequence reaction products and for the analysis and purification of synthetic oligonucleotides. For the size fractionation of DNA sequencing reaction products the thin gel system of Sanger and Coulson (1978) was used. The gel plates (20 X 50 cm), one of which had "rabbit ears" (the front plate) were cleaned thoroughly with pyroneg in solution followed by absolute ethanol. The front plate was then treated with "replicoat" (2% dimethyldichlorosilane in 1,1,1- trichloroethane). Simultaneously the back plate was treated with a solution comprised of 750 µl 10% glacial acetic acid and 75 µl silane (gamma-[methylacryloxy]-propyltrimethoxysilane) in 25 ml absolute ethanol. After allowing the plates to dry, they were polished with soft paper tissue to remove excess coating. The gel "sandwiches" were then assembled with 0.35 mm plastic card spacers (Raven Scientific Ltd) using "nitto" tape (BRL). The "gradient" gels routinely used were made from 40 ml 0.5 X TBE gel mix and 7 ml 5 X TBE gel mix. The gel mixes were polymerised by the addition of 25 % (w/v) ammonium persulphate and TEMED to 0.2% (v/v) each. The gradients were poured in the following manner. Using a 10 ml pipette, 6 ml of the clear 0.5 X TBE gel mix was taken up, followed by the 7 ml of blue 5 X TBE gel mix. After allowing the passage of two air bubbles through the

pipette to mix the contents slightly, the contents of the pipette was carefully poured into the gel sandwich followed by the remaining 34 ml 0.5 X TBE gel mix. A plastic card gel comb was placed in the top of the gel and the gel mix left to polymerise. Up to a maximum of 48 (12 X 4) samples could be loaded on one gel.

For the analyses and purification of oligonucleotides, 20% acrylamide, 50% urea, 1 X TBE gels were routinely used. These were poured into gel sandwiches as already described except that 1 mm spacers and combs were used, and that both glass plates (20 X 50 cm) were treated with "replicoat". The acrylamide/urea mixture was made up by gently melting the appropriate volume of 40% acrylamide stock with the urea and 10 X TBE and adjusting the final volume with distilled water. The gel was polymerised by the addition of 25% (w/v) ammonium persulphate solution and TEMED to 0.2% (v/v). Once poured and set the gels were usually pre-run at 600 V for 30 min.

2.2.10.3. SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS

Gradient (10-30%) gels were used for the size-fractionation of the products of in vitro transcription/translation assays (2.2.21.).

The acrylamide gradient was produced by the mixing of "light" and "dense" solutions in a gradient mixer. The dense solution consisted of 18 ml 40% (w/v) acrylamide stock solution, 3.8 ml 3% (w/v) bis-acrylamide solution and 4.5 ml lower Tris. The light solution consisted of 5.4 ml 40% (w/v) acrylamide stock solution, 2.3 ml 3% (w/v) bis-acrylamide solution, 4.5 ml lower Tris, and distilled water to give a final volume of 15 ml. Both solutions were polymerised by addition of 18 $\mu$l 25 % (w/v) ammonium persulphate and 14 $\mu$l TEMED. Once the gel had been poured it was overlaid with 30% methanol, 25% lower Tris (v/v). Following polymerisation, the overlay solution was removed and the top of the gel rinsed with water. The stacking gel was then poured directly on top of the gradient gel. This

consisted of 1.1 ml 40% acrylamide stock solution, 2.5 ml upper Tris, made up to 10 ml with distilled water, and was polymerised by the addition of 50 $\mu$l 25% (w/v) ammonium persulphate and 85 $\mu$l TEMED.

Typically an equal volume of SDS PAGE sample buffer was added to the protein sample which was then boiled for 3 min prior to loading into the wells of the stacking gel submersed in 1 X SDS PAGE running buffer. The gradient gel was run at a constant current of 40 mA for 8 hour. The gel was then fixed by a 15 min immersion in fixing solution (25% v/v isopropranol, 10% v/v acetic acid) followed by 5-10 min immersion in Amplify solution. The gel was then dried down in a Biorad gel drier under vacuum at 65$^{o}$C for 2-3 hour prior to autoradiography (2.2.13.4.).

### 2.2.11. RECOVERY OF DNA FROM AGAROSE GELS BY ELECTROELUTION

DNA was extracted from agarose gels essentially as described by McDonnell et al.. (1977).

After resolution of the desired DNA fragment by agarose gel electrophoresis and staining with ethidium bromide, the band was localised with the aid of a long wave UV transilluminator (Blak-ray, Ultra-Violet Products Inc., San Gabriel, Ca, U.S.A.). The desired DNA band was excised from the gel with a scalpel and placed in an appropriate length of 1/4" prepared dialysis tubing and submerged in 0.05 X Tris-borate buffer. The ends of the tubing were sealed and the dialysis bag placed lengthways at right angles to the current in a horizontal gel tank partially filled with 0.05 X Tris-borate buffer. Electroelution was carried out at 240 V for 2 hour or at 40 V overnight. After this time the current was reversed for several second to disengage the eluted DNA from the dialysis tubing. The Tris-borate buffer, containing the eluted DNA, was removed from the tubing and extracted once with isoamyl alcohol, once with phenol, four times with water saturated diethyl-ether, ethanol precipitated, and resuspended in an appropriate volume of 1 X TE buffer (2.2.16.).

## 2.2.12. "SOUTHERN" TRANSFER OF DNA

### 2.2.12.1. TRANSFER TO NITROCELLULOSE MEMBRANE

DNA was transferred from agarose gels to nitrocellulose membrane by the method of Southern (1975).

After photographing the stained agarose gel, the DNA was nicked to facilitate transfer of the larger DNA fragments. This was effected either by irradiation for 15 min with a short wave UV transilluminator (Chromato-vue, Transilluminator model C-61) or, more routinely, by partial depurination by submerging the gel in two changes of 0.25 M HCl for 10 to 15 min (Wahl et al., 1979). The gel was subsequently soaked twice in 250 ml of denaturing solution (0.5 M NaOH, 1 M NaCl) for 15 min each, followed by a 10 min wash in distilled water. The gel was then soaked in neutralising solution (0.5 M Tris-HCl (pH 7.5), 3 M NaCl) for 60 min. After a 10 min wash in 2 X SSC buffer the gel was placed on a "wick" of Whatman 3MM filter paper (soaked in 20 X SSC) draped over a glass plate resting on top of a plastic tray containing 1 litre of 20 X SSC transfer solution. The ends of the wick were immersed in this buffer. The edges of the gel were sealed from contact from the wick by sliding strips of X-ray film underneath. A sheet of nitrocellulose membrane, soaked in 2 X SSC for 10 min, was then placed on top of the gel, care being taken not to trap air bubbles between the two surfaces. Three sheets of Whatman 3MM filter paper, cut to the size of the gel and again soaked in 2 X SSC, were placed on top, and then a wad (10 cm) of blotting paper placed on top. A small weight (approximately 250-500 g) was placed on top to ensure an even contact between the gel and the membrane. Transfer was allowed to proceed for at least 18 hour, after which time, the membrane was removed, washed briefly in 2 X SSC, dried at ambient temperature and then baked at 80°C for 2 hour under vacuum. The gel was restained with ethidium bromide and photographed to visualise the extent of transfer of DNA.

2.2.12.2. TRANSFER TO NYLON MEMBRANE

The transfer of DNA from agarose gels to "zeta probe" nylon membrane (Biorad Laboratories) was carried out essentially as for transfer to nitrocellulose membrane with some modifications (Reed and Mann, 1985).

The transfer solution used was 0.4 M NaOH. Following depurination the gel was briefly rinsed in distilled water and immediately positioned on the Whatman 3MM "wick" system described previously. Transfer was carried out for 2 to 12 hour and the membrane rinsed in 2 X SSC and air dried prior to use. Baking of the membrane was unnecessary as the alkaline solvent induces covalent fixation of DNA to the membrane (Reed & Mann, 1985).

**2.2.13. DNA-DNA HYBRIDISATIONS AND AUTORADIOGRAPHY**

2.2.13.1. IN SITU COLONY HYBRIDISATION

The method employed was a modification of that described by Grunstein and Hogness (1975).

Bacterial colonies to be screened were picked and "streaked" (2 to 3 mm in length) onto the ordered grid pattern of 9 cm nitrocellulose filter discs (Schleicher & Schuell) which had been placed on the surface of agar plates containing the selective antibiotic. Simultaneously the colonies were streaked in identical positions on agar master plates also supplemented with the selective antibiotic. Up to 100 colonies could be streaked on a single filter. Following overnight incubation at 37°C the master plates were stored at 4°C and the filters peeled off the replicas. The following steps were performed by placing the filters, colony sides up, onto trays containing 3 layers of Whatman 3MM filter paper soaked in the appropriate solution. First, the filters were treated with 0.5 M NaOH for 10 min, followed by 3 treatments with 1 M Tris-HCl (pH 7.4) for 2 min each. Finally the filters were treated with 1.5 M NaCl; 0.5 M Tris-HCl (pH 7.4) for 5 min. The

filters were then immersed in 1 X SSC, supplemented with 1 mg $ml^{-1}$ proteinase K and 0.5% (w/v) SDS, and incubated at 37°C for 1 hour. The filters were then rinsed in 2 X SSC, air dried, and baked in a vacuum at 80°C for a minimum period of 30 min. The filters were stored at ambient temperature until used in a hybridisation experiment. The filters were stored for periods up to 3 months in this way. The hybridisation conditions used varied with the type of labelled DNA probe used as outlined in section 2.2.13.2.

2.2.13.2. DNA-DNA HYBRIDISATIONS

Hybridisation conditions employed depended on the type of DNA probe being used. "Nick-translated" DNA, "oligo-labelled" DNA, and labelled M13 template DNA were hybridised to DNA immobilised on either nitrocellulose or nylon membrane as outlined in (I). In contrast, when a $^{32}P$ end-labelled oligonucleotide probe (21 bases or less in length) was used the conditions were as described in (II). The amount of labelled probe used varied from $10^5$ to $10^7$ cpm $ml^{-1}$ hybridisation solution.

(I) The membrane was wet in 2 X SSC and then transferred to a plastic bag. A sufficient volume (50 ml or more) of prehybridisation solution was added and the plastic bag sealed with a Calor heat sealer with the careful exclusion of all air bubbles. The prehybridisation solution consisted of 50% formamide, 5 X SSC, 5 X Denhardts reagent, 0.1% SDS (w/v), 100 $\mu g$ $ml^{-1}$ sheared denatured (100°C/5 min) salmon sperm DNA, and 1 mM $Na_2EDTA$. prehybridisation was carried out at 37°C for a minimum period of 1 hour. After this time the majority of the pre-hybridisation solution was removed from the bag by making a small incision in one corner, leaving some 10 to 20 ml of solution remaining. The labelled probe DNA (which varied in amount from as little as 10 ng to 1 $\mu g$ depending on the specific activity) was then carefully added to the bag, the contents gently mixed, and the bag resealed after the careful removal of any air bubbles. If the probe consisted of double stranded DNA it was first denatured in a boiling water bath for 5 min.

Hybridisation was carried out at 37°C for at least 16 hour. After hybridisation, the membrane was washed in several changes of 2 X SSC, 0.1% (w/v) SDS, at 65°C, 15 min each. After a final 45 min wash in 0.5 L of 0.1 X SSC, 0.1% (w/v) SDS, 65°C, the membrane was rinsed in 2 X SSC at room temperature. The membrane was then air dried and autoradiographed (2.2.13.4.).

(II) The membrane was wet in 2 X SSC and then placed in a sealable perspex box and an ample volume of prehybridisation solution added (50 to 100 ml). The prehybridisation solution consisted of 6 X SSC, 10 X Denhardts reagent, 0.2% (w/v) SDS , and 100 $\mu g\ ml^{-1}$ sheared heat denatured (100°C/5 min) salmon sperm DNA. Prehybridisation was carried out at 65°C for 15 to 30 min. The 5'-end-labelled oligonucleotide probe was then added to a final concentration of 1 ng $ml^{-1}$ (with a specific activity approximating $10^7$ cpm $ml^{-1}$). Hybridisation was carried out at the calculated dissociation temperature for the oligonucleotide less 5°C ($T_d$-5 °C). The $T_d$ was defined to be the temperature at which 50% of the duplexes dissociate in 6 X SSC (Wallace et al., 1979) and calculated as follows:

$$T_d(^{o}C) = 4(G+C) + 2(A+T)$$

Hybridisation was typically carried out for 2 hour or until sufficient time had elapsed to enable the probe to reach a calculated value of 1 to 3 X $Cot_{1/2}$; $Cot_{1/2}$ is the calculated time required for half-renaturation of the probe (Maniatis et al., 1982) and was calculated as follows:

$$1/X \times Y/5 \times Z/10 \times 2 = \text{number of hour to achieve } Cot_{1/2}$$

where,

X = The weight of probe added (in $\mu g$).
Y = The complexity of the probe (length of probe in kb).
Z = The volume of the hybridisation reaction (in mls).

Following hybridisation the membrane was washed with several changes of 6 X SSC, 0.1% (w/v) SDS, for 5 min each at the hybridisation temperature (unless otherwise stated). The membrane was finally rinsed with 6 X SSC at room temperature and either wrapped in saran wrap or air dried prior to autoradiography (2.2.13.4.).

### 2.2.13.3. REHYBRIDISATION OF MEMBRANES

If a previously used membrane was to be rehybridised with a different probe, the bound probe was removed by several washes with boiling 0.1 X SSC, 0.5% (w/v) SDS, for 5 min each. The membrane was finally rinsed in 2 X SSC at room temperature and air dried. The membrane was then examined by autoradiography to ensure that all the probe had been removed before proceeding with the rehybridisation.

### 2.2.13.4. AUTORADIOGRAPHY

Routine autoradiography was carried out with Kodak X-Omat S, X-ray film which was positioned on top of the membrane or immobilised gel inside a light proof cassette. This arrangement was left undisturbed for periods from 1 to 72 hour at room temperature. When necessary, the sensitivity of the autoradiography was increased by the following methods. One or two calcium-tungstate-phosphor intensifying screens (Du Pont Cronex Lightning-Plus) were used in the assembly in the cassette (Swanstrom & Shank, 1978) and the autoradiography carried out at -70$^{o}$C. Alternatively, the film was sensitised by "preflashing," with a filtered photographic flashgun, prior to autoradiography (Laskey and Mills, 1977). Finally, films with greater sensitivity were employed, such as Kodak XAR5 and Amersham's Betamax film, and autoradiography carried out at room temperature.

## 2.2.14. QUANTITATION OF DNA

### 2.2.14.1. SPECTROPHOTOMETRIC QUANTITATION

When sufficient quantity of DNA was available, it's concentration was determined by measuring the optical density of the sample at 260nm and 280nm relative to an appropriate buffer blank. These measurements were performed using matched quartz cuvettes and a Pye Unicam SP6 spectrophotometer. The $OD_{260}$ reading enabled the calculation of the concentration of DNA in the sample. An $OD_{260}$ of 1.0 was taken to represent an approximate concentration of 50 $\mu g\ ml^{-1}$ for double stranded DNA, 40 $\mu g\ ml^{-1}$ for single stranded DNA, and 20 $\mu g\ ml^{-1}$ for oligonucleotides (Maniatis et al., 1982). The ratio obtained between the readings at 260nm and 280nm ($OD_{260}/OD_{280}$) provided an estimate of purity of the DNA sample. Pure DNA preparations have a $OD_{260}/OD_{280}$ ratio of 1.80. Values significantly less than 1.8 were an indication of contamination with either protein or phenol in the sample. In such cases accurate quantitation of the amount of DNA in the sample was not possible.

### 2.2.14.2. ETHIDIUM BROMIDE FLUORESCENT QUANTITATION OF DNA

When accurate spectrophotometric quantitation of DNA in a sample was not possible (i.e., when either quantity or purity were insufficient) the agarose spot plate method was used (Maniatis et al., 1982).

Equal volumes (usually 1 $\mu l$) of the DNA sample under investigation and a series of DNA concentration standards were spotted onto a 1% agarose slab containing 0.5 $\mu g\ ml^{-1}$ ethidium bromide which had been cast in a 9 cm Petri dish. After standing for 2 hour, by which time any UV absorbing contaminants would have diffused away, the agarose slab was removed from the petri dish and placed on a short wave UV transilluminator. The spots were photographed and the concentration of the DNA sample estimated by comparing the intensity of the fluorescence in the sample with that of the standard solutions. The standard solutions were pre-

pared from Lambda DNA which had been digested jointly with HindIII and EcoRI, purified by isopropranol precipitation in the presence of 2 M ammonium acetate, resuspended in 1 X TE buffer and their concentration and purity determined spectrophotometrically.

Alternatively, equal volumes (typically 2 to 5 $\mu$l) of the DNA sample and the series of DNA concentration standards were mixed with a one-fifth volume of tracking dye and electrophoresed through a 0.8%, 1 X TBE agarose mini-gel containing 0.5 $\mu$g $ml^{-1}$ ethidium bromide. Electrophoresis was carried out at 60 mA until the bromophenol blue dye front had migrated 2 cm. The gel was then destained by immersing it in 1 X TBE buffer containing 0.01 M $MgCl_2$ for 10 min. The gel was then placed on a short wave UV transilluminator and photographed. The intensity of fluorescence of the DNA sample under investigation was compared with that of the DNA concentration standards to give an estimate of the quantity of DNA in the sample. In this method the DNA concentration standards were derived from a single DNA fragment which had been isolated, purified, and its concentration and purity determined spectrophotometrically.

### 2.2.15. PRECIPITATION OF DNA

#### 2.2.15.1. ETHANOL PRECIPITATION

DNA was routinely precipitated from solution by the addition of a one-tenth volume of 3 M sodium acetate (pH 5.5) and two volumes of chilled absolute ethanol (-20°C), followed by incubation at -20°C for at least 1 hour. Alternatively, incubation was at -70°C for 15 min. After precipitation, the DNA was recovered by centrifugation (Sorval RC5B, SHTM rotor, 12,000 r.p.m., 4°C, 10 min). The supernatant was carefully decanted and discarded. The DNA pellet was either vacuum dried or washed in an appropriate volume of 70% (v/v) ethanol. If washed, the DNA was recovered by a further centrifugation step prior to vacuum drying. The dessicated DNA pellet was resuspended in an appropriate volume of 1 X TE buffer.

2.2.15.2. ISOPROPRANOL PRECIPITATION

If it was desired to rapidly, and selectively, precipitate DNA from a mixed solution of DNA and other materials the sample was precipitated with a 0.4 volume of 5 M ammonium acetate and two volumes of isopropranol. After a 10 min incubation at room temperature the DNA was recovered by centrifugation (Sorval RC5B, SHTM rotor, 12,000 r.p.m., 4°C, 10 min). The DNA was vacuum dried after a brief wash in 70% ethanol and subsequent centrifugation step. The DNA was resuspended in an appropriate volume of 1 X TE buffer.

**2.2.16. PHENOL EXTRACTION OF DNA**

Protein was removed from DNA samples by extraction with phenol, pre-equilibrated with TE buffer. An equal volume of this TE-saturated phenol was mixed with the aqueous DNA sample and vortex mixed for 30 sec. After standing for 5 min the emulsion was separated by centrifugation (MSE Microcentaur, 600 r.p.m., 2 min). The aqueous phase was removed (usually the top phase) and transferred to a fresh tube. Traces of phenol were removed by extracting this phase 3 times with a double volume of water-saturated diethyl ether. Residual ether was removed by incubation of the sample at 55°C for 10 min. The purified DNA sample was then ethanol precipitated.

**2.2.17. DNA SEQUENCING**

The method employed was the "chain-terminating" procedure first described by Sanger et al.. (1977). DNA fragments to be sequenced were first cloned into the the replicative form (RF) of the M13 cloning vectors mp8, mp18, mp9, and mp19. These vectors contain multiple cloning sites in the alpha-peptide of the beta-galactosidase gene. Insertion of foreign DNA into this polylinker sequence destroys active alpha-peptide synthesis such that recombinants are easily identified on media containing the chromogenic substrate, X-gal (BCIG). Single-stranded template DNA is prepared from the mature phage particles of the recombinants and is used

directly in the sequencing reaction. An oligonucleotide primer ("sequence universal primer"), which is complementary in sequence to the viral DNA, is hybridised 3' to the insert DNA sequence. Thus the sequence complementary to this insert DNA is determined by sequential extension with Klenow polymerase in the presence of both deoxyribonucleotides and the chain-terminating dideoxyribonucleotide analogues. Subsequent size fractionation of the reaction products by denaturing polyacrylamide gel electrophoresis enables the DNA sequence to be read.

2.2.17.1. SUB-CLONING OF DNA INTO M13 VECTORS

Routinely, 50 ng of either singly or doubly digested M13 vector was ligated with a twofold molar excess of target DNA fragments. The ligation products were used to transfect competent E. coli JM101. After transformation, the cells were added to a 5 ml test tube, in a 47°C heating block, containing 3 ml molten H-top soft agar , 30 $\mu$l of IPTG (20 mg $ml^{-1}$ in distilled water), 50 $\mu$l of X-gal (40 mg $ml^{-1}$ in DMF), and 200 $\mu$l of exponentially growing E. coli JM101 cells. The mixture was gently mixed and poured on a 2 X YT agar plate and allowed to set before the plate was inverted and incubated at 37°C overnight. Recombinants were identified as colourless plaques owing to the insertional inactivation of the functional beta-galactosidase alpha-peptide encoding sequence.

2.2.17.2. SONICATION OF DNA

In certain cases it was desirable to use the "shot-gun" approach to sequence large DNA molecules. This procedure involves shearing the DNA randomly into much smaller fragments and cloning these fragments into M13 vectors (Messing et al., 1977; Messing & Vieira, 1982). These are then sequenced at random to ultimately reveal the complete sequence. The sonication method followed was that of Deininger (1983), using an MSE soniprep 150 sonicator.

Before use, the probe was boiled in 0.5% (w/v) SDS for 10

min and then rinsed well in sterile distilled water. 25 $\mu$g of the gel purified DNA fragment was placed in a 1.5 ml Eppendorf tube and the volume made up to 100 $\mu$l with 1 X TE buffer. This was kept on ice throughout the procedure. The probe was inserted into the Eppendorf tube as far as possible without touching the tube and allowed to cool for 10 min by contact with the ice. With the sonicator switched to maximum output, and an amplitude setting of 3, two to four 8 second bursts, with 1 min cooling periods between each, were then applied to the DNA solution. The solution was recovered and adjusted to a volume of 150 $\mu$l with 15 $\mu$l 10 X T4 DNA polymerase buffer (2.1.4.), 15 $\mu$l dNTP mix (dATP, dTTP, dCTP, and dGTP; 0.25 mM each), 5 $\mu$l T4 DNA polymerase, 2 $\mu$l Klenow polymerase, and 13 $\mu$l distilled water. The reaction mixture was then incubated at 15$^{o}$C for 3 to 4 hour. The sample was then electrophoresed on a 1.5% agarose mini gel, alongside kilobase ladder DNA size markers, until the bromophenol blue dye front had migrated 2 to 3 cm from the well. DNA fragments between 300 and 900 bp in length were excised from the gel and the DNA extracted by electroelution, and ethanol precipitated.

2.2.17.3. PREPARATION OF TEMPLATE DNA

Recombinant plaques (colourless), between 12 and 96 at a time, were picked into 2 ml aliquots of 2 X YT broth, containing 100 $\mu$l per 100 ml of an overnight culture of E. coli JM101, and incubated at 37$^{o}$C for 6 hour with vigorous aeration. The cultures were then transferred to Eppendorf tubes and centrifuged (Sorval RC5B, SHTM rotor, 12,000 r.p.m., 4$^{o}$C, 10 min). The supernatants (c. 1.25 ml vol) were decanted into fresh Eppendorf tubes containing 250 $\mu$l 20% PEG (6,000); 2.5 M NaCl, and the contents mixed well by inversion. After being left to stand for a minimum of 20 min, the tubes were again centrifuged (Sorval RC5B, SHTM rotor, 12,000 r.p.m., 4$^{o}$C, 10 min). The supernatants were discarded by aspiration leaving the phage pellets in the tubes which were again briefly centrifuged for 30 sec as above to collect all remaining traces of supernatant in the bottom of the tubes. These traces were discarded by careful

aspiration. With the aid of an Eppendorf shaker, the phage pellets were resuspended in 100 $\mu$l volumes of 1 X TE buffer and extracted once with 50 $\mu$l volumes of TE saturated phenol. Following separation of the emulsions by centrifugation (MSE Microcentaur, 6000 r.p.m., 2 min, room temp), the aqueous phases (top phase) were transferred to fresh Eppendorf tubes and traces of phenol removed by extracting twice with 0.5 ml water-saturated diethyl ether. The samples were then ethanol precipitated, washed with 70% ethanol, dried under vacuum, and resuspended in 30 $\mu$l 1 X TE buffer.

2.2.17.4. ANNEALING OF TEMPLATE AND PRIMER

Templates to be sequenced first required annealing with the appropriate oligonucleotide primer. This was carried out in capped Eppendorf tubes. 5 $\mu$l of template DNA was mixed with 1 $\mu$l TM buffer, 3.5 $\mu$l distilled water, and 0.5 $\mu$l primer (2 ng $\mu l^{-1}$ in oligonucleotide storage buffer). The contents were annealed by incubation of the tube in a 80$^{o}$C water bath for 3 min, after which the water bath (and tube) was allowed to slowly cool to room temperature. The tube was then briefly centrifuged (IEC Centra 4X with centra rack rotor) for 5 sec to ensure recovery of the contents in the bottom of the tube.

2.2.17.5. SEQUENCING REACTIONS (POLYMERISATION)

These reactions were performed in capless 1.5 ml Eppendorf tubes which had been autoclaved; the tubes being positioned in ordered centra racks.

For each template, 2 $\mu$l of the annealed mix was transferred to the inside wall of 4 individual tubes; these being for the four independent sequencing reactions: with the "T", "C", "G", and "A" deoxy/dideoxynucleotide mixes which also contained [alpha-$^{32}$P] dATP and Klenow polymerase. 2 $\mu$l aliquots of these nucleotide/polymerase mixes were transferred, also onto the inside wall but not mixing with the aliquoted template, to the appropriate reaction tubes. The

reactions were started by brief centrifugation (IEC Centra 4X with centra rack rotor) and the reaction allowed to proceed for 15 min at room temperature. 2 $\mu$l aliquots of sequence "chase" (0.25 mM dNTPs) was then added to the inside wall of the reaction tubes and again mixed by centrifugation. Following a further 15 min incubation at room temperature, 4 $\mu$l of formamide dye was added to each tube and mixed by centrifugation.

2.2.17.6. LOADING AND RUNNING OF SEQUENCE GELS

Denaturing polyacrylamide gradient gels (6 % w/v) were used routinely. The order of loading was T,C,G,A, for the reaction mixtures for each template sequenced. Immediately before loading the samples were heated in a boiling water bath for 3 min and the gel wells flushed out with 1 X TBE running buffer. Approximately 1 to 2 $\mu$l of sample was loaded in each well using a drawn-out capillary and a constant current of 25 mA was applied. The gels were routinely run for 3.5 hour, by which time the bromophenol blue tracking dye had just run off the bottom of the gel. After electrophoresis the gel, adhered to the back plate, was fixed by immersion in 10% glacial acetic acid, 10% methanol, for 15 min. After being left to stand in a fume cupboard for 1 hour the gels were dried by incubation at 80$^{o}$C for 2 hour, and finally autoradiographed.

2.2.17.7: COMPILATION AND ANALYSIS OF SEQUENCE DATA

All computer analyses performed were with the aid of the programmes supplied by DNASTAR Inc.

## 2.2.18. SYNTHETIC OLIGONUCLEOTIDES

### 2.2.18.1. SYNTHESIS OF OLIGONUCLEOTIDES

Oligodeoxyribonucleotides (oligonucleotides) were synthesised by "solid phase oligonucleotide synthesis" using an Applied Biosystems Model 380A DNA synthesiser. The coupling chemistry employed was the Phosphoramadite method which has inherently high coupling yields, typically 95-100%. The principles of this method of synthesis have been fully described elsewhere (Gait et al.. 1982). Individual syntheses were carried out on a one micro-Mol scale using a controlled-pore glass column. The amount of trityl cation released in each cycle (a direct indication of the stepwise yield) was measured spectrophotometrically at 498 nm (User bulletin, Applied Biosystems, 1984). At the completion of synthesis the oligonucleotides were automatically cleaved from the column and released into a glass collection vial in 1.5 ml of 27% ammonium hydroxide.

### 2.2.18.2. DEPROTECTION OF OLIGONUCLEOTIDES

The sample (in 1.5 ml 27% v/v ammonium hydroxide) was removed from the synthesiser and a further 1 ml of ammonium hydroxide (35% v/v) added and the glass vial sealed. The vial was then incubated in an aluminium heat block at 55°C for a minimum period of 4 hour (maximum of 12 hour). The sample was allowed to cool to ambient temperature prior to opening the vial. The sample was desalted by passing the entire sample (2.5 ml volume) through a 10 ml prepacked Sephadex G-25 (Pharmacia, PD-10) column which had been pre-equilibrated with oligonucleotide storage buffer. The oligonucleotide was recovered by gravitational elution with 3.5 ml oligonucleotide storage buffer. The concentration was measured spectrophotometrically. An $OD_{260}$ reading of 1.0 was taken to represent a concentration of 20 $\mu g\ ml^{-1}$ (Maniatis et al.. 1982). The concentration was typically adjusted to 100 $\mu g\ ml^{-1}$ in oligonucleotide storage buffer and stored frozen at -20°C.

### 2.2.18.3. GEL ANALYSIS OF OLIGONUCLEOTIDES

The custom synthesised oligonucleotides were analysed for size and purity by $^{32}$P-dATP end-labelling of the crude product (after deprotection) followed by resolution through denaturing, 20% polyacrylamide gels. Subsequent autoradiography revealed the relative purity of the full-length oligonucleotides. Where premature termination of synthesis was evident the oligonucleotide was resynthesised. The correct exposure times used were, therefore, very critical as overexposure of certain "chain-length" bands would be under represented on a subsequent densitometer scan of the autoradiogram.

### 2.2.18.4. PURIFICATION OF OLIGONUCLEOTIDES

The following method was used for the purification of sequence universal primer.

500 $\mu$g (25 $OD_{260}$ units) of oligonucleotide was freeze-dried and resuspended in 50 $\mu$l 98% formamide (i.e., 2 $\mu$l per $OD_{260}$ unit) without tracking dye. The sample was heated at 90$^{o}$C for 3 min immediately prior to loading onto a preparative 20% acrylamide/50% urea gel. The gel was prepared as described in section 2.2.10.2 and prerun at 600 V for 30 min. The oligonucleotide solution was loaded at a concentration of 2.5 $OD_{260}$ units per cm of gel as recommended by the manufacturer (User bulletin, Applied Biosystems DNA synthesiser). The loading of the sample was crucial as overloading would cause difficulty in isolating the desired (n-0) band and underloading would hamper visualisation of the product by "UV shadowing". Formamide sequence dye was loaded in two outer wells of the gel and electrophoresis carried out at 400 V for 16 hour, by which time the bromophenol blue dye front had migrated 25 cm from the wells. One of the gel plates was then carefully removed and saran wrap placed on top of the gel. This film was then carefully removed with the gel attached to it and laid to rest on a fluorescent thin layer chromatography plate (TLC). Illumination from above with a medium range UV light transillumi-

nator ("UV shadowing") enabled visualisation of the slowest migrating band, the desired full-length product band. This dark UV absorbing band was excised with a scalpel, taking care to avoid any possible overlapping n-1 band.

The oligonucleotide DNA was extracted from the acrylamide by a modification of the "crush-soak" method first described by Smith (1980). The excised acrylamide was placed in a 0.5 ml Eppendorf tube which had been pierced several times at the tip with an 18-guage hypodermic needle. This tube was placed within a capless 1.5 ml Eppendorf tube such that it was supported. Centrifugation (MSE Microcentaur, 13,000 r.p.m., 15 min) of this arrangement successfully extruded the acrylamide, in a fine particulate form, into the larger tube. 1 ml of oligonucleotide elution buffer was added to the tube and the contents incubated at 37$^{o}$C with vigorous shaking for 3 hour. The acrylamide "slurry" was then centrifuged (MSE Microcentaur, 6,000 r.p.m., 5 min) and the supernatant decanted into a fresh tube. A further 2 extractions were similarly carried out. The supernatants collected were pooled and passed through a 10 ml prepacked Sephadex G-25 column (Pharmacia PD-10) which had been pre-equilibrated with distilled water. The eluted oligonucleotide was freeze dried and resuspended in a one-tenth volume of oligonucleotide storage buffer. The concentration of the sample was determined, the sample diluted to a concentration of 2 ng $\mu l^{-1}$, and stored in 100 $\mu$l aliquots at -20$^{o}$C.

### 2.2.19. SITE-DIRECTED MUTAGENESIS

The method employed was a derivation of that described by Carter et al.. (1985) (D.A. Barstow, personal communication).

The target DNA was first cloned into a suitable M13 mp vector. The recombinant M13 derived template DNA was then annealed with an appropriately phosphorylated 21 base mutagenic oligonucleotide which contained a single, centralised, base mismatch. The mismatch was the complement of

the resultant base change to be created in the target sequence. A subsequent Klenow polymerase-directed "extension reaction", primed by the phosphorylated mutagenic oligonucleotide, in a repair-deficient E. coli host, resulted in the generation of a mutant (-) M13 strand. This (-) strand was then faithfully copied by the compromised host to give the mutant (+) M13 strand. Resultant M13 clones were screened by hybridisation experiments with the mutagenic oligonucleotide, and those mutants identified subjected to DNA sequence analysis.

### 2.2.19.1. PHOSPHORYLATION OF MUTANT OLIGONUCLEOTIDES

The freshly synthesised mutagenic oligonucleotide, which had been deprotected, was adjusted to a concentration of 100 ng $\mu l^{-1}$. A 10 $\mu l$ aliquot of the oligonucleotide was mixed with 2 $\mu l$ 10 X kinase buffer, 1 $\mu l$ 100 mM DTT, 2 $\mu l$ 10 mM rATP, 1 $\mu l$ T4 polynucleotide kinase, and 4 $\mu l$ distilled water. The reaction mixture was incubated at 37$^{o}$C for 30 min and then the remaining enzyme inactivated by heat treatment (70$^{o}$C/10min).

### 2.2.19.2. EXTENSION REACTIONS AND TRANSFORMATION OF E. coli BMH71-18

The M13 template DNA (target DNA) was adjusted to an approximate concentration of 2 $\mu g$ per 15 $\mu l$ in 1 X TE buffer. 15 $\mu l$ of this DNA was added to a tube containing 4 $\mu l$ of the phosphorylated oligonucleotide and 2 $\mu l$ of TM buffer and annealed as described in 2.2.17.4. Once the annealed template had returned to room temperature the following were added: a further 5 $\mu l$ TM buffer, 4 $\mu l$ of a dNTP solution (5 mM dATP, dTTP, dCTP, and dGTP), 1 $\mu l$ rATP, 2 $\mu l$ 100 mM DTT and 10 $\mu l$ distilled water. The reaction mixture was placed on ice and then 2 $\mu l$ T4 DNA ligase (BRL, 2.5 u $\mu l^{-1}$) and 0.8 $\mu l$ Klenow polymerase (BCL, 5 u $\mu l^{-1}$) was added and the mixture incubated at 15$^{o}$C for 16 hour. The reaction was stopped with the addition of 200 $\mu l$ 10 mM $Na_2EDTA$, 100 mM Tris-HCl (pH 8.1).

A range of aliquots (1, 5, and 10 $\mu$l) of the arrested extension reaction mixture was used to transform 200 $\mu$l of amounts of competent E. coli BMH 71-18 exactly as previously described . The transformed cells were split in half and added to molten (45$^{o}$C) 3 ml soft agar overlays, which were supplemented with 200 $\mu$l of exponentially growing E. coli JM101, poured over 2 X YT agar plates, and allowed to set. The plates were incubated overnight at 37$^{o}$C.

2.2.19.3. COLONY SCREENING OF MUTANTS

Typically 100 M13 plaques were picked from each mutagenesis experiment and streaked onto 2 X YT agar plates in an asymmetric pattern, thereby facilitating the later correct identification of the mutants. The plates were incubated overnight at 37$^{o}$C and then "plaque lifted". A plain 9 cm nitrocellulose disc was laid to rest on the surface of the agar plate, on which the phage-infected bacterial colonies had grown, and left to stand at room temperature for 1 min. The filter was carefully lifted away, with the bacterial colonies adhering to it, and processed as described elsewhere for in situ colony hybridisation with the omission of the 2 X SSC rinse and proteinase K treatment. The filter was placed on a sheet of Whatman 3MM filter paper, colony side up, and allowed to air dry at room temperature. The filter was then baked at 80$^{o}$C in a vacuum for 30 min.

The filter was then wet in 6 X SSC for 5 min and hybridised with the mutagenic oligonucleotide which had been end-labelled with [gamma$^{32}$P] dATP. This was performed exactly as described elsewhere (2.2.13.2.) with the exceptions that sheared heat denatured salmon sperm DNA was omitted from the prehybridisation solution and that the the hybridisation was carried out at a $T_d$-15$^{o}$C for the oligonucleotide. The filter was then washed with 3 changes of 100 ml 6 X SSC at room temperature, 1 min each, and wrapped in Saran wrap and autoradiographed with preflashed film at -70$^{o}$C for 1 hour. This revealed the asymmetric pattern of the bacterial colonies. The filter was then progressively washed (just 1 X 100 ml 6 X SSC) at increasing tempera-

tures, with subsequent autoradiography of the filter inbetween each wash, until the $T_d$ ($^oC$) of the mutagenic oligonucleotide had been reached. As this $T_d$ temperature was approached the non-mutant colonies began to lose their hybridisation signal relative to the positive mutant colonies (as a consequence of the lower $T_d$ due to the mismatch). At the $T_d$ of the mutagenic oligonucleotide (or in some cases a few degrees higher) the only hybridisation signals obtained were from the positive colonies.

2.2.19.4: PLAQUE PURIFICATION AND DNA SEQUENCING

For each mutagenesis experiment, 4 hybridisation positive colonies were chosen. These were individually transferred to 20 ml 2 X YT broth and vortex mixed. Typically, a 10 $\mu$l volume of these diluents was streaked over the dried surface of a 2 X YT agar plate. The plates were then overlaid with 3 ml molten (47$^oC$) H-top agar containing 200 $\mu$l exponentially growing E. coli JM101, left to set, and incubated overnight at 37$^oC$. Isolated plaques were chosen (usually 3 for each mutant) and template DNA prepared from them. The entire cloned insert of each template was subjected to DNA sequence analysis as described in 2.2.17.5. This was generally accomplished by using a number of custom synthesised oligonucleotide primers. DNA sequence analysis not only confirmed the mutated sequence but also checked that no other sequence alterations had occurred. Template DNA (1 $\mu$l) of the chosen mutant(s) was used to transfect competent E. coli JM101 cells and the transfected cells used to inoculate 10 ml 2 X YT broth for overnight incubation to provide the "phage inoculum" for the isolation of double-stranded RF DNA.

## 2.2.20. DETECTION OF SINGLE-STRANDED PLASMID DNA IN B. SUBTILIS

The method used was that described by S.D. Ehrlich (personal communication) and was used for the detection of single-stranded plasmid replication intermediates.

An overnight L-broth culture of B. subtilis harbouring the plasmid under investigation was grown with vigorous aeration at 37°C in the presence of appropriate antibiotic. The culture was diluted 100-fold in 10 ml of fresh broth and again incubated at 37°C with vigorous aeration. The $OD_{650}$ was monitored and the cells were placed on ice when a an $OD_{650}$ of 0.6 was reached. 5 ml of culture was then centrifuged (MSE Microcentaur, 13,000 r.p.m., 2 min), the cells washed in 1 ml cold (4°C) B. subtilis washing buffer, recentrifuged, and resuspended in 193 ul B. subtilis lysis buffer. The sample was then incubated for 10 min at 0°C followed by a further 10 min at 37°C. Following the addition of 7 $\mu$l 35% (v/v) sarcosyl, the sample was gently vortex mixed and incubated for 20 min at 70°C. The lysate was allowed to cool on ice for 2 min prior to vortex mixing for 1 min at maximum speed. The lysate was then extracted twice with an equal volume of TE saturated phenol and finally once with an equal volume of chloroform:isoamylalcohol (24:1 v/v). DNAase-free RNAase (2 $\mu$l of 20mg $ml^{-1}$ stock) was then added and the sample incubated at 37°C for 10 min. To 60 $\mu$l of the lysate, on ice, 7 $\mu$l 10 X S1 nuclease buffer and 2 $\mu$l S1 nuclease (20,000 u $ml^{-1}$) was added. The sample was then incubated at 37°C for 10 min and then returned to the ice. A 0.8% (1% w/v) agarose gel was then loaded on the left and on the right with 34 $\mu$l of nuclease S1 digested lysate and 30 $\mu$l of the non-digested remaining lysate. Following overnight electrophoresis, the gel was photographed alongside a ruler and cut in two halves. Half of the gel was stored at 4°C while the other half was denatured. The denatured and non-denatured gel halves were then blotted to nitrocellulose membrane in 20 X SSC transfer solution. The membrane was then hybridised with a nick-translated plasmid probe. Following autoradiography, single-stranded plasmid DNA was visualised as a fast migrating DNA species present in only the undigested (S1 nuclease) sample and the sample transferred from the undenatured gel.

### 2.2.21. FILTER MATING PROCEDURE

Filter matings were performed to facilitate the conjugal transfer of pAMβ1 cointegrates from a B. subtilis donor to C. acetobutylicum recipient (Oultram & Young, 1985).

Mid-exponential phase aerobic cultures of B. subtilis (3 ml) in L-broth, and mid-exponential phase anaerobic cultures of C. acetobutylicum (8 ml) in TYG broth were centrifuged (Sorval RC5B, SS24 rotor, 5,000 r.p.m., 10 min) and the bacterial pellets resuspended in 2.5 ml holding buffer (1 mM $MgSO_4$ in 25 mM potassium phosphate, pH 7.0). The mating mixture was prepared by combining the suspensions of bacteria in holding buffer, harvesting the organisms onto a nitrocellulose filter (pore size, 0.45 μm; Millipore), and placing the filter, bacteria lowermost, on RCM agar. It was not necessary to employ anaerobic conditions when carrying out these manipulations. After incubation of the mating mixture for 24 hour in an anaerobic environment, the bacteria were recovered, and resuspended in 1 ml holding buffer. The numbers of donor and recipient bacteria were assessed by serial dilution in holding buffer and then plating samples on appropriate selective media. Transcipients carrying the pAMβ1 cointegrates were selected by plating on the selective medium appropriate.

### 2.2.22. IN VITRO TRANSCRIPTION/TRANSLATION ASSAYS

These were performed with a transcription/translation kit obtained from Amersham International. Reactions were carried out as specified by the supplier. Typically 5-10 μl aliquots of the reaction products were subjected to SDS PAGE electrophoresis (10-30% gradient gel) alongside mid-range protein size markers which were labelled with $^{14}C$ (supplied by Amersham International).

### 2.2.23. CATECHOL 2,3-DIOXYGENASE (C23O) ACTIVITY ASSAY

#### 2.2.23.1. AGAR PLATE DETECTION METHOD

Functional expression of the xylE gene was detected by spraying antibiotic selection plates with an aqueous solution of 0.5 M catechol. Colonies that express the gene become yellow due to the conversion by catechol 2,3-dioxygenase to 2-hydroxymucomic semialdehyde (Zukowski et al., 1983).

#### 2.2.23.2. SPECTROPHOTOMETRIC ASSAY

For enzymatic assays, the levels of C23O produced were measured spectrophotometrically (Sala-Trepat & Evans, 1971) from sonic extracts prepared essentially as described by Zukowski et al., 1983; Zukowski & Miller, 1986.

Cell cultures were grown overnight in 10 ml of appropriate broth and harvested by centrifugation (Sorval RC5B, SS24 rotor, 6,000 r.p.m., 10 min). The cells were washed in 5 ml 20 mM phosphate buffer, pH 7.2, buffer, recentrifuged as above, and resuspended in 1 ml 100 mM phosphate buffer, 10% acetone (v/v), pH 7.5. The cell suspensions were cooled on ice and then disrupted by three thirty sec sonications (MSE Soniprep 150) with 30 sec cooling intervals between bursts. The extracts were then centrifuged (Sorval RC5B, SHTM rotor, 12,000 r.p.m., 4°C, 15 min) to remove cellular debris. The supernatant ("extract") was decanted to a fresh tube and kept on ice. C23O activity was determined by following the increase in absorbance at 375 nm due to the accumulation of 2-hydroxymuconic semialdehyde; cuvettes contained in a total volume of 3 ml: 250 µM of phosphate buffer, pH 8.0, 0.2 µM of catechol and extract (between 1 µl and 100 µl depending on the amount of activity present. All measurements were carried out at 30°C. One unit of enzyme was defined as that amount which oxidised one micromole of catechol per min at 30°C. The molar extinction coefficient of the product used was that of Nozaki (1970), being 4.4 X $10^4$; one unit of enzyme activity corresponded

to an increase in absorbance of 14.8 $min^{-1}$. Specific activity was expressed as units per milligram of protein.

### 2.2.24. CARBOXYPEPTIDASE $G_2$ PLATE ASSAY

E. coli cells harbouring a functional carboxypeptidase $G_2$ gene were detected by the Minimal media/folate agar plate method described by Minton & Clarke (1985). This method is reliant on the observation that a functional crboxypeptidase $G_2$ gene enables E. coli strains to utilise folate as a sole carbon source with the production of a bright yellow product. Cells were simply plated onto minimal medium/folate agar plates and incubated for 48 hour at 37°C, by which time those carboxypeptidase $G_2$ positive colonies present were easily detected by their bright yellow colour.

### 2.2.25. PROTEIN DETERMINATION

The Coomassie Brilliant blue G250 binding assay was used to estimate protein levels in sonic extracts. 0.2 ml of dye reagent concentrate (Biorad Protein Assay Bulletin 1069) was added to 0.8 ml of sample, containing 2 to 25 $\mu g$ of protein. After gentle mixing, and a period of 5 to 60 min at room temperature, the absorbance at 595 nm was read against a reagent blank. The protein concentration was determined from a calibration curve using BSA as standard.

# CHAPTER THREE

# CHARACTERISATION OF A CLOSTRIDIAL REPLICON

## 3.1. INTRODUCTION

As no suitably marked plasmids had been identified from any saccharolytic clostridia, it was decided to undertake the *in vitro* construction of a plasmid vector by combining the replicative machinery of a cryptic plasmid with a suitable genetic marker known to be phenotypically expressed in the intended clostridial host (e.g., an antibiotic resistance gene). The availability of such a vector should then facilitate the development of a transformation procedure.

At this time the most extensively characterised small plasmids, derived from saccharolytic clostridia, were the cryptic plasmids pCB101 (5.6 kb), pCB102 (7.4 kb), and pCB103 (6.2 kb) isolated from *C. butyricum* (strains NCIB 7423 and NCTC 6084). First described by Minton & Morris (1981), these same plasmids were also isolated and partially characterised by Luczak *et al.* (1985). Both groups of researchers simultaneously published restriction endonuclease maps for all three plasmids (Collins *et al*.. 1985; Luczak *et al*., 1985).

Additional studies were directed towards the identification of the minimal replicon of these plasmids. In the absence of a reliable transformation procedure for saccharolytic clostridia, studies were performed in *B. subtilis*. Various restriction endonuclease fragments derived from pCB101 and pCB102 (Collins *et al*.. 1985) and pCB103 (Minton & Clarke, unpublished data) were cloned into Gram-positive replication-deficient plasmid vectors (i.e., "replicon cloning vectors") and the ability of the resultant chimaeras to replicate in *B. subtilis* tested. Collins *et al*. (1985) identified a 3.3 kb *Sau*3A fragment of pCB101 which conferred on the replicon cloning vector used (pJAB1) the ability to autonomously replicate in *B.subtilis*. They also

identified a 2.0 kb fragment of pCB102 which, upon appropriate selection, caused the chimaeric construction with pJAB1 to integrate into the B. subtilis chromosome. Although the pCB101 chimaera (pRB1) appeared to replicate autonomously in both $rec^+$ and $rec^-$ strains of B. subtilis. it appeared to be present in low copy number. Furthermore, pRB1, appeared to be unstable as it was rapidly lost from the population when selection was removed (Collins et al., 1985). Even so, at the time of this study pCB101 still represented the best candidate for use in the construction of a clostridial cloning vector.

An additional attraction to employing pCB101 in the intended vector construction studies was the earlier report that this plasmid might encode a bacteriocin (butyricin); if so determined, it would represent the first identified "marked" plasmid from a saccharolytic Clostridium. The suggested role of either of the two co-resident plasmids, pCB101 and pCB102, in bacteriocin production (Minton & Morris, 1981) was plausible as these two plasmids were isolated from the bacteriocinogenic strain, C. butyricum NCIB 7423 (Clarke & Morris, 1976), and bacteriocin production has been associated with other clostridial plasmids (Garnier & Cole, 1986; Ionesco et al.. 1974, 1976; Li et al., 1980; Mihlec et al.. 1978).

For this study it was decided to further characterise the plasmid pCB101. By determining the complete nucleotide sequence of pCB101, the objectives were twofold: (1) to study the region of pCB101 that appears to comprise the minimal replicon and hopefully gain some understanding of the mechanisms involved in the replication of this plasmid that might be beneficial to the vector construction program, and (2) to examine the remainder of the plasmid for open reading frames (ORFS), to which functions might be

open reading frames (ORFS), to which functions might be assigned, e.g., bacteriocin production.

## 3.2. RESULTS

### 3.2.1. SUB-CLONING OF THE MINIMAL REPLICON OF pCB101

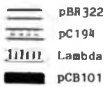
The plasmid isolation techniques described for C. butyricum (Minton & Morris, 1981) proved relatively inefficient, resulting in low yields of pCB101 DNA. The recombinant plasmid pRB1 (Collins et al., 1985) was therefore chosen as a suitable source of the 3.3 kb Sau3A fragment of pCB101; propagation of pRB1 in E. coli would therefore enable the isolation of a sufficient quantity of the 3.3 kb Sau3A fragment for nucleotide sequence determination.

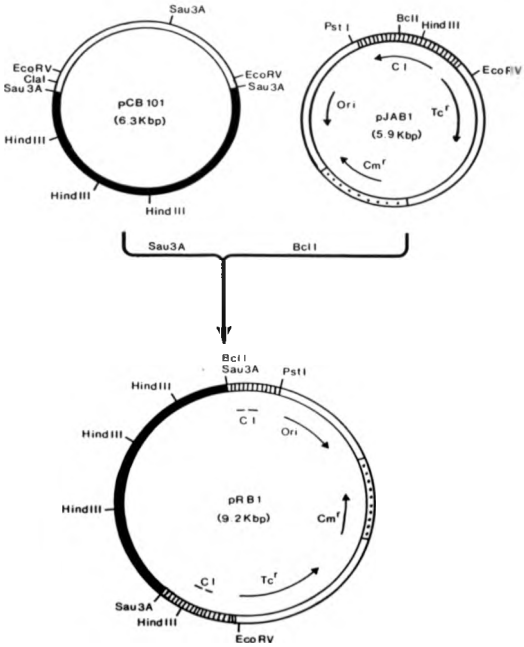
A strain harbouring pRB1 (Fig. 3a) was obtained from M. Young (University of Aberystwyth) and employed for the large scale isolation of plasmid DNA. The plasmid DNA obtained, however, was found to be substantially smaller than expected, with the largest DNA fragment generated by digestion with Sau3A being only 2.4 kb in size. The deleted plasmid was designated pRB1Δ1. A further three independently isolated clones carrying equivalent plasmids to pRB1 were obtained from M.Young and large scale plasmid preparations undertaken. On the basis of the sizes of the fragments released after digestion with Sau3A, all three plasmids were identical to pRB1Δ1. These plasmids were therefore designated pRB1Δ2-4.

To determine the extent of the observed deletions extensive restriction analysis of pRBΔ1-4 was undertaken. This involved digestion of the four plasmid DNA preparations, both singly and in various combinations, with HindIII, TaqI, BglI, AvaII, AccI, PstI, HpaI, NcoI, AatII, and Sau3A. The

FIGURE 3a DERIVATION OF THE CHIMAERIC PLASMID pRB1

The various regions of the plasmids are as follows:



cal and that the deleted DNA was predominantly of a clostridial origin. The deletion event had resulted in the loss of a 2.5 kb region of the pRB1 restriction map (Fig. 3b) encompassing the clostridial DNA insert.

The results obtained indicated that plasmids generated by inserting pCB101 DNA into pJAB1 were structurally unstable in E. coli. This factor had not come to light in the study by Collins et al. (1985), as large scale preparation of pRB1 had not been undertaken (M. Young, personal communication). In the study of Luczak et al. (1985) similar plasmids had been generated using an alternative vector to pJAB1. These resultant plasmids were structurally stable (W.L. Staudenbauer, personal communication). This would suggest that it was the cloning vector employed rather than the pCB101 DNA per se which induced instability on recombinant plasmids. It was therefore elected to sub-clone the 3.3 kb Sau3A fragment of pCB101 into the smaller and better characterised vector, pUC8. Plasmid DNA prepared from C. butyricum NCIB 7423 was digested to completion with Sau3A and the 3.3 kb DNA fragment derived from pCB101 isolated by electroelution. This purified DNA fragment was ligated with BamHI cut, dephosphorylated, pUC8 DNA and the ligation products used to transform competent E. coli JM83. The plasmid DNA of six presumptive recombinant clones was prepared by the small scale isolation technique and analysed by digestion with Sau3A and subsequent agarose gel electrophoresis (1% w/v). In each case the presence of a single 3.3kb DNA insert was demonstrated. One clone was chosen and employed for the large scale isolation of plasmid DNA. Subsequent digestion with Sau3A and agarose gel electrophoresis (1% w/v) of this DNA, designated pCB1, indicated that it was structurally stable. Plasmid pCB1 is illustrated in Fig. 3c.

FIGURE 3b RESTRICTION ENDONUCLEASE MAP OF pRB1Δ1-4

The restriction endonuclease sites are as follows: AccI, A; AatII, Aa; AvaIII, Av; BgII, B; EcoRI, E; HinIII, H; BamHI, Ba; NcoI, N; PstI, P; SalI, S.

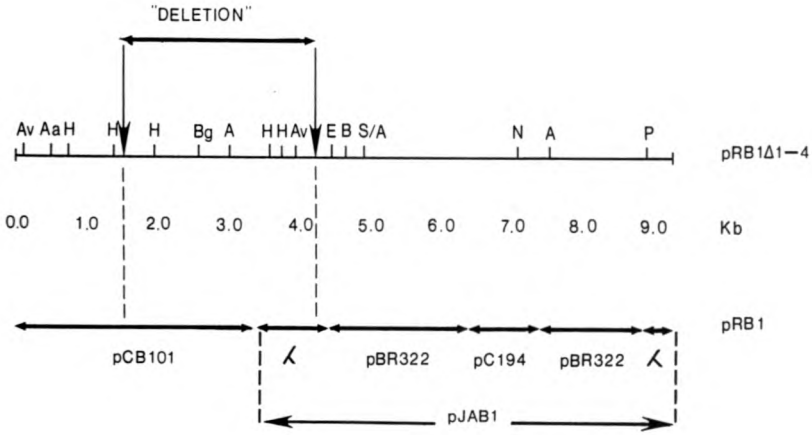
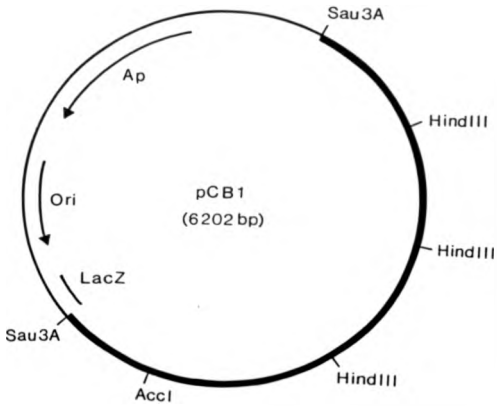


FIGURE 3c RESTRICTION ENDONUCLEASE MAP OF pCB1

The various regions of the plasmids are as follows: ——— pUC8

▬▬▬ pCB101



### 3.2.2. NUCLEOTIDE SEQUENCE DETERMINATION OF THE MINIMAL REPLICON OF pCB101

Approximately 100 μg of pCB1 DNA was digested to completion with Sau3A and the 3.3 kb pCB101 derived DNA fragment isolated by electroelution. A random population of blunt-ended sub-fragments (300-900 bp) were generated by the sonication procedure and inserted into the SmaI site of M13mp8 (2.2.17.). Template DNA (for sequencing reactions) was prepared from 250 recombinants and sequenced using the "chain-termination" method (2.2.17.). The sequence data obtained was compiled into one complete sequence (82 % of which was determined for both strands) using the programs of DNASTAR Inc. The total length of the Sau3A fragment was determined to be 3,480 bp and to contain a previously undesignated HindIII site (Collins et al., 1985; Luczak et al., 1985). Features of the sequence are discussed later (3.2.5.) in conjunction with the sequence of the remainder of pCB101.

### 3.2.3. NUCLEOTIDE SEQUENCE ANALYSIS OF pRB1Δ1-4

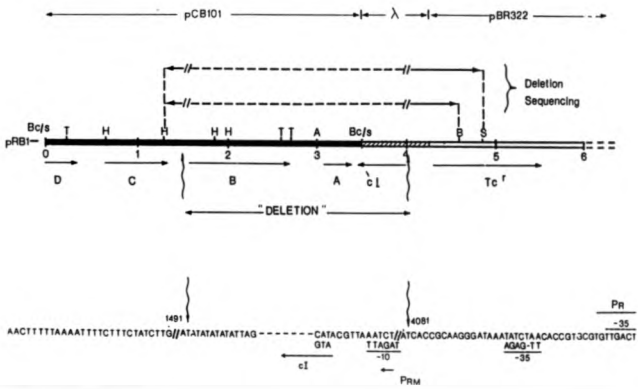
Having determined the complete nucleotide sequence of the 3.48 kb Sau3A fragment the extent of the deleted DNA in pRB1Δ1-4 was determined.

The strategy initially adopted was to clone a 0.73 kb BamHI-HindIII fragment, which spanned the deletion end points of all four plasmids, into M13mp8 and M13mp9 (Fig. 3d). The appropriate fragments (0.73 kb) were isolated from pRBΔ1-4 and ligated with BamHI-HindIII cut M13mp8 and M13mp9. Upon subsequent transformation of competent E. coli JM101, recombinant plaques were only obtained with M13mp8; even after further attempts, M13mp9 recombinants were inexplicably unattainable. In view of this problem it was

FIGURE 3d SCHEMATIC REPRESENTATION OF THE pRB1ω-4 DELETION EVENT

The restriction endonuclease sites are as follows: Acc1, A; BamH1, B; Bcl1, Bc; HindIII, H; Sal1, S; Sau3A, s; Taq1 T.

The nucleotide sequence at the points of the deletion is indicated below the restriction map. DNA subfragments isolated to facilitate nucleotide sequencing across the deletion junction are indicated above the restriction map.



elected to isolate the overlapping 1.0 kb SalI-HindIII fragment for cloning into M13mp9 (Fig. 3d). In contrast to the previous result, this fragment was successfully cloned into M13mp9. Template DNA was prepared from both the M13 mp8 (BamHI-HindIII) and M13 mp9 (SalI-HindIII) recombinants and sequencing was undertaken using universal sequence primer and a custom synthesised 17'mer oligonucleotide primer (complementary to the pRB1 sequence at position 4300; Fig. 3b, 3d).

The nucleotide sequence at the deletion junction of pRB1 1-4 established that they were identical. In each case, a 2.95 kb region of DNA had been deleted from pRB1 extending from position 1491 to 4081 (Fig. 3b, 3d). The reasons for these specific deletions were at the time unclear as there appeared to be no homology between the two sequences at the sites of the deletion event; a feature sometimes encountered with topoisomerase-mediated illegitimate recombination (Drlica, 1984). One obvious feature is the sequence at position 1492 (ATATATATATAT) which is both repetitive and palindromic (Fig. 3d). The fact that the clostridial DNA was unstable when cloned in pJAB1, and not in pUC8, suggested that a sequence within the lambda DNA segment of the vector might be responsible.

### 3.2.4. NUCLEOTIDE SEQUENCE DETERMINATION OF THE REMAINDER OF pCB101

#### 3.2.4.1. CONSTRUCTION OF pCB2

Working from the restriction map of Luczak *et al.* (1985), the large HindIII fragment (c. 4.6 kb) was chosen as the source of the remaining pCB101 DNA (Fig. 3a). Plasmid DNA (i.e., pCB101 and pCB102) was isolated from *C. butyricum* NCIB 7423, cleaved with HindIII and the products resolved

by preparative agarose gel (1 % w/v) electrophoresis. The desired fragment DNA, which co-migrated with a similar sized fragment derived from pCB102 (Luczak et al., 1985), was excised from the gel and the DNA extracted by electroelution. This mixed fragment DNA was ligated with pAT153 DNA which had been cleaved with HindIII and dephosphorylated. Following transformation of E. coli W5445, recombinants were detected on the basis of their $Ap^r$, $Tc^s$ phenotype. Subsequent digestion of small scale plasmid DNA preparations with PstI followed by size-fractionation of the products by agarose gel (1%, w/v) electrophoresis, revealed which recombinants contained the desired 4.6 kb pCB101 HindIII fragment. Two clones were chosen for large scale isolation of plasmid DNA. Restriction analysis with AccI, ClaI, EcoRV, and Sau3A indicated that they were both identical, with the pCB101 HindIII fragment inserted in the same orientation. One of these clones, pCB2 (Fig. 3e), was chosen as substrate for ensuing sub-cloning to facilitate nucleotide sequencing.

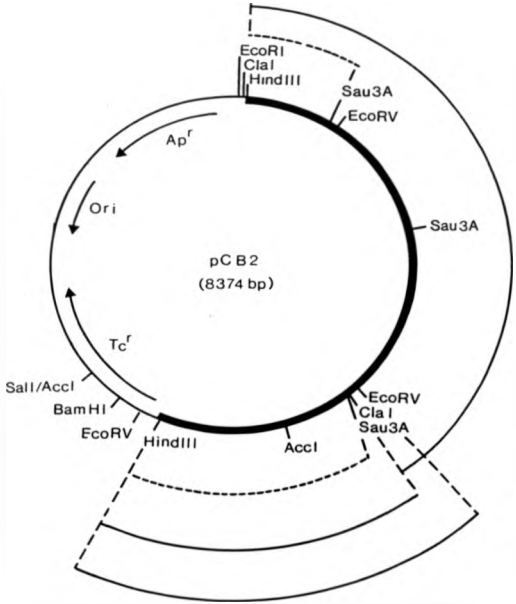
3.2.4.2. SEQUENCING OF THE CLAI FRAGMENT OF pCB2

In order to avoid the excessive duplication of nucleotide sequence determination that would result from the random sequencing of the 4.6 kb HindIII fragment of pCB2 (i.e., some 2.1 kb DNA), it was elected to initially sequence the 3.2 kb ClaI fragment. This fragment, by virtue of the unique ClaI sites in the pCB101 derived HindIII fragment and the pAT153 vector, contained the majority of the remaining pCB101 sequence (Fig. 3e). The small region of remaining sequence (that which spans the ClaI site itself) was to be sequenced separately (3.2.4.3.).

FIGURE 3e RESTRICTION ENDONUCLEASE MAP OF pCB2

The various regions of the plasmids are as follows: ——— pAT153
▬▬▬ pCB101

Regions of pCB101 previously sequenced are indicated by the dashed lines. Regions isolated for further DNA sequence analysis are indicated by the solid line



Approximately 100 μg of pCB2 DNA was digested to completion with ClaI and the desired 3.2 kb ClaI fragment DNA isolated by electroelution. A random population of blunt-ended sub-fragments (300-900 bp) was generated by the sonication procedure and inserted into the SmaI site of M13mp8 (2.2.17.). Following transformation of competent E. coli JM101, template DNA was prepared from 250 recombinants plaques and nucleotide sequencing undertaken. The sequence data obtained (90 % of which was obtained for both strands) was compiled into one complete sequence using the computer programs of DNASTAR Inc. The total length of the ClaI fragment was determined to be 3,208 bp and lacked a previously reported Sau3A site (Collins et al.. 1985), suggesting that it may have either been incorrectly mapped or it's presence was simply proposed to account for 0.3 kb of DNA of the published restriction map (Collins et al., 1985). The features of the sequence are presented later (3.2.5.).

3.2.4.3. SEQUENCING OVER THE CLAI JUNCTION OF pCB2

To complete the sequence of pCB101 it was necessary to sub-clone both the HindIII-ClaI and the HindIII-EcoRV fragments (1.2 kb each) of pCB2 (Fig. 3e). Nucleotide sequence determination in the direction initiating at the ClaI and EcoRV termini of the two fragments would thus (1) complete the sequence of pCB101 by joining the 2 continuous sequences obtained so far, (2) locate the "missing" Sau3A site reported by Collins et al. (1985) if it existed at all, and (3) verify the existence of single sites for ClaI and EcoRV in this region.

Aliquots (10 μg) of pCB2 DNA were digested with combinations of HindIII-ClaI and HindIII-EcoRV and the desired

preparations (HindIII-ClaI and HindIII-EcoRV) were then ligated with M13mp9 DNA preparations which had been cleaved with HindIII-AccI and HindIII-SmaI respectively. Following transformation of competent E. coli JM101, template DNA was prepared from six recombinants from each cloning. Subsequent nucleotide sequencing using universal sequence primer enabled the joining of the two sequences, verified the presence of single sites for ClaI and EcoRV in that region, and established the absence of an additional Sau3A site. The complete nucleotide sequence of pCB101 is given in Fig. 3f.

### 3.2.5. FEATURES OF THE pCB101 SEQUENCE

#### 3.2.5.1. THE pCB101 DNA SEQUENCE

The complete nucleotide sequence was determined to be 6065 bp which agrees well with those values determined experimentally (Collins et al., 1985; Luczak et al. 1985). Differences at the nucleotide sequence level included an additional HindIII site at position 4388 (Fig. 3f) and the absence of the Sau3A site reported by Collins et al. (1985) to exist in the region proximal to the Sau3A site at position 2548 (Fig. 3f). A circular restriction endonuclease map is given in Fig. 3g.

The overall dA + dT content of pCB101 was determined to be 72.8% which is comparable to that of the C. butyricum chromosome, 71-73 % (Gottschalk et al., 1981), and the value reported for the only other clostridial plasmid sequenced to date, pIP404 (75 %) (Garnier & Cole, 1988a).

FIGURE 3f COMPLETE NUCLEOTIDE SEQUENCE OF pCB101

Sequence features are illustrated as follows: Direct repeats, solid arrows; inverted repeats, broken arrows; palindromes, solid lines.
Major restriction endonuclease sites are indicated.

```
                                                                                EcoRV
TTTAGTATTTTTTATGCTTAAATACGTAATGGGAGGATATTTAAAAAAATGATATCTAATGTTGAATTTTTTTCT   75

TATGTTTTTCCTAGACTTGTTCAAAGTGCTTTTGTAATGGGTATTAGTATGTTTTCTTTAGCGTTTATATTCCGT  150

TTTGGAATTGACTTACTTTTATCTGTTTTTAAAATAATAGCAAAATAAAAAATTGGAGGTAATGTAATGACTAAA  225
                                                                   ORF G  m  t  k
TTAAAGAAAATTTTATATTGTGGTCGTACTGGTGTTATTGCAAGTGCATTAACAACTGTTTCAGCTTTTGCAGCA  300
 l  k  k  i  l  y  c  g  g  t  g  v  i  a  s  a  l  t  t  v  a  s  f  a  a
GAGGGAAGTGATAGCTCAACAGTTATAACTGGAATAACTGATGGTTTAACTTCTGGAAAAACTGAATTTGTAACT  375
 e  g  s  d  s  s  t  v  i  t  g  i  t  d  g  l  t  s  g  k  t  s  f  v  t

                                                                        ScaI
GCTCTTGCTGCTATTTCTGGAATTGCTATTGGATTCTTTGTTGTAAAGTTTCTTGTAAAACAAGGTATTAAGTAC  450
 a  l  a  a  i  s  g  i  s  i  g  f  f  v  v  k  f  v  v  k  q  g  i  k  y

                                                                StuI
                                                  ___________________ ___
TTTTCTACTATTGCAAATAAAGCCTAGGATTAATTATAAAAAAAGATTTTCGGCATATTGGCGCAGGCCTTTAAA  525
 f  s  t  i  s  n  k  g  .

 ___________________
TGTGCCGAAAATCTTTTTTTATTTTAAATAGAGGTGTTATATGAAAAAAATTATGAAAAGATTTTTAATTTTATT  600
                                                ORF F  m  k  k  i  m  k  r  f  l  i  l  f
TTTATGTTTTAGTTTATTAAGTGGAATTTATACAAAAAAAGCAAATGCTATAGGCGTAGTTGATGATGCTTTTTA  675
 l  c  f  s  l  l  s  g  i  y  t  k  k  a  n  a  i  a  v  v  d  d  a  f  y
TATAGTTGCCGGTGTTGTAGTTACTGGTGCTACAATTTATGCTTTAGTTGACCATACTCTGTCTATGGGTTCAAG  750
 i  v  a  g  v  v  v  t  g  a  t  i  y  a  l  v  d  h  t  l  s  m  g  s  s
TTCTCATACTGCAAATGATTCTGCAAATATATCTTCTTCTTTATCTAATACACAAAAGTTATCACTTTATTTAAC  825
 s  h  t  a  n  d  s  a  n  i  s  s  s  l  s  n  t  q  k  l  s  l  y  l  t
TGTTAAAACATTAGAAAATGGTTCTAAATTTCTTACTTGGACTAAAAGTGGACTAGATTGGCTTACTAACTTTAT  900
 v  k  t  l  s  n  g  s  k  f  l  t  w  t  k  s  g  l  d  w  l  t  n  f  i
AGATACTACAATAAATACATCTTATGTTAATAATTATGGTTCTATTCAAATTTCACAATCTAATACTGTTACTTC  975
 d  t  t  i  n  t  s  y  v  n  n  y  g  s  i  q  i  s  q  s  n  t  v  t  s
TCCACTTGTGTTAAGTTCTCAATTTACAGACCAAAGTTGTGTATTTAGATATGATATTGGTTATGGTCAGACTTC 1050
 p  l  v  l  s  s  q  f  t  d  q  s  c  v  f  r  y  d  i  g  y  g  q  t  s
CAGAAGTCTTTACTATGGTTCTTTTGCTGGACCTTATTATGTTGGCGCTTTTTTTTATAATTACAATTATTATAT 1125
 r  s  l  y  y  g  s  f  a  g  p  y  y  v  g  s  f  f  y  n  y  n  y  y  i

TGGTGTTTTTAATGCAAATGAAAATGTTTATTCACTTCCTGTTTCAAGCGTTGTTAATTCATTTGAGGTTGAATT 1200
 g  v  f  n  s  n  s  n  v  y  s  l  p  v  s  s  v  v  n  s  f  e  v  e  f
TCCAGAGCGTTCATGTGACGTTAAAGATTATCTCCAGACGTTGTTGATATAATTTTTTTAATCCAGAGCCTGA 1275
 p  e  r  s  c  d  v  k  d  y  s  p  d  v  v  d  y  n  f  f  n  p  s  p  e
ACCAGATGAAGATAATAGTTATAAACCTAAAATCGGTTTTCCATTACGTGAATATGATAAAAATAAAGATGGATA 1350
 p  d  e  d  n  s  y  k  p  k  i  g  f  p  l  r  e  y  d  k  n  k  d  g  y
TGAGCCAGTTATTAATCCAGATACTGGCGAAGAATACGAACCATATGAAACTCCTTATACCGGTGATGATATAAA 1425
 e  p  v  i  n  p  d  t  g  e  e  y  e  p  y  e  t  p  y  t  g  d  d  i  k
```

```
                                                  Sau3A
                                                          ①
AGAAAATCCGGATATTTTAGTGATTCCAGATGCTTCTACTAATCCAGATCCAAATCCAGAACCTACACCAGACCC 1500
 e  n  p  d  i  l  v  i  p  d  g  s  t  n  p  d  p  n  p  e  p  t  p  d  p
 ②
AAATCCAGAACCTACACCAGACCCAGACCCAAATCCAGATACGGGAAATGATCAAGATAAACTTCCTGAAAATGA 1575
 n  p  e  p  t  p  d  p  d  p  n  p  d  t  g  n  d  s  d  k  l  p  s  n  d

                                   TaqI
TACTCCAGTTTTAGATTTTTCTCCACTTTATGGAGCTACTCGAAAATTTCCTTTTTGTTTACCATGGGATATTTA 1650
 t  p  v  l  d  f  s  p  l  y  g  s  t  r  k  f  p  f  c  l  p  w  d  i  y
TGATTGTATTAAACTTTTCAAAAGTGAAGAAGAACCATTTAAATACGAATTTAAAGAAACTAAATATTCAGACAC 1725
 d  c  i  k  v  f  k  s  e  e  e  p  f  k  y  e  f  k  e  t  k  y  s  d  t

GGAAATTATTATAATTCCAAATTTTGAAATAGATTTTTCAAAGATACCTTATATTGATTTAATGAGAAATATTTT 1800
 e  i  i  i  i  p  n  f  e  i  d  f  s  k  i  p  y  i  d  l  m  r  n  i  l
AAAATTATTCCTTTATTATCAATTTATACGTTTCTTAATGATAAGCATAAGAAAGCTGATGTCGTCTTAATGTAT 1875
 k  l  f  l  y  y  q  f  i  r  f  l  m  i  s  i  r  k  v  m  w  s  .
                                                                ORF E  s  y
TCTTTTATAATAAAAACTATAAATGCTTTAATACGTGGATTAGGAGCTGTATTAACTTTTATTTTTAAAGTTTTA 1950
 s  f  i  i  k  t  i  n  s  l  i  r  g  l  g  s  v  l  t  f  i  f  k  v  l
CCAGATTCACCTTTTCAAAAGCATGTTTTAGAAAATCAAGTTATTATGCCTTATTTAAAAATGATTAATTATTTT 2025
 p  d  s  p  f  q  k  h  v  l  e  n  e  v  i  s  p  y  l  k  s  i  n  y  f
GTCCCTGTCGCTGAAATATTAATAACTTTTCAGGCTCTTTTAATTGCTATTGGTGTTTATTATATTTATCAAGTT 2100
 v  p  v  s  e  i  l  i  t  f  q  s  l  l  i  s  i  g  v  y  y  i  y  q  v

GTTGCTCGTTGGGTAAAAATTATTAGTTAGGAGTTTTTTTATATGATAAGTTTTTATAGTGGTACCCCTGGAAGT 2175
 v  s  r  w  v  k  i  i  s  .        ORF D  s  i  s  f  y  s  g  t  p  g  s
GCTAAATCTTTAAAAATGGCAAAAGAAGTTCTTATGTGCTTAAAGACATTTAAAAAGAATGTAATAGCAAACGTT 2250
 g  k  s  l  k  s  a  k  e  v  v  s  w  l  k  t  f  k  k  n  v  i  s  n  v
GAGATAGATAGAGATTATATTTTAAAAGGTAAGCAGGCTGGAGAATTTTTCTACCTTGCTAATGAAGAATTTTCT 2325
 e  i  d  r  d  y  i  l  k  g  k  q  g  g  s  f  f  y  l  s  n  e  s  f  s
CCAGATTATTTTTATGAGTATGCAATTAAATATCATGATATTGGTAAAGAACATCAAACACTTATTGTCGTTGAT 2400
 p  d  y  f  y  e  y  s  i  k  y  h  d  i  g  k  s  h  q  t  l  i  v  v  d

                                                          Eco RV
GAAGCACAGACTATTTTTTCCCCTAGTGCAGTAAAATTATTTAGTCAAAAAGAACCTAGATATCGTCAAAATTGG 2475
 e  s  q  t  i  f  s  p  s  s  v  k  l  f  s  q  k  s  p  r  y  r  q  n  w
                                                                   Sau3A
TTGGATTTTTTTACGCAGCATAGACATCTTGGGTTCGATATAATTTTAATTAGTCAATTCGATAGATTAATAGAT 2550
 l  d  f  f  t  q  h  r  h  l  g  f  d  i  i  l  i  s  q  f  d  r  l  i  d
CCACAAGTAAGATGTTTATTTGAATATAATTTTTGTACACAGAAAAGCGAATAATTTTGGATTGCTCGGACAAATT 2625
 p  q  v  r  c  l  f  s  y  n  f  v  h  r  k  s  n  n  f  g  l  l  g  q  i
```

```
TTTACTCTTATGCATATTAGTTTATTTGTTCAAATTCAAAAATGGTACGGTATGAATGAAGTATGTAGTAAAACA 2700
 f  t  l  m  h  i  m  l  f  v  q  i  q  k  w  y  g  m  n  e  v  c  m  k  t

                                                                 TaqI
TTTTATACTTATTCTAAGAAATACGCTAAGATATACAATTCATATGCTTTTAGAGAGAAAGTAATCGAAATATTG 2775
 f  y  t  y  m  k  k  y  m  k  i  y  n  m  y  m  f  r  e  k  v  i  e  i  l
ACACAGAAATATGGTTCTAAAAAAATGGAAGAGTTAATGGGGTTTAAAAATGCTAAAAAGGTATCTTAATAAATA 2850
 t  q  k  y  g  m  k  k  m  e  e  l  m  g  f  k  n  a  k  k  v  m  -
TGAGTTTTATGGTTTAGTTATTGATATATTCTTAATACTGTATGATTTTTGTCTTATAGTAAAATATTTCTAAAA 2925

ACATTTGACATTTCAGCATCATTATTATATCATATAAATATCAAGTTCATCTTTCACGTGAAACATTTAATAGTA 3000
 AatII
ACGACGTCGTTACAAAATATTTGGAGGTATTTATATGAGAGTTAATATTTCAATCCCAGATGAGCTTAAACAATT 3075
                                ORF C'  m  r  v  n  i  m  i  p  d  e  v  k  q  f
TTTTGAAGATTATAGTAAAAAAACAGGAGTTCCACAAAGTTCACTTATGGCTTTAGCTCTTTCTGAGTATAAAGA 3150
 f  m  d  y  m  k  k  t  g  v  p  q  m  m  l  m  a  l  m  l  m  e  y  k  d

                                                        HindIII
TAAGATAGAAAGGAGTTTATCTAATGACAAATAATGACATTATTGATGCTAATAAGCTTAGTCAGTTAACTAGTT 3225
 k  i  e  r  m  l  m  n  d  k
                        ORF C  m  t  n  n  m  i  i  d  m  n  k  l  m  q  l  t  m  m
CTATAGGTTCTATTAAATTAGAAATGAGAACATGTGCTATGTCTTTTTTTCGCATAGGAGAAAAACTCTTTAATA 3300
  i  g  m  i  k  l  m  m  r  t  c  m  m  m  f  f  r  i  g  m  k  l  f  n

TTCAAGAAAGTGAAGTTTATAAGATTAAAGGATATAAAAGTTTCGTTTCTTTTTGCAAAGGTGAGTTTGGAATTT 3375
i  q  m  m  v  y  k  i  k  g  y  k  m  f  v  m  f  c  k  g  m  f  g  i
CTAAAGTTCAATCTTATCGTTTTGTTGGTATTTATAAAAAATTTAGTTCTCAAAGATACTCTAGTTTACTTATT 3450
m  k  v  q  m  y  r  f  v  g  i  y  k  k  f  m  m  e  r  y  m  m  f  t  y
CACAATTGATAGAAATGCTTTCTCTTAAAGATTCTGATATAGATAAAGTTAATCCAGGTATGACTGTTAAGGAAA 3525
m  q  l  i  e  m  l  m  l  k  d  m  d  i  d  k  v  n  p  g  m  t  v  k  e

                                                Xmn I
TACGAACATTAAAAAAAGATTTAAAAGAATCTAAAGAAAAAGAAGATATTCAATTTGGAGAACAAACTGTTGTTG 3600
i  r  t  l  k  k  d  l  k  e  m  k  e  k  e  d  i  q  f  g  e  q  t  v  v

ACATTCAAGAAAGTGAAGTTGTTGAAGTTCCTGAAAAAACTGATAAAGAAATTATAGAAGAAAAAAATCTCAGAA 3675
d  i  q  e  m  e  v  v  e  v  p  e  k  t  d  k  e  i  i  e  e  k  n  l  r
TTTCAAATCAAGAACTTTTAATACAAAGATTATATGAGCGATTAAGAATTGAGAGTAAGAGAAGAAGATAAAAAAA 3750
i  m  n  q  e  l  l  i  q  r  l  y  e  r  l  r  i  e  m  k  e  k  d  k  k
TAGAAAAATTATCTGAAGAACTTAAAGAAACTCTAAATGAACTGAAAGCTCTTAAAGAGAAGAAAAACAGAGAAA 3825
i  e  k  l  m  e  e  l  k  e  t  l  n  e  l  k  m  l  k  e  k  k  n  r  e
```

Hind III

TACATAAGCTAAGAACTAAGCTTAAAGAACAAAGTAATAATTAAAGATTACAACTCTTCTTTCCCTAAAAGATTA 3900
i h k l r t k l k e q s n n

CACTTTGGCAGACCTTGTGTCGGCAATGCTTATGCTTATAAGTTAACGAATACTGGAATATATGTTGTTAGTGGT 3975

ACGGGCATCGCAGATAGTAGCGTAGCGTGCCATTAACTTTTTAAAATTTTCTTTCTATCTTGATATATATATATT 4050

AGAAATAAGTTAACCCATAAAAAAAATTTACTTTAGAAAAGTGGTGATAATTTGAGTTTAAAGGATTCTTTGAAA 4125

ORF B m s l k d s l k

AAAGCTGAAATAGATAAAGATTTATTAAAAGATATTATTAATAATATTGAATATAATAAGACAATAATTAAATAT 4200
k s e i d k d l l k d i i n n i e y n k t i i k y

TATGAACGTTTTTCTAAGGACTCAAGTATAGTATTTAAAAATACTATATTAGAAAATAAAAGAGATTCTATTTCT 4275
y e r f s k e s s i v f k n t i l e n k r d s i s

TTCTGTAATAAATATTGGCTTTTAGATAAATACGAAAAGCATAAGATAAAGATTTTAAGAAGACTAATTTATGT 4350
f c n k y w l l d k y e k h k i k d f k k t n l c

Hind III

CATGATAAATTTTGTAGTAATTGCAAAAAGGTTAAACAAGCTTCCAGAATGGCAAAATATATCCCGCAACTGAA 4425
h d k f c s n c k k v k q s s r s s k y i p q l e

CAGTATAAAGGTCAGTTGTATCATTTAACTTTAACATTACCTAATTGTTCTGCAAATGATTTAAGGTTAACTATA 4500
q y k g q l y h l t l t l p n c s g n d l r l t i

AAACATATGTCTAAATGTTTTTTAAGACTTGTACACTTTTTAGATGGTAGCAAGAAGATACGTGGTATAAATTT 4575
k h s s k c f l r l v q f l d g r k k i r g i n f

Hind III

GCTCAAGCTTTTCAAGGAGCAATAAGGTCTTTAGAAGTTACGTTTAAAAATCATAACTATCATCCACATTATCAC 4650
g q a f q g s i r s l e v t f k n d n y h p h y h

CTTGCATTAGTATTGAATAATTTTAAAATGACTGATAAAAAGTATAAGAATAAATATTCTTATAATAATAAACAT 4725
v s l v l n n f k s t d k k y k n k y s y n n k h

GGTATTAAAGAGTTGACTAGGCTTTTTTCAGCAGAAGAAATTTTAATTCAAAAGATATGGTATTTATTAATAAAT 4800
g i k e l t r l f s s e e i l i q k i w y l l i n

AATATAACTGTTACTAGACAAAACATTGAAGCTCTTAAAGATGGATATTCTTGTTGCATGGAAAAATTTTCAGAA 4875
n i t v t r q n i e s l k d g y s c c s e k f s e

AATGATTATGCAGAGCTTTTTAAATATATGACAAAGGTTACTGGTGAAGATGGTTTTACATTAACTTATGAAAAT 4950
n d y s e l f k y s t k v t g e d g f t l t y e n

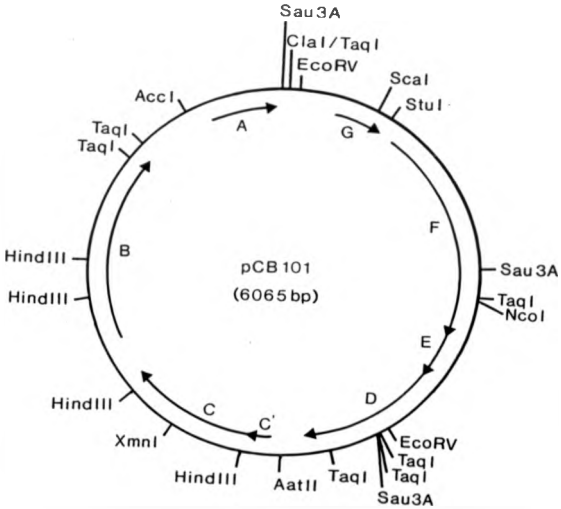
TTTGTTGCTTTATATTATGGACTTTATAGAATTAAACAGATTCAAGGATATGGAGTTTTGTATAATATAACTGAT 5025
f v e l y y g l y r i k q i q g y g v l y n i t d

GATGGAGATTTAGAAAAATTAGAAGAAGAGTATGAAAATTACATTGAACAGTTAAAACAAAAAGAAAGCCCAGTG 5100
d g d l e k l e e e y e n y i e q l k q k e s p v

```
                                                  Taq I
GCAAGTTATGAAGCTCCACAGGACTTAATAAAAGATACTGAATACATGTTAATTTCTCGAAAGTCATTTTTCAAG 5175
 a s y e a p q d l i k d t e y m l i s r k s f f k

                    --- --- --- - ---> . <---- - --- --- ---
TATCTCAATGAATAATTTTATTATAATACAAGGGGTAGAAAATTTCTATCTCTTTTATTTTAAAAGAAAATTTCT 5250
 y l n e .
          Taq I
TCAATAGTTTTTCGAGTGCATTAGCTATTTTTCGTGCTGTATCTAAATTACAAGTATTTTTTCTTTCTGCTTTT 5325

ACTATAGTGCTTAATGATAGTTCTGTTTTCCTTGCTAATTCTTCTTGAGTTAATCCTTTCTTAATTCTCATTTCT 5400

                    -- --- - ---> .  <-- - ------
TTAAGTGTCATTTTCTTTCTCCTTAAAATCTAATTTTACCCCTAATCAAATTTTAATTTAATTATAATCGCATATT 5475

                                                                  ----->
                                                 ----- - ------> -  <------ -
                  ---------- --  - --> . <- - --- ----------
GTAAAAATGTCCAGTTACTTCAATACTAGTAAAAAAATGATGGTAATTGTCTACTACTTTTTATTTACTACTACA 5550

 . <------
 -----
AAATTACTAGTTATAATAAGTATGTAAATAAGTAACGACATCGTTACTAAAAAAAATGGAGGTATTATTTATCAA 5625
                                                ORF A m e v l f m k

                                                 ---------
ACAAGATGCTTTAATAACTGGGGTTAAAAAGTACAATTTTAAAAACAATCATGGTGAAAATGTTAAGGGTGCTAA 5700
 q d m l i t g v k k y n f k n n d g e n v k g a k

                                 ----------
AATAAGTTATCTAATTGATGAATCAGACACCGATGGTGAAAATATTACTCGTTACTTACCTCGTCAAGCAAGTAT 5775
 i m y l i d e m d t d g e n i t g y l p r q a m i
TAAAGATATGTCTTTAATTTCTGGTTTAGTTGAACTTCCAGGAGTTTATGAAATTAATTTTAAAACTGTTTCAGG 5850
 k d m m l i m g l v e l p g v y e i n f k t v m g
AGCTAACAACAAAGAAGAATTAAAAATAATTGGTATAGATTTTCAGAAATCTTGTATTTTAAAAATGTTTTTGT 5925
 m n n k m m l k i i g i d f q k m c d f k n v f v
TGGTTAATTAAAATGGATAGCGATTTTTTTGAATTAGATTCTGAACAAGAAGTAACTACAGGTTCAGCAGTTGAC 6000
 G .
                    Sau3A                                   Cla I
                    <---                                    <---
ATTAGTTTAAAGTTGGACAAGGTCAATGATCTATTGTTAATTCAAAATATTTTTATCTTATCCGAT 6065
                                                            Taq I
```

FIGURE 3g PHYSICAL MAP AND GENETIC ORGANISATION OF pCB101

The arrows indicate the size and direction of the 8 open reading frames (ORFS). The locations of the major restriction endonuclease sites are shown.



### 3.2.5.2. GENETIC ORGANISATION OF pCB101

The principal feature of the nucleotide sequence was the presence of 8 open reading frames (ORFs) accounting for 77% of the plasmid. An identical figure has been reported for pIP404 (Garnier & Cole, 1988a). The ORFs were identified on the basis of size (> 40 codons), appropriate translational initiation signals (McLaughlin et al., 1981b; Moran et al.. 1982), overall codon usage. In addition, in vitro transcription and translation of some of the ORFs of pCB101 has also been achieved (3.2.7.3.).

All of the ORFs are contained on the same strand of pCB101 with predicted molecular weights of the ORF products ranging from 5,192 Da for ORF C' to 49,450 Da for ORF F (Table 3.1). ORFs A, B, and G appear to be monocistronic, whereas ORFs D, E, and F and ORFs C' and C might by polycistronic owing to their close proximity (Fig. 3f, 3g).

The average dG + dC content of the ORFs, 28%, is comparable to that of the plasmid as a whole (27.2%). The % dG+dC content of the intergenic regions is not significantly lower as reported for pIP404 (Garnier & Cole, 1988a), with the exception of one region lying within nucleotide positions 4013 and 4090 (Fig. 3f). This region is 86 % dA + dT and includes the 12 bp palindromic and repetitive sequence (5'-ATATATATATATAT-3') found at the deletion junction of the pRB1 deletion variants (3.2.3.2.).

### 3.2.5.3. TRANSLATIONAL SIGNALS

The 5' ends of the 8 ORFs were screened for sequences complementary to the 3' end of the 16 S rRNA of B. subtilis and in all cases a suitably positioned ribosome binding

site was found (Table 3.2). As is common for translation initiation signals for genes from Gram-positive bacteria (McLaughlin et al.. 1981b; Moran et al.. 1982) significant complementarity is evident. Similar observations have been reported for pIP404 (Garnier & Cole, 1988a).

The codon usage of the 8 ORFs was analysed using the programmes of DNASTAR Inc. and is presented in Table 3.3. As expected for an organism with a genome rich in dA + dT, there is a pronounced preference for codons containing these nucleotides. This is particularly apparent in the "wobble" position where 86% of the codons end in dA or dU; an identical observation is reported for pIP404 (Garnier & Cole, 1988a). However, when a choice between dA or dU exists it is generally dU that is chosen; the converse preference is reported for pIP404 (Garnier & Cole, 1988a). When the frequency of codon utilisation in pCB101 was compared with that of E. coli (Ikemura, 1981; Grosjean & Fiers, 1982), major differences were seen for most amino acids with the exception of those encoding lysine and glutamine (Table 3.3) Interestingly, the codon usage displayed by pCB101 is closer to that of B. subtilis (Ogasawara, 1985).

The frequency of amino acid utilisation for pCB101 and pIP404 was also examined. As can be seen from Table 3.4, the frequencies for strongly basic, strongly acidic, hydrophobic, and polar amino acids utilised by the two plasmids are very similar.

All of the ORFs of pCB101, with exception to ORF B, possess typical methionine start codons (AUG); ORF B, however, starts with the atypical methionine codon TTG (Shuttleworth et al.. 1987).

TABLE 3.1

CALCULATED MOLECULAR WEIGHTS OF THE PUTATIVE pCB101 ORF PRODUCTS

| ORF | MOLECULAR WEIGHT (Daltons) |
|---|---|
| A | 11,383 |
| B | 42,910 |
| C' | 5,192 |
| C | 27,134 |
| D | 27,589 |
| E | 9,924 |
| F | 49,470 |
| G | 8,742 |

TABLE 3.2

LOCATION OF REGIONS UPSTREAM OF THE OPEN READING FRAMES OF pCB101 AND PUTATIVE TRANSLATIONAL INITIATION SIGNALS

Boldface letters indicate sequences complementary to the 3'-end of *B subtilis* 16 S rRNA (Moran *et al*., 1982). Initiation codons are underlined.

| ORF | LOCATION | SEQUENCE |
|---|---|---|
| A | 5601-5623 | 5'-AAAAAATGGAGGTATTATTTATG-3' |
| B | 4082-4104 | 5'-TTTAGAAAAGTGGTGATAATTTG-3' |
| C' | 3015-3037 | 5'-AAATATTTGGAGGTATTTATATG-3' |
| C | 3154-3176 | 5'-GATAGAAAGGAGTTTATCTAATG-3' |
| D | 2127-2145 | 5'-TTAGTTAGGAGTTTTTTTATATG-3' |
| E | 1850-1872 | 5'-AGAAAGGTGATGTGGTCTTAATG-3' |
| F | 546-568 | 5'-ATTTTAAATAGAGGTGTTATATG-3' |
| G | 197-219 | 5'-AAAAAATTGGAGGTAATGTAATG-3' |
| 3'-end of *B. subtilis* 16 S rRNA | | 3'-UCUUUCCUCCACUAG-5' |

TABLE 3.3

CODON UTILISATION OF THE ORFs OF pCB101

AND COMPARISON WITH pIP404, E. COLI AND B. SUBTILIS

| AMINO ACID | CODON | OCCURRENCE IN OPEN READING FRAME: A | B | C' | C | D | E | F | G | TOTAL | % USE OF ONE CODON RELATIVE TO ALL OTHERS WITHIN ONE AMINO ACID GROUP: pCB101 % | * pIP404 % | * E. coli % | * B. subtilis % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Arg | CGU | 1 | 2 | 0 | 1 | 1 | 2 | 3 | 0 | 10 | 22.2 | 3.1 | 58.0 | 35.2 |
| | C | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 2.2 | 0.0 | 35.0 | 17.5 |
| | A | 0 | 1 | 0 | 2 | 0 | 0 | 1 | 0 | 4 | 8.8 | 2.1 | 2.3 | 9.1 |
| | G | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 1.0 | 3.2 | 11.1 |
| | AGA | 0 | 5 | 1 | 7 | 7 | 0 | 5 | 0 | 25 | 55.5 | 80.2 | 1.1 | 27.7 |
| | G | 0 | 4 | 1 | 0 | 0 | 0 | 0 | 0 | 5 | 11.1 | 13.6 | 0.3 | 9.4 |
| Leu | CUU | 1 | 7 | 2 | 7 | 4 | 1 | 9 | 1 | 32 | 24.2 | 12.1 | 8.6 | 26.1 |
| | C | 0 | 1 | 0 | 2 | 1 | 0 | 0 | 0 | 4 | 3.0 | 0.0 | 6.6 | 9.8 |
| | A | 1 | 0 | 0 | 2 | 0 | 0 | 1 | 0 | 4 | 3.0 | 9.3 | 1.8 | 6.3 |
| | G | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 2 | 1.5 | 1.2 | 64.0 | 21.8 |
| | UUA | 5 | 25 | 2 | 8 | 9 | 8 | 19 | 4 | 80 | 60.6 | 71.3 | 5.8 | 22.4 |
| | G | 0 | 6 | 0 | 1 | 3 | 0 | 0 | 0 | 10 | 7.5 | 6.1 | 8.1 | 13.6 |
| Ser | UCU | 3 | 10 | 1 | 12 | 5 | 1 | 18 | 3 | 53 | 43.8 | 18.4 | 26.6 | 24.5 |
| | C | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 3 | 2.4 | 2.8 | 25.6 | 12.0 |
| | A | 2 | 4 | 2 | 2 | 1 | 1 | 9 | 2 | 23 | 19.0 | 29.6 | 8.3 | 18.7 |
| | G | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0-0 | 0.0 | 0.6 | 11.4 | 10.0 |
| | AGU | 2 | 4 | 3 | 9 | 8 | 1 | 9 | 2 | 38 | 31.4 | 41.9 | 6.5 | 10.7 |
| | C | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 1 | 4 | 3.3 | 6.7 | 21.6 | 24.1 |
| Thr | ACU | 3 | 12 | 0 | 7 | 3 | 3 | 17 | 9 | 54 | 64.2 | 41.1 | 23.8 | 14.9 |
| | C | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 2 | 2.3 | 5.3 | 50.6 | 14.1 |
| | A | 0 | 4 | 1 | 3 | 4 | 0 | 9 | 2 | 23 | 27.3 | 49.1 | 5.8 | 43.3 |
| | G | 0 | 1 | 0 | 0 | 2 | 0 | 2 | 0 | 5 | 5.9 | 4.5 | 19.7 | 27.9 |
| Pro | CCU | 1 | 1 | 0 | 1 | 3 | 3 | 10 | 0 | 19 | 33.9 | 39.7 | 9.0 | 33.6 |
| | C | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0 | 0.0 | 6.0 | 9.8 |
| | A | 1 | 3 | 2 | 1 | 2 | 1 | 25 | 0 | 35 | 62.5 | 55.5 | 19.9 | 19.1 |
| | G | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 2 | 3.5 | 4.8 | 65.1 | 37.5 |
| Ala | GCU | 3 | 6 | 2 | 3 | 4 | 6 | 7 | 5 | 36 | 60.0 | 52.5 | 27.9 | 27.5 |
| | C | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 3.3 | 6.9 | 18.7 | 20.0 |
| | A | 1 | 6 | 0 | 0 | 5 | 0 | 4 | 5 | 21 | 35.0 | 57.7 | 22.9 | 27.1 |
| | G | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1.6 | 3.3 | 30.5 | 25.4 |
| Gly | GGU | 7 | 7 | 0 | 4 | 7 | 1 | 12 | 5 | 43 | 53.7 | 30.3 | 47.8 | 25.4 |
| | C | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 3 | 3.7 | 4.2 | 40.8 | 29.6 |
| | A | 2 | 7 | 1 | 4 | 4 | 2 | 6 | 5 | 31 | 38.7 | 58.2 | 4.6 | 31.5 |
| | G | 1 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 3 | 3.7 | 7.1 | 6.8 | 13.5 |

| AMINO ACID | CODON | OCCURRENCE IN OPEN READING FRAME: A | B | C* | C | D | E | F | G | TOTAL | % USE OF ONE CODON RELATIVE TO ALL OTHER WITHIN ONE AMINO ACID GROUP: pCB101 % | pIP404 % * | E. coli % * | B. subtilis % * |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Val | GUU | 7 | 7 | 3 | 11 | 5 | 6 | 21 | 5 | 65 | 72.2 | 45.3 | 37.5 | 31.4 |
| | C | 0 | 0 | 0 | 0 | 1 | 2 | 0 | 0 | 3 | 3.3 | 2.0 | 12.8 | 24.7 |
| | A | 0 | 3 | 0 | 0 | 7 | 2 | 3 | 3 | 18 | 20.0 | 50.0 | 22.9 | 24.5 |
| | G | 0 | 1 | 0 | 0 | 0 | 0 | 3 | 0 | 4 | 4.4 | 2.7 | 26.8 | 19.4 |
| Ile | AUU | 6 | 13 | 1 | 12 | 10 | 8 | 18 | 7 | 75 | 55.9 | 36.2 | 37.3 | 50.0 |
| | C | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 4 | 2.9 | 3.2 | 62.2 | 39.4 |
| | A | 4 | 14 | 1 | 9 | 7 | 6 | 12 | 2 | 55 | 41.0 | 60.6 | 0.5 | 10.6 |
| Lys | AAA | 10 | 31 | 4 | 28 | 19 | 4 | 22 | 5 | 123 | 76.4 | 80.6 | 76.7 | 75.4 |
| | G | 2 | 16 | 1 | 6 | 6 | 1 | 3 | 3 | 38 | 23.6 | 19.4 | 23.3 | 24.6 |
| Asn | AAU | 6 | 24 | 1 | 10 | 9 | 3 | 29 | 1 | 83 | 91.2 | 86.7 | 24.2 | 53.1 |
| | C | 3 | 2 | 0 | 1 | 1 | 0 | 1 | 0 | 8 | 8.8 | 13.3 | 75.8 | 46.9 |
| Gln | CAA | 2 | 8 | 2 | 9 | 8 | 2 | 6 | 1 | 38 | 73.0 | 9.9 | 26.6 | 54.2 |
| | G | 1 | 6 | 0 | 1 | 4 | 1 | 1 | 0 | 14 | 27.0 | 10.1 | 73.4 | 45.8 |
| His | CAU | 0 | 7 | 0 | 1 | 5 | 1 | 2 | 0 | 16 | 88.8 | 75.0 | 38.9 | 68.6 |
| | C | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 2 | 11.2 | 25.0 | 61.1 | 31.4 |
| Glu | GAA | 7 | 22 | 2 | 26 | 11 | 3 | 22 | 1 | 94 | 81.7 | 75.8 | 73.4 | 69.5 |
| | G | 0 | 4 | 2 | 6 | 4 | 0 | 4 | 1 | 21 | 18.3 | 23.2 | 26.6 | 30.5 |
| Asp | GAU | 7 | 19 | 3 | 8 | 9 | 1 | 27 | 2 | 76 | 87.4 | 89.9 | 51.0 | 63.8 |
| | C | 1 | 1 | 0 | 1 | 0 | 0 | 8 | 0 | 11 | 12.6 | 10.1 | 49.0 | 36.2 |
| Tyr | UAU | 2 | 25 | 2 | 6 | 12 | 6 | 28 | 1 | 82 | 84.5 | 93.5 | 40.6 | 61.8 |
| | C | 2 | 3 | 0 | 1 | 4 | 0 | 4 | 1 | 15 | 15.5 | 6.5 | 59.4 | 36.2 |
| Cys | UGU | 1 | 6 | 0 | 1 | 2 | 0 | 5 | 1 | 16 | 84.2 | 88.5 | 42.0 | 45.7 |
| | C | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 3 | 15.8 | 11.5 | 58.0 | 54.3 |
| Phe | UUU | 5 | 16 | 2 | 9 | 17 | 6 | 25 | 5 | 85 | 90.4 | 83.2 | 43.5 | 64.0 |
| | C | 0 | 2 | 0 | 1 | 3 | 0 | 3 | 1 | 9 | 9.6 | 16.8 | 56.5 | 36.0 |
| Met | AUG | 2 | 6 | 2 | 5 | 7 | 3 | 6 | 1 | 33 | 100 | 100 | 100 | 100 |
| Trp | UGG | 0 | 2 | 0 | 0 | 3 | 1 | 4 | 0 | 10 | 100 | 100 | 100 | 100 |
| | | 103 | 361 | 45 | 231 | 233 | 86 | 434 | 86 | 1579 | | | | |

* Taken from Garnier & Cole, 1988a

TABLE 3.4

FREQUENCY (%) OF AMINO ACID UTILISATION IN THE PROTEINS AND PUTATIVE ORF PRODUCTS BY pCB101 and pIP404

| | | | pCB101 (AV) | pIP404* (AV) |
|---|---|---|---|---|
| A. | STRONGLY BASIC AMINO ACIDS | LYSINE | 10.2 | 10.1 |
| | | ARGININE | 2.6 | 3.7 |
| B. | STRONGLY ACIDIC AMINO ACIDS | ASPARTATE | 4.8 | 5.3 |
| | | GLUTAMATE | 6.81 | 7.1 |
| C. | HYDROPHOBIC AMINO ACIDS | ALANINE | 4.8 | 4.7 |
| | | ISOLEUCINE | 9.3 | 9.7 |
| | | LEUCINE | 8.2 | 9.4 |
| | | PHENYLALANINE | 5.9 | 4.3 |
| | | TRYPTOPHAN | 0.46 | 0.9 |
| | | VALINE | 6.7 | 5.7 |
| D. | POLAR AMINO ACIDS | ASPARAGINE | 4.8 | 7.5 |
| | | CYSTEINE | 0.87 | 1.0 |
| | | GLUTAMINE | 3.3 | 2.6 |
| | | SERINE | 7.9 | 6.9 |
| | | THREOCINE | 5.18 | 4.3 |
| | | TYROSINE | 5.27 | 5.3 |

* Garnier and Cole, 1988a

### 3.2.5.4. PUTATIVE TRANSCRIPTIONAL REGULATORY SIGNALS

The sequences upstream of the ORFs were screened for possible promoters by comparison with the consensus sequence for "extended" promoter elements for Gram-positive genes presented by Graves & Rabinowitz (1986). Likely candidates are shown in Table 3.5. Sequence conservation outside of the cananonical -10 and -35 regions appears to be evident, particularly around the -10 region, e.g., T's at -16 and -5, and A's at -6, and -4.

Likewise, the nucleotide sequence of pCB101 was examined for appropriately situated dyad symmetries which could correspond to rho-independent transcriptional terminators. The sequences of two candidates are shown in Table 3.6. The corresponding values of free energy of interaction ($\Delta G$) are given. The best candidate, which follows ORF G ($\Delta G$ = -31 Kcal $mol^{-1}$), is also followed by a short poly U stretch.

### 3.2.5.5. ADDITIONAL SEQUENCE FEATURES

Using the computer programs of DNASTAR inc., the remainder of the pCB101 sequence was examined for sequences that might contribute to the formation of secondary structure either at the DNA or RNA level, or perhaps have some regulatory role, i.e., direct repeats and palindromic elements.

The N-terminal portion of ORF B (nt: 4227-4279, Fig. 3f, 3h) contains a sequence that could form two adjacent loops with $\Delta G$ values of -5.8 and -1.2 Kcal $mol^{-1}$ according to the calculations of Tinoco et al. (1973). The formation of a stable structure is implied here as this region appeared to be responsible for "pile ups" in the nucleotide sequencing reactions. Approximately 300 bp downstream of ORF B there

Table 3.5

PUTATIVE PROMOTER SEQUENCES OF pCB101

| ORF | LOCATION | -45 | -35 | | -15 | -10 | |
|---|---|---|---|---|---|---|---|
| A | 5528-5577 | ACTAGTTTTTAT | TTACTA | GTACAAA | ATTACTAGT | TATAAT | AAGTATGTAA |
| b | 4003-4052 | CTTTGGCAGACC | TTGTGT | CGGCAAT | GCTTATGCT | TATAAG | TTAACGAATA |
| C' | 2918-2968 | TTCTAAAAAGAT | TTGACA | TTTCAGC | ATCATTATTA | TATCAT | ATAAATATCA |
| C | 3108-3158 | CACAAAGTTCAC | TTATGG | CTTTAGC | TCTTTCTGAG | TATAAA | GATAAGATAG |
| D | 2051-2101 | CTTTTCAGGCTC | TTTTAA | TTGCTAT | TGGTGTTTAT | TATATT | TATCAAGTTG |
| E | 1796-1849 | ATTTTAAAATTA | TTCCTT | TATTATC | AATTTATACG | TTTCTT | GATAAGCATA |
| F | 508-577 | TTGGCGCAGGCC | TTTAAT | TGTGCCG | AAAATCTTT | TTTTAT | TTTAAATAGA |
| G | 146-195 | TCCGTTTTGGAA | TTGACT | TACTTTT | ATCTGTTTT | TAAAAT | AATAGCAAAA |
| Gram-positive consensus * | | Ta AAAAA | TTGAcA | a A | a T TG | TATAAT | AAtAt |
| E. coli consensus * | | a t | TTGACa | t t | t tg | TAtAAT | |

Bold face letters indicate sequence complementarity with the consensus sequence for Gram-positive promoters (Graves & Rabinowitz, 1986*). Height of letters in consensus sequence indicates relative degree of conservation.

Table 3.6

POTENTIAL TRANSCRIPTIONAL TERMINATORS

$\Delta G$ values calculated according to Tinoco *et al.* (1973).

| ORF | nt position | |
|---|---|---|
| ORF B | 5196-5237 | TTATAATACAAGGGGTAGAAAATTTCTATCTCTTTTATTTTAA |
| | | $\Delta G$ = -6.4 kcal $mol^{-1}$ |
| ORF G | 486-547 | ATAAAAAAAGATTTTCGGCATATTGGCGCAGGCCTTTAAATGTGCCGAAAATCTTTTTTTAT |
| | | $\Delta G$ = -31 kcal $mol^{-1}$ |

is another region (nt: 5499-5456; Fig. 3f) which exhibits a capability of forming three different loop structures with addition, some 50 bp upstream of this region lies a sequence (nt: 5423-5456; Fig. 3f) capable of forming a loop with a $\Delta G$ value of -4.4 Kcal/mol (Fig. 3j).

A number of other palindromic elements were also identified. The 12 bp palindromic element previously described (3.2.3.2.) lies some 50 bp upstream of ORF B (nt: 4038-4049; Fig. 3f). Within ORF B two palindromic elements comprised solely of adenine and thymine bases are evident: a 16 bp and an 18 bp palindrome (nt: 4157-4172, and 4786-4801, Fig. 3f). A tandem pair of palindromic elements, 18 and 16 bp in length, are found some 20 bp upstream of ORF C' (nt: 2974-3013, Fig. 3f), lying between the ribosome binding site and the putative promoter sequences (Fig. 3f). A 14 bp palindrome is evident within ORF D (nt: 2354-2367, Fig. 3f) as is a 12 bp element within ORF G (nt: 320-332, Fig. 3f).

A very striking pair of overlapping direct repeat sequences (27 and 17 bp) are found within ORF F (nt: 1475-1525, Fig. 3f). A pair of 8 bp direct repeat sequences flank the start codon of ORF D (nt: 2136-2158, Fig. 3f). Within ORF C lies a pair of 18 bp direct repeat sequences which are separated by some 286 nucleotides (nt: 3300-3317 and 3603-3620, Fig. 3f). ORF A contains a small 12 bp repeated sequence separated by some 46 nucleotides (nt: 5675-5686 and 5732-5743, Fig. 3f).

FIGURE 3h

pCB101 DNA SEQUENCE: SECONDARY STRUCTURE PREDICTION

DNA SEQUENCE IN N-TERMINAL PORTION OF ORF B RESPONSIBLE FOR SEQUENCE PILE UPS

Nucleotide position: 4227 -> 4279; See Fig. 3f

ΔG values calculated according to Tinoco et al., (1973)

```
                                A
                              G   T
                              A   T
                                G : C
                                A : T
        A   A                 A     A
      T       A                 A : T
        T : A                   A : T
        T : A                 T     T
        A : T                       C
        T : A                   A : T
        G : C                   A : T
        A : T                   A : T
        T : A                   G : C
        A : T                   A : T
C A A G T : A T - - - - - - - - - - - T   G T A A T
```

$\Delta G = -5.8$ Kcal.mol$^{-1}$ $\Delta G = -1.2$ Kcal.mol$^{-1}$

FIGURE 3i pCB101 DNA SEQUENCE: SECONDARY STRUCTURE PREDICTION

DNA SEQUENCE 300 bp DOWNSTREAM OF ORF B CAPABLE OF FORMING
3 DIFFERENT STRUCTURES

ΔG values calculated according to Tinoco et al. (1973)

(A) Nucleotide position: 5518 - 5557

```
  T   A
 T     T
T       T
 T     T
  T : A
  G : C
  A : T
  T : A
  C : G
  A : T
  T : A
C       C
  T : A
G       A
  T : A
  T : A
  A : T
  A : T
  T : A
  G : C
```

$\Delta G = -10.2$ kcal.mol$^{-1}$

(B) Nucleotide position: 5583 - 5561

```
  A   A
A       A
C       T
  A : T
  T : A
  G : C
  A : T
  T : A
  C : G
  A : T
```

$\Delta G = -6.8$ kcal.mol$^{-1}$

(C) Nucleotide position: 5499 - 5548

```
  T   G
T       T
A       C
  A : T
  T : A
  G : C
G       T
  T : A
A       G
G         T
T       T
  A : T
  A : T
  A : T
A       A
  A : T
  A : T
  A : T
  T : A
  G : C
  A : T
  T : A
  C : G
  A : T
  T : A
```

$\Delta G = -10.6$ kcal.mol$^{-1}$

FIGURE 3j

pCB101 SEQUENCE: SECONDARY STRUCTURE PREDICTION

Nucleotide position: 5423 - 5456

ΔG value calculated according to Tinoco *et al.*, (1973)

```
      C
  A       C
T           C
  T       T
    T : T
    T : A
    A : A
  A       C
    T : A
  C       A
    A : T
    A : T
    A : T
    A : T
    T : A
    T : A
```

$\Delta G = -4.4$ Kcal.mol$^{-1}$

### 3.2.6. FURTHER CHARACTERISATION OF THE pCB101 MINIMAL REPLICON

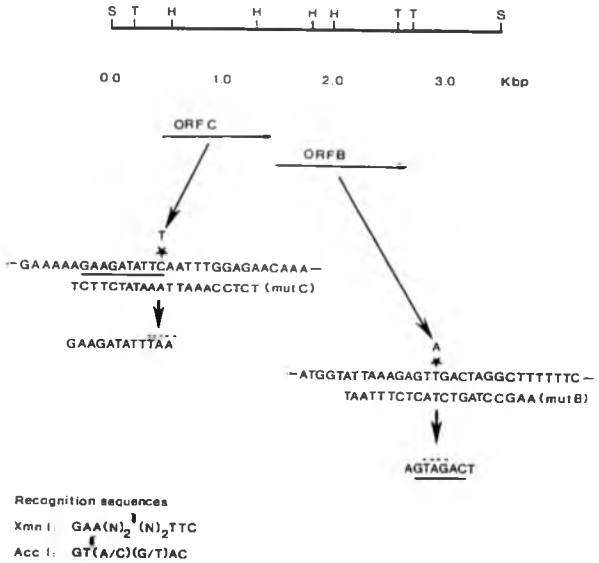
Having obtained the complete nucleotide sequence of pCB101 it was elected to further analyse the minimal replicon; i.e., the 3.48 Sau3A fragment possessing the two major ORFs B and C. Initially, it was elected to try and ascertain which of the translation products of ORF B and ORF C are essential to replication. The strategy chosen was to employ site-directed mutagenesis to create individual point mutations in the two ORFs resulting in the generation of translational stop codons. Subsequent delivery of the modified 3.48 kb Sau3A fragments into *B. subtilis* on board a Gram-positive replication-deficient plasmid vector would therefore identify the ORF(s) responsible for the establishment of autonomous replication.

#### 3.2.6.1. SITE-DIRECTED MUTAGENESIS OF THE pCB101 MINIMAL REPLICON

Two 21 base mutagenic oligonucleotides, mutB and mutC, were synthesised. The mutB oligonucleotide (5'-TAATTTCTCATCT-GATCCGAA-3') was designed to change the "T" at position 4739 of the pCB101 map (Fig. 3f, 3k) to an "A" thus creating both an in frame translational stop codon (TAG) within ORF B and a unique AccI site. Similarly, the mutC oligonucleotide (5'-TCTTCTATAAATTAAACCTCT-3') was designed to change the "C" at position 3576 of the pCB101 map (Fig. 3f, 3k) to a "T", resulting in the creation of an in frame translational stop codon (TTA) in ORF C with the concomitant loss of an XmnI site. The altered restriction enzyme sites were to aid in the later screening of successful mutants by digestion of M13 RF DNA.

FIGURE 3k SITE-DIRECTED MUTAGENESIS STRATEGY TO CREATE FRAME-SHIFT MUTATIONS IN THE pCB101 ORFS B AND C

Translational stop codons and restriction endonuclease recognition sequences are indicated by dashed and solid lines respectively.




Recognition sequences

Xmn I: $GAA(N)_2(N)_2TTC$

Acc I: $GT(A/C)(G/T)AC$

As a prerequisite for in vitro site-directed mutagenesis, some of the previously isolated 3.48 kb Sau3A fragment DNA (3.2.2.) was sub-cloned into M13mp9. Template DNA was prepared from 6 recombinant plaques and each subjected to DNA sequence analysis. Sequencing with universal sequence primer revealed that all six clones were identical, the orientation being such that the coding strand of the pCB101 DNA was generated as the "+" viral single-strand form.

The mutagenesis experiments with both mutB and mutC oligonucleotides were carried out with the prepared template DNA. The calculated dissociation temperatures for mutB and mutC mutagenic oligonucleotides were 56°C and 52°C respectively. Template DNA was prepared from four plaque-purified hybridisation-positive isolates for each mutagenesis experiment. The sequence of the mutants was checked by DNA sequencing employing custom synthesised oligonucleotide primers (primers "a" and "b") complementary to the pCB101 sequence at nucleotide positions 2281-2297 and 3978-3994 (Fig. 3f) for mutB and mutC mutations respectively (Table 3.7). Successful mutants for each experiment were chosen and employed for large scale preparation of M13 RF DNA. Subsequent digestion of the RF DNA preparations with Sau3A and AccI for mutB, and Sau3A and XmnI for mutC, followed by agarose gel electrophoresis (1% w/v) of the restricted products alongside similarly cleaved M13mp9 RF DNA, revealed that the majority of the pCB101 DNA had been deleted.

As a consequence of the observed deletions arising in the prepared M13 RF DNA it was elected to analyse the nucleotide sequences of the mutant template DNA preparations that were used for the RF preparations in addition to the original M13mp9 template DNA that was employed as substrate in

the mutagenesis experiments. As available template DNA was limited, initially only three additional 17'mer oligonucleotide primers (primers "c", "d", and "e"; Table 3.7) complementary to widespread regions of the pCB101 Sau3A fragment were synthesised. Subsequent DNA sequencing employing primers "a", "b", "c", "d", "e," (Table 3.7) and universal sequence primer, of most of the remaining mutant template DNA revealed that only pCB101 DNA 5' to the regions selected in the mutant screening was present (Fig. 31). Sequencing of the non-mutant substrate template revealed the presence of pCB101 DNA throughout, but with an underlying faint sequence resembling a possible deleted derivative. Aliquots (1 $\mu$l) of the remaining template DNA preparations were then used to transfect competent E. coli TG1 for the isolation of additional template DNA for further sequence analysis. Five additional oligonucleotide primers ("f", "g", "h", "i", and "j", Table 3.7), complementary to the pCB101 sequence at positions to facilitate the sequencing of the remainder of the Sau3A insert DNA, were synthesised. The newly prepared template preparations were then sequenced with universal primer and the 10 additional primers. This revealed that all of the pCB101 derived DNA had now been deleted from M13 implying an increased instability of the cloned clostridial DNA in M13 with further growth in E. coli TG1.

As similar instability problems with M13 have been resolved by employing the recA equivalent strain of E. coli TG1, E. coli TG2 (G.P.C. Salmond, personal communication), the mutagenesis experiments were repeated in this strain. Both mutagenesis experiments were repeated exactly as performed with E. coli TG1. However, upon nucleotide sequence analysis of successful mutants, the same instability problems were encountered.

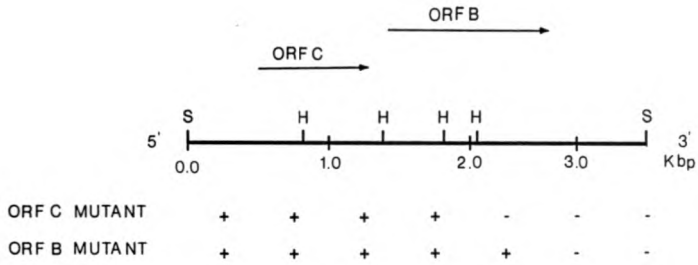
TABLE 3.7

LIST OF CUSTOM SYNTHESISED OLIGONUCLEOTIDE SEQUENCE PRIMERS

| PRIMER | LOCATION OF pCB101 COMPLEMENT |
|---|---|
| "a" | 2281 - 2297 |
| "b" | 3668 - 3684 |
| "c" | 4218 - 4234 |
| "d" | 4278 - 4294 |
| "e" | 3252 - 3268 |
| "f" | 5110 - 5126 |
| "g" | 5403 - 5419 |
| "h" | 3978 - 3994 |
| "i" | 2931 - 2947 |
| "j" | 4531 - 4547 |

FIGURE 31 SCHEMATIC REPRESENTATION OF DNA SEQUENCE ANALYSIS OF THE pCB101 ORF B AND ORF C MUTANTS

+/- denotes the presence and absence of pCB101 DNA



### 3.2.6.2. CONSTRUCTION OF THE ORF C FRAME-SHIFT MUTANT, pM21C127

An alternative approach to characterise ORFs B and C was devised whereby frame-shift mutations could be introduced separately into the two ORFs by sub-cloning. Subsequent delivery of the mutated Sau3A fragments into B. subtilis would therefore identify the ORF(s) essential to replication. By this stage of the investigation collaborative studies undertaken at the University of Aberystwyth had provided evidence that ORF B was essential for replication of pCB101 based plasmids in B. subtilis (Brehm et al., 1986). The three regions of the pCB101 Sau3A fragment depicted in Fig. 3m were cloned into the Gram-positive replication deficient plasmid pJH101. Of the three chimaeras constructed only one, pAP20 (Fig. 3m), was incapable of autonomous replication in B. subtilis. Thus the replicative functions of pCB101 were more accurately assigned to a 2.5 kb fragment, between the TaqI site at position 2765 and the TaqI site at position 5263 (Fig. 3f, 3m). Furthermore the loss of 106 bp TaqI fragment (nt: 5157-5263) appeared to destroy the replicative functions of pCB101 (i.e., pAP20 is rep$^-$). The simplest explanation for this impairment of replication is the removal of 9 codons from the C-terminus of ORF B. The functional significance of ORF C, however, remained to be determined.

Some of the previously isolated Sau3A fragment DNA (3.2.2.) was ligated with the multicopy cloning vector pMTL21C (Fig. 3n) which had been cleaved with BamHI and dephosphorylated. The ligation products were used to transform competent E. coli JM83 and small scale plasmid DNA preparations were isolated from six recombinant clones. Restriction enzyme analysis of these preparations with a combination of Hin-

dIII and EcoRI revealed that all six isolates were identical, with the Sau3A fragment inserted in pMTL21C in the orientation shown in Fig. 3o. One of these isolates, pM21C125, was employed for the large scale isolation of plasmid DNA.

Approximately 50 µg pM21C125 DNA was digested to completion with XmnI and BglII and the products size fractionated on a preparative agarose (1% w/v) gel. The three smallest fragments generated (1.82, 1.84, and 0.8 kb) were extracted from the gel by electroelution and self ligated to yield the derivative pM21C126, some 2.5 kb smaller than the parental plasmid (Fig. 3p). Small scale plasmid preparations were isolated from four clones and subjected to restriction enzyme analysis with XmnI. All four isolates were found to be identical giving the expected restriction digestion pattern. One of the isolates was employed for the large scale isolation of plasmid DNA.

Approximately 50 µg amounts of pM21C125 and pM21C126 plasmid preparations were digested with combinations of ScaI/XmnI and StuI/ScaI respectively, and the digestion products size-fractionated on separate agarose (1% w/v) gels. The 4.25 kb ScaI/XmnI fragment of pM21C125 and the 2.2 kb StuI/ScaI and the 0.6 kb ScaI/StuI fragments of pM21C126 were isolated from the gels by electroelution. Subsequent ligation of these isolated fragments, followed by transformation of *E. coli* JM83, yielded the construct pM21C127 (Fig. 3q). Four isolates were chosen and small scale plasmid preparations derived from them. Subsequent restriction enzyme analysis with BglII and XhoI confirmed that all four isolates were pM21C127 as depicted in Fig. 3q.

FIGURE 3m DIAGRAMATIC REPRESENTATION OF THE 3.48 Kb SAU3A SUBCLONES

The restriction endonuclease sites are as follows: AccI, A; HindIII, H; Sau3A, S; and TaqI, T. Rep +/- denotes the ability to replicate in B. subtilis

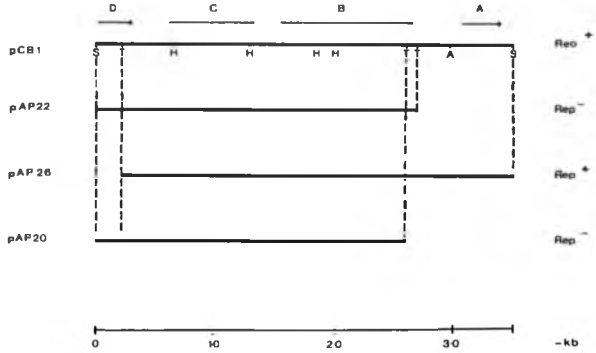
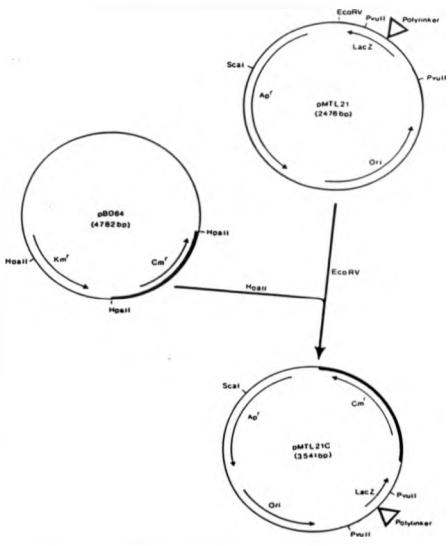


FIGURE 3n DERIVATION OF pMTL21C

Chimaeric plasmid pMTL21C was constructed by the combination of the chloramphenicol resistance determinant of pBD64 (Gryczan et al, 1980) as a 1070 bp HpaII fragment, and the multicopy E. coli cloning vector pMTL21 (Chambers et al, 1988). This vector offers blue/white selection of recombinants on X-gal containing media as the unique EcoRV site lies external to the polylinker region. The sequence of the polylinker, displaying all the unique cloning sites, is given below.




```
                                                  AccI
                                                  SalI
                                   XhoI                                   SmaI
                HindIII SphI PstI StuI    BglII NcoI  MluI  AscII  XbaI  BamHI   KpnI  SacI  EcoRI
ATGACCATGATTACCCCAAGCTTGCATGCCTGCAGGCCTCGAGATCTCCATGCACGCCTGACCTCGACTCTAGACGATCCCCCGCTACCCACCTCGAATTCACTCGCC
 M  T  M  I  T                                                                                          L  A
```

Polylinker

FIGURE 3o DERIVATION OF pM21C125

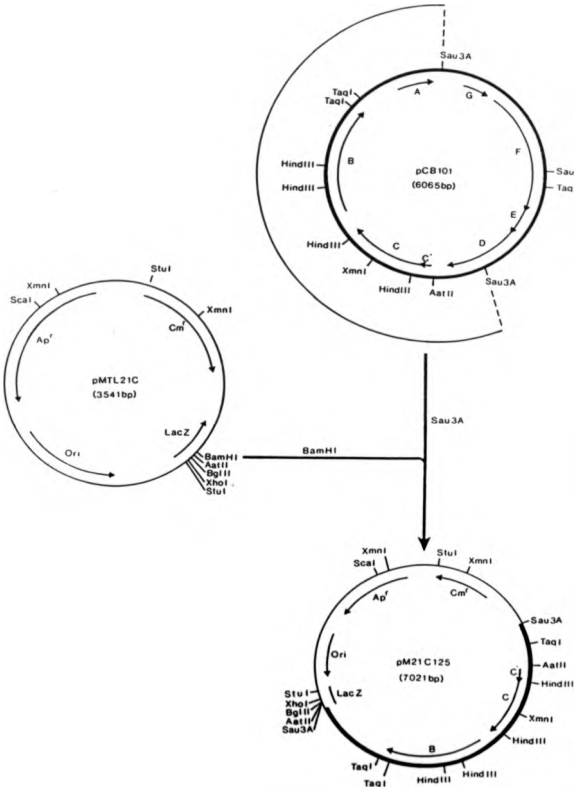


FIGURE 3p DERIVATION OF pM21C126

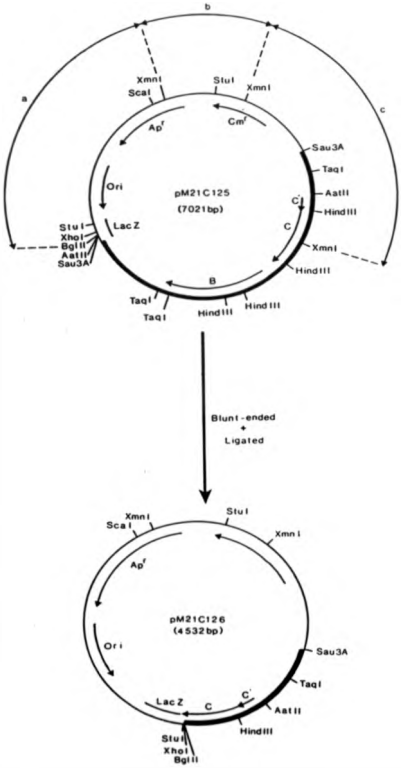
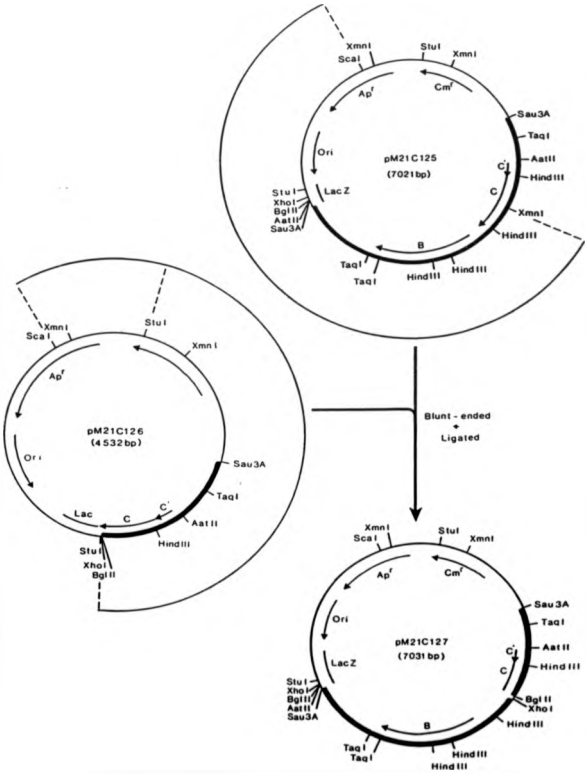


FIGURE 3q DERIVATION OF pM21C127



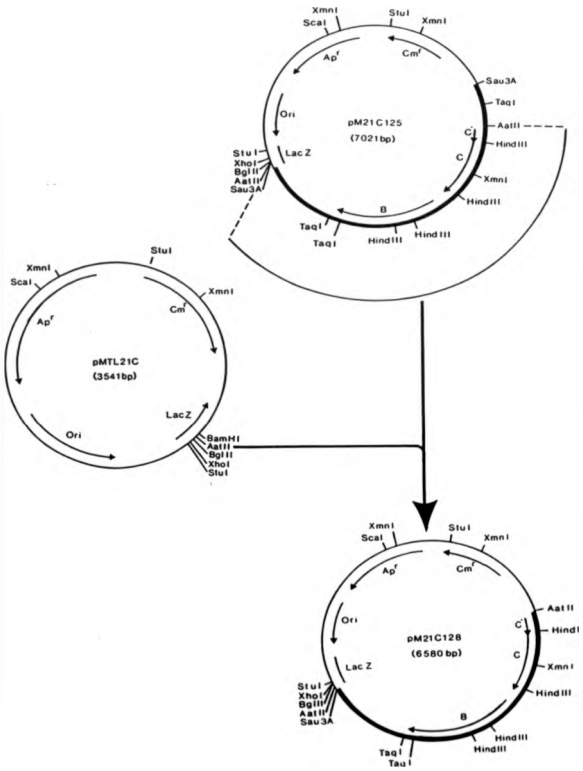
Equal aliquots (5 $\mu$g) of both pM21C125 and pM21C127 were used to transform naturally competent B. subtilis 168 cells which were subsequently plated out onto L-agar supplemented with chloramphenicol and incubated for 48 hours at 37$^{o}$C. Transformants were only obtained with pM21C125 DNA thus suggesting that ORF C is also in some way essential for replication of pCB101.

3.2.6.3. CONSTRUCTION OF pM21C128

A pM21C125 derivative devoid of the tandem palindromic sequences lying immediately upstream of ORF C' (i.e., nt position 2974-3013 of the pCB101 map; Fig. 3f) was constructed to ascertain whether this sequence might be involved in replication of the plasmid. Such a deletion derivative was constructed by virtue of the unique AatII site lying within the sequence of the first palindrome (nt position 2927; Fig. 3f).

Approximately 50 $\mu$g pM21C125 DNA was digested with AatII and the products size-fractionated on a preparative agarose gel (1% w/v) The desired 3 kb AatII fragment, containing intact ORFs B, C, and C' without the adjacent palindromic sequence, was extracted from the gel by electroelution. This fragment was subsequently ligated with pMTL21C DNA which had been cleaved with AatII and dephosphorylated. Transformation of competent E. coli JM83 yielded many recombinants, six of which were chosen and small scale plasmid preparations derived from them. These preparations were digested with HindIII and the products size-fractionated through a 1% (w/v) agarose gel. One isolate, pM21C128, with the insert orientation as shown in Fig. 3r was chosen and employed for the large scale isolation for plasmid DNA.

FIGURE 3r DERIVATION OF pM21C128



Equal aliquots (5 $\mu$g) of pM21C128 and pM21C125 DNA preparations were used to transform naturally competent B. subtilis 168 cells and after 48 hours incubation at 37°C the only transformants obtained were from pM21C125. This therefore suggested that the palindromic sequences upstream of ORF C' are in some way essential to replication of pCB101.

### 3.2.6.4. "NICK-SITE" SEARCH ANALYSIS

As recent reports had demonstrated that several of the Gram-positive plasmids studied to date replicate via a single-strand intermediate in an analogous fashion to the bacteriophages of E. coli. the pCB101 sequence was analysed for both the ØX174 and M13 type "nick-sites" present in the plus ("+") strand origins of replication.

Using the programs of DNASTAR Inc. the pCB101 sequence was searched for the presence of suitably located sequences that might bear some resemblance to either the ØX174 type of "nick-site" (Langeveld et al.. 1978), also shared by pC194 and pUB110 (Gros et al. 1987), or the M13 type "nick-site" shared by pT181 (Koepsel & Khan, 1987), pC221, and pNS1 (this thesis). A striking homology was found between the pCB101 sequence at position 40212-4048 (Fig. 3f) and the "nick-site" sequence common to ØX174 (Langeveld et al.. 1978), pC194, and pUB110 (Gros et al.. 1987). This sequence is located some 50 bp upstream of the initiation codon for ORF B and includes the majority of the 12 bp repetitive and palindromic sequence found at the deletion point of pRB1 (3.2.3.2.). Homology between these sequences is particularly striking and is shown in Fig. 3s. Also shown in Fig. 3s is the amino acid sequence homologies of the replication proteins; this limited region of homology is thought to be that responsible for the recognition of

FIGURE 3a NICK-SITE SEQUENCE HOMOLOGIES

A. Nucleotide sequence homology in the vicinity of the nick-site ( ⇩ ) of pCB101 (this study), pC194 (Michel & Erlich, 1986a), pUB110 (Gros *et al.*, 1987), and ØX174 (Langeveld *et al.*, 1978). Identity is indicated by colons.

B. Amino acid sequence homology between the replication proteins of ØX174 (van Mansfeld *et al.*, 1986), pC194 (Horinouchi & Weisblum, 1982a), and pUB110 (McKenzie *et al.*, 1986), and pCB101 (this study). The tyrosine of the ØX174 replication protein, covalently linked to DNA, is indicated by the arrow.

A.

```
                          NICK  SITE
                                ⇩
5' - TTTTCTTTCT-ATCTTGATATATATA  -  pCB101
5' - :G::::::::T::::::::::C:::::  -  pUB110
5' - ::::::::::T::::::::::ATA:GG  -  pC194
5' - :G:G::CC:CC:A::::::::TA:::   -  ØX174
```

B.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Glu | Leu | Phe | Lys | Tyr | Met | Thr | Lys | Val | Thr | Gly | - pCB101 |
| Glu | Thr | Ala | Lys | Tyr | Pro | Val | Lys | Asp | Thr | Asp | - pUB110 |
| Glu | Met | Ala | Lys | Tyr | Ser | Gly | Lys | Asp | Ser | Asp | - pC194 |
| Glu | Leu | Ala | Lys | Tyr | Val | Asn | Lys | Lys | Ser | Asp | - ØX174 |

the "+" origin of replication and also contains the tyrosine residue of the ØX174 protein which is covalently attached to DNA upon nicking of the replication origin (van Mansfeld et al., 1986).

### 3.2.6.5. SEARCH FOR A MINUS ORIGIN OF REPLICATION

It has been reported for most of the Gram-positive plasmids so far studied that they possess a minus ("-") origin of replication analogous to those present in the single-stranded bacteriophages of E. coli (Gruss et al.. 1987). This element (pal A) is a sequence of extended hyphenated dyad symmetry capable of hairpin formation and also possesses a highly conserved sequence, $RS_b$ (Novick et al., 1984b). The pCB101 sequence was analysed for a similar element.

Using the programs of DNASTAR Inc. no sequence resembling the pal A sequences of pT181, pC221, pC194 (Gruss et al.. 1987) and pLS1 (del Solar et al.. 1987) was found.

### 3.2.6.6. ANALYSIS OF pCB101 FOR A SINGLE STRANDED REPLICATION INTERMEDIATE

Prompted by the finding of a ØX174 type of "nick-site" in the pCB101 sequence, the next logical step was to analyse log phase cultures of B. subtilis 168 carrying the pCB101 chimaeric plasmid construct, pM21C125 (3.2.6.2.), for the presence of a single-stranded replication intermediate.

Small scale crude lysates were prepared from exponential cells of B. subtilis harbouring pM21C125 (3.2.6.2.) and analysed for the presence of single-stranded plasmid DNA that would be present if pCB101 replicated via a rolling-

circle mechanism. This analysis revealed the presence of single stranded pM21C125 DNA in addition to both monomeric and multimeric double stranded forms of the plasmid (Fig. 3T).

### 3.2.7. ANALYSIS OF THE pCB101 ORF PRODUCTS

#### 3.2.7.1. ANALYSIS OF THE HYDROPATHIC CHARACTER OF THE ORF PRODUCTS

Using the computer program devised by Kyte and Doolittle (1982) the hydrophilicity and the hydrophobicity of all the putative pCB101 ORF products were evaluated. The results are displayed graphically in Fig. 3u which might suggest that they are indeed proteins (T. Atkinson, personal communication). In addition, similar plots for the replication proteins of other Gram-positive plasmids are compared with the plots for ORFs B and C' (Fig. 3v).

#### 3.2.7.2. AMINO ACID SEQUENCE COMPARISONS

Using the computer programs of DNASTAR Inc. the presumptive pCB101 encoded proteins were compared with all the translated ORFs of all the other Gram-positive plasmids sequenced at that time; namely, pCB102 (Minton, unpublished data), pIP404 (Garnier & Cole, 1988a), pC194 (Horinouchi & Weisblum, 1982a), pUB110 (McKenzie et al.. 1986, 1987), pFTB14 (Murai et al.. 1987), pRBH1 (Muller et al.. 1986), pT181 (Khan & Novick, 1983), pC221 (Projan et al.. 1985), pE194 (Horinouchi & Weisblum, 1982b; Villafane et al.. 1987), pLS1 (Lacks et al.. 1986), pSN2 (Khan & Novick, 1983), pNE131 (Lampson & Parisi, 1986), pIM13 (Monod et al.. 1986), and pNS1 (Noguchi et al.. 1986). In addition, the presumptive proteins were compared with all the protein sequences available in the PIR database of DNASTAR Inc.

FIGURE 3t DETECTION OF SINGLE-STRANDED pCB101 DNA

Autoradiogram derived from agarose gel of lysates prepared from *B. subtilis* containing pM21C125. Samples were treated (lanes A and C) or not treated (lanes B and D) with nuclease S1. The DNA was denatured (lanes A and B) or not (lanes C and D) prior to transfer to nitrocellulose and was then hybridised with $^{32}$P-labelled pM21C125.
Supercoiled monomeric plasmid DNA (SC) and single-stranded DNA (SS) are indicated.

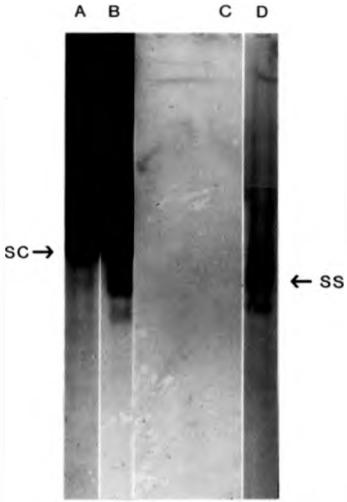


FIGURE 3u HYDROPATHIC CHARACTER OF THE PUTATIVE ORF PRODUCTS OF pCB101

Hydropathy scale indicates relative hydropathy as determined by Kyte and Doolittle (1982). Positive values represent hydrophilicity. Negative values represent hydrophobicity.

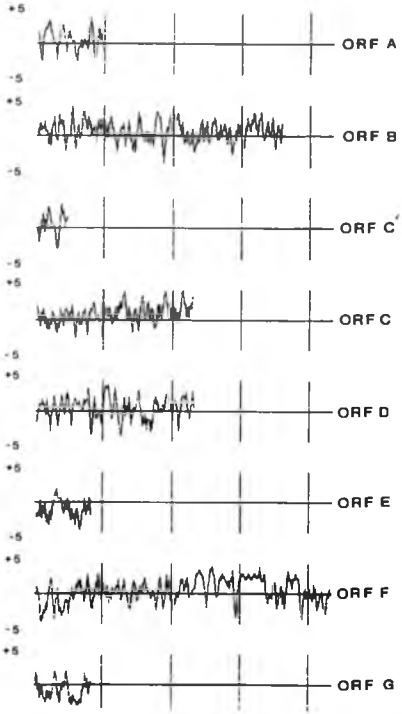
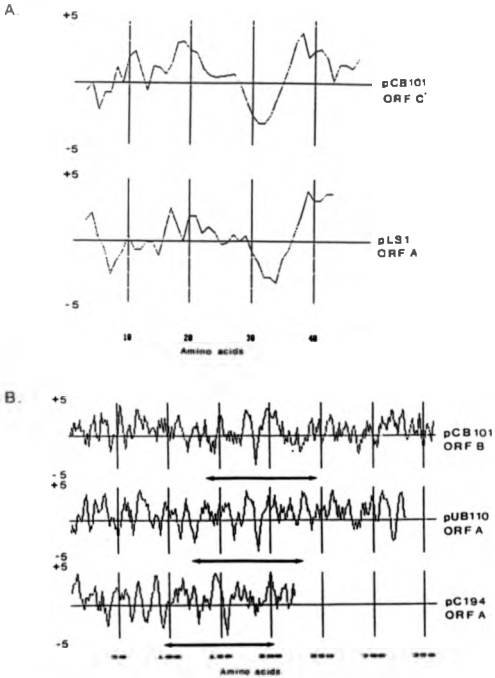


FIGURE 3v HYDROPATHY PLOT COMPARISONS

A. Hydropathy plots of pCB101 ORF C' (this study) and pLS1 ORF A (Lacks *et al.*, 1986).

B. Hydropathy plots of pCB101 ORF B (this study), pUB110 ORF A (McKenzie *et al.*, 1986), and pC194 ORF A (Horinouchi and Weisblum, 1982a). Regions of similarity are indicated by solid arrows.



No significant homologies were discovered except in the cases of ORF B and ORF C'. ORF B exhibits a 30% overall homology with the replication proteins of pC194 (ORF A), and pUB110 (ORF A). As can be seen from the alignment of these amino acid sequences in Fig. 3w, the three proteins exhibit 3 significant regions of homology (aa positions: 116-122, 176-183, and 263-273, Fig. 3w). Another interesting feature is the conservation of three cysteine residues in all three proteins (aa positions: 60, 83, and 91, Fig. 3w). This finding therefore adds weight to the hypothesis that the ORF B protein is involved in plasmid replication.

ORF C' shows homology with the similar sized repA ORF (ORF A) of pLS1 (Lacks et al., 1986) and is shown in Fig. 3x. These two small peptides also possess very similar hydrophobicity plots (3v).

### 3.2.7.3. IN VITRO TRANSCRIPTION AND TRANSLATION OF THE pCB101 ORFS

Equal quantities (2.5 μg) of pAT153, pCB2 (3.2.4.1.), pMTL21C, pM21C125, pM21C127 (3.2.6.2.), and pM21C128 (3.2.6.3.) were extracted once with phenol and the purified DNA preparations resuspended in 12.5 μl volumes of 1 X TE buffer. These samples were used directly in in vitro transcription and translation reactions. The reaction products were subsequently subjected to SDS polyacrylamide electrophoresis through a 10-30 % gradient gel, which was dried under vacuum prior to autoradiography.

As seen in Fig. 3y, autoradiography revealed the presence of three protein bands for pM21C125 which are not present with pMTL21C. The molecular weight of these proteins was estimated as 43,000, 26,000, and 34,000 Da by comparison of

FIGURE 3w AMINO ACID ALIGNMENT OF pCB101, pC194, AND pUB110 REPLICATION PROTEINS

Identities are indicated as follows:

* - identical for all three proteins

embolden - identical for 2 proteins

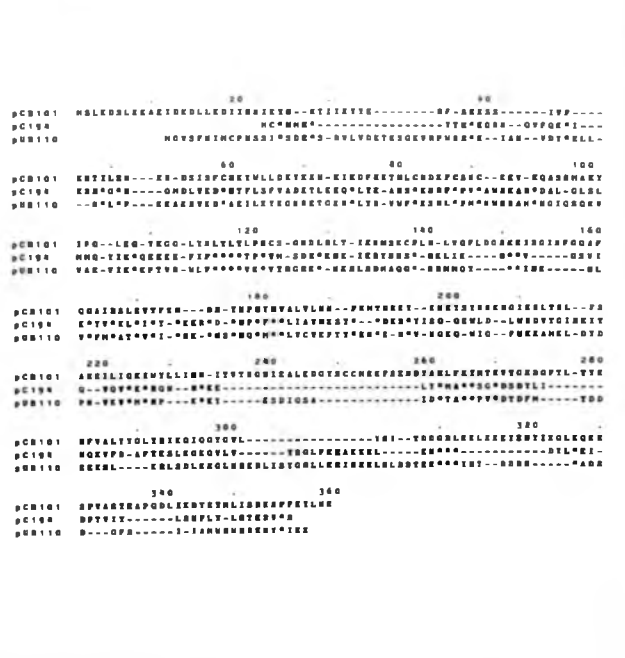


FIGURE 3x AMINO ACID ALIGNMENT OF pCB101 ORF C' AND pLSI ORF A

Identical amino acids are indicated. Chemically similar amino acids are indicated by :.

```
pCB101 ORF C'    MRVNISIPDEVKQFFEDYSKKTGVPQSSLMALALSEYKDKIERSLSNDK
                  R  I:: : V : :E   :  G:  S    :AL  YK   E:
pLS1 ORF A      MKKRLTITLSESVLENLEKMAREMGLSKSAMISVALENYKKGQEK
```

their electrophoretic mobilities with the protein size markers. The weak band migrating at about 43,000 Da agrees well with the calculated value of 42,910 Da for ORF B (Table 3.1). Similarly the weak band with an estimated molecular weight of around 26,000 Da correlates well with the calculated value of 27,134 Da for ORF C (Table 3.1). The strong band corresponding to a protein with an estimated molecular weight of 33,000 Da correlates well to that of a protein arising from the translation of a polycistronic mRNA comprising ORFs C' and C. This might seem feasible as the two ORFs overlap by some 11 bp (Fig. 3f).

The gel pattern for pM21C127 agrees well with that expected for premature translational termination of ORF C, with the ORF B product unchanged. Early termination at the point of the frame-shift mutation in ORF C (Fig. 3k) would thus give rise to proteins with calculated molecular weights of 16,000 and 20,000 Da for the monocistronic and bicistronic ORF C mRNAs respectively. The observed presence of protein bands with approximate molecular weights of 15,500 and 21,000 Da, with the concomitant loss of the 26,000 and 33,000 Da proteins observed with pM21C125, supports this hypothesis.

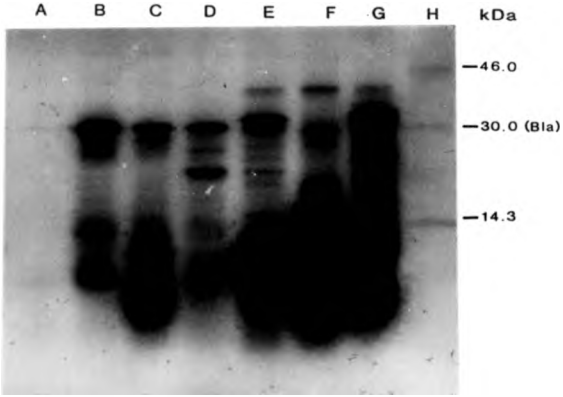
The gel pattern for pM21C128 is similar to that for pM21C125 but different in two respects. Firstly, the band corresponding to the ORF C'/C polypeptide is present in reduced quantity suggesting that the removal of the upstream palindromic sequence has had a deleterious effect on the expression of this product. Secondly, inexplicably however, this construct exhibits altered electrophoretic mobility and intensity of the small (c. 5,000 and 10,000 Da) vector derived polypeptides (Fig. 3y).

FIGURE 3y PROTEIN SYNTHESIS IN AN E. COLI BASED IN VITRO TRANSCRIPTION/TRANSLATION SYSTEM

Proteins labelled with [$^{35}$S] methionine were separated by electrophoresis on an SDS-Polyacrylamide gradient gel (10-30%) and revealed by autoradiography.

Lane A No DNA Control
B pCB2 DNA
C pMTL21C DNA
D pAT153 DNA
E pM21C128 DNA
F pM21C127 DNA
G pM21C125 DNA
H Protein size markers

The products of the beta-lactamase gene carried by both vectors pAT153 and pATL21C are indicated, as are the molecular weight markers.



No protein bands were seen with pCB2 that were not present in pAT153 therefore suggesting that none of the other ORFs of pCB101 (i.e., A,D,E,F, & G) are expressed in E. coli.

### 3.2.7.4. ANALYSIS OF THE pCB101 ORFS FOR DNA-BINDING REGIONS

Using the programs of DNASTAR Inc. based on the alogorythms derived by Robson (1974) and Chou & Fasman (1978), the ORFs of pCB101 were analysed for the presence of Helix-Turn-Helix motifs common to proteins known to interact with DNA (Pabo & Sauer, 1984). In addition, alignments were made with the DNA-binding regions of a number of known DNA-binding proteins. These being the proteins encoded by Lambda cI (Sauer & Anderegg, 1978), Lambda cro (Hsiang et al. 1977), trpR (Gunsalus & Yanofsky, 1980), 434 cI (Grosschedl & Schwartz, 1979), lex A (Horii et al.. 1981), spoIID (Lopez-Diaz et al., 1986), gerE (Cutting & Mandelstam, 1986), and spoIIIC (Errington personal communication, cited Holland et al.. 1987).

The only putative DNA-binding motif identified was the sequence QNIEALKDGYSCCMEKFSEN occurring at position 240-260 (Fig. 3w) of the ORF B amino acid sequence. An alignment of this sequence with the other DNA-binding motifs is illustrated in Fig. 3z. This observation may be significant as the location of this region is immediately adjacent to the amino acid sequence proposed to interact with the identified "nick-site" (Fig. 3s).

FIGURE 3z

COMPARISON OF THE PUTATIVE DNA BINDING REGION OF ORF B

WITH DNA BINDING REGIONS OF OTHER REPORTED DNA BINDING PROTEINS

Homologous amino acid (aa) residues are indicated by underlining.

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Q | T | K | T | A | K | D | L | G | V | Y | Q | S | A | I | N | K | A | I | H | λcro |
| Q | E | S | V | A | D | K | M | G | M | G | Q | S | G | V | G | A | L | F | N | λcI |
| Q | R | E | L | K | N | E | L | G | P | G | I | A | T | I | T | R | G | S | N | trpR |
| Q | A | E | L | A | Q | K | V | G | T | T | Q | Q | S | I | E | Q | L | E | N | λ434cI |
| R | A | E | I | A | Q | R | L | G | F | R | S | P | N | A | A | E | E | H | L | lexA |
| T | R | G | F | G | H | G | V | G | M | S | Q | Y | G | A | N | F | M | A | K | spoIID |
| I | S | N | A | M | Q | K | L | G | V | K | G | R | S | Q | A | V | V | E | L | gerF |
| Q | R | E | I | A | K | E | L | G | I | S | R | S | Y | V | S | R | I | E | K | spoIIIC |
| Q | N | I | E | A | L | K | D | G | Y | S | C | C | M | E | K | F | S | E | N | pCB101 ORF B |

(aa 240-260)

← HELIX → . ← TURN → . ← HELIX →

## 3.3. DISCUSSION

The results obtained in this study indicate that the replication machinery of the *C. butyricum* plasmid pCB101 reside within a 2.5 kb region of DNA between nucleotide 2764 and 5267 (Fig. 3f). A number of observations suggest that pCB101, like some other plasmids of Gram-positive origin (Gros *et al.*. 1987; Koepsel & Khan, 1987; Puyet *et al.*. 1988), replicates via an asymmetric rolling circle mechanism analogous to the isometric and filamentous bacteriophages of *E. coli* (Baas, 1985).

The first piece of evidence for the above hypothesis is the observed amino acid sequence homology between the protein encoded by ORF B and the replication proteins of pC194 and pUB110 (3.2.7.2.). This suggests that the ORF B translation product is functionally equivalent to a replication protein with topoisomerase (nicking-closing) activities (Drlica, 1984; Baas, 1985). This is further supported by two other findings. Firstly, that, the C-terminal region of ORF B possesses a typical DNA-binding motif (Pabo & Sauer, 1984) juxtaposed to the region of homology with ØX174 (van Mansfeld *et al.*. 1986) and pC194 (Gros *et al.*. 1987) containing the tyrosine residue which is believed to be covalently attached to the 5'-end of the leading strand upon nicking of the replication origin. Secondly, there is an apparent similarity in the hydrophilicity profiles of some 100 amino acids comprising the central regions of the replication proteins of pC194, pUB110, and the pCB101 ORF B product (Fig. 3v).

A second compelling piece of evidence supporting the proposal that pCB101 replicates via a rolling circle mechanism is the presence of the putative "nick-site" 5' to ORF B.

plus origins of the pC194/pUB110 family of plasmids. In the case of pUB110, 23 of 26 nucleotides are identical (Fig. 3a). Although no experiments were undertaken to confirm that the origin of replication resides within this region, indirect evidence was obtained supporting this contention.

Initial analysis of pRBΔ1-4 (3.2.3.2.) did not elucidate the mechanism responsible for the specific deletion, other than the proposed involvement of the repetitive clostridial sequence (5'-ATATATATATAT-3') and possibly some lambda derived vector sequence. Once the presence of the putative nick site was established it became apparent that the exact point of deletion in all four plasmids was the same position cleaved by the replication proteins of the E. coli bacteriophage ØX174 (Langeveld et al.. 1978), pC194 (Michel & Ehrlich, 1986a), and pUB110 (Gros et al., 1987). Studies with hybrid plasmids between pC194, pBR322, and bacteriophages f1 and M13, which gave rise to deletion variants in E. coli. led Michel & Ehrlich (1986a, 1986b) to propose a model of illegitimate recombination for the mechanism responsible. In this model, deletion formation results from the erroneous termination of replication occurring at sequences resembling the correct termination signal which is related to those which flank the "nick-site". Thus in the case of pRB1Δ1-4, a similar recombination event, occurring between the "nick" site sequence of pCB101 and some related sequence within the Lambda DNA segment, could be responsible for the specific deletions encountered. This hypothesis is supported by the observed structural stability of pCB1 (Fig. 3c) in E. coli where no Lambda derived sequences are present.

The proposed rolling-circle mode of replication of pCB101 is further established by the demonstration of single-stranded DNA, believed to be a replication intermediate

stranded DNA, believed to be a replication intermediate (3.2.6.6.) in an analogous way to that demonstrated for other Gram-positive replicons (te Riele et al., 1986a, 1986b). However, not all Gram-positive plasmids replicate in this fashion; examples include the C. perfringens plasmid pIP404 (Garnier & Cole, 1988b) and vectors based on the pAMβ1 replicon (N. Minton, personal communication). Consistent with the proposed theory, the pCB101 sequence was searched for a minus origin of replication (palA). Although no sequence exhibiting homology with the palA sequences characterised to date (Baas, 1985; Gruss et al., 1987; del Solar et al., 1987) was identified, a putative sequence that could be proposed is that which is responsible for "pile-ups" in the nucleotide sequence reactions (3.2.6.5.). However, it is more likely that this sequence might be involved in the control of replication as it lies within the N-terminal portion of ORF B. An interesting observation arising from the continued culture of pM21C125 in B. subtilis, during the course of these experiments, was the visualisation of this plasmid upon staining of agarose gels with ethidium bromide (3.2.6.6.). Collins et al. (1985) reported that pCB101 chimaeric plasmids were only detectable in B. subtilis by hybridisation experiments, reflecting it's low copy number in this host. However, a similar observation has been made with other pCB101 chimaeric constructs in B. subtilis and shown to be the result of a chromosomal mutation in the host (W.L. Staudenbauer, personal communication). It therefore seems likely that a similar host mutant has been selected accidentally in this study.

Although experiments to identify any regulatory RNA countertranscripts were beyond the scope of this study, it would seem feasible that replication control and incompatibility of pCB101 might be effected by short countertran-

for pT181 (Novick, 1984a), pC221 (Projan et al.. 1985), pC194 (Alonso & Tailor, 1987), pE194 (Villafane et al.. 1987), and pLS1 (del Solar et al.. to be published, cited Puyet et al.. 1988). It is perhaps worth speculating that the palindromic sequence lying within the N-terminal portion of ORF B (attributable to sequence reaction "pile-ups") could possibly be responsible for some kind of RNA-RNA interaction involved in control of replication by affecting the transcription of ORF B. The pair of tandem palindromes lying immediately upstream of ORF C' might also be implicated in the control of replication as disruption of this sequence (pM21C128) results in the loss of replication of this construct in B. subtilis (3.2.6.3.).

The minimum requirement for replication in plasmids pC194 and pUB110 is the presence of a plus origin and a functional replication protein. This does not appear to be the case with pCB101. In the first instance, the introduction of a frame-shift into ORF C destroys the replicative ability of pCB101. This mutation does not appear to effect the expression of ORF B, as a polypeptide corresponding to the ORF B protein is still produced during in vitro transcription/translation assays. The simplest explanation of these results is that the ORF C polypeptide is also required.

The similar hydrophilicity profiles and amino acid homology between ORF C'(45 aa) and pLS1 Rep A protein (49 aa) (Fig. 3v, 3x) suggests that the two small encoded polypeptides might be functionally similar. It has been reported for pLS1 that Rep A (ORF A) protein is translated from a polycistronic mRNA together with the replication protein Rep B (ORF B), and that the Rep A protein exhibits a strong affinity for DNA (Puyet et al., 1988). These authors further suggested that Rep A is involved in the control of

replication. It appears conceivable that the pCB101 ORFs C' and C might also be transcribed as one polycistronic unit as in vitro transcription/translation products have been obtained which exhibit electrophoretic mobilities which would be expected from the translation of ORF C' and ORF C together, in addition to translation of ORF C alone (3.2.7.3.). This is supported by the finding of appropriately sized polypeptides arising from the in vitro transcription/translation of pM21C127 which possesses a frame-shift mutation in ORF C (3.2.6.2.; 3.2.7.3.). However, unlike the situation in pLS1, in pCB101 the ORF C'/ORF C 11 bp overlap terminates in a translational stop codon in the ORF C' reading frame. Therefore, to account for the presence of ORF C'/ORF C sized protein observed (3.2.7.3.), either there is a mistake in the nucleotide sequence or a ribosomal frame-shifting event occurs. The latter appears most likely as this region was sequenced more than once on each strand. Such frame-shift events have been reported in the translation of RNA phage genomes (Atkins et al.. 1979; Kastelein et al.. 1982) and yeast mitochondrial genes (Fox & Weisss-Brummer, 1980). The lack of finding of any amino acid sequence homology between ORF C and any other ORF from any other Gram-positive plasmid (3.2.7.2.) suggests that it may be uniquely involved in the regulation of replication of pCB101.

The deletion problems encountered with the M13mp9/pCB101 minimal replicon recombinant (3.2.6.1.) appeared to be one of complete structural instability, independent of any host recombination functions as the problem was not resolved in a E. coli TG2 host. Similar observations have been made when the introduced DNA contained a strong promoter or other feature deleterious to it's maintenance in E. coli (N. Minton, personal communication). However, not much is known about the strength of clostridial promoters per se.

known about the strength of clostridial promoters per se. but ORFs B, C', and C, contained within the minimal replicon, do possess sequences upstream of their initiation codons which resemble typical E. coli and vegetative B. subtilis promoters (Moran et al., 1982) (Table 3.5), and that these ORFs are expressed in an E. coli based in vitro transcription/translation system (3.2.7.3.). Alternatively, and perhaps most likely, this observed instability might have arisen as a direct consequence of the presence of two replicons in the same chimaera, both of which replicate via a similar rolling circle type mechanism.

One of the most striking features of the pCB101 sequence, and clostridial genes in general, is the remarkably high dA + dT content of 73%. From a statistical analysis of the relationship between DNA base composition and protein content in microorganisms, Elton (1983) concluded that such a value is close to the theoretical limit for a protein-coding sequence based on the conventional genetic code. Also, the dG + dC content of the eight ORFS of pCB101 is not significantly higher than the value for the non-coding regions of the plasmid, as reported for the C.perfringens bacteriocinogenic plasmid pIP404 (Garnier & Cole, 1988a).

Of the 8 ORFs encoded by pCB101 only those involved in replication of the plasmid appear to be expressed in an E. coli based in vitro transcription/translation system (3.2.7.3.). However, pCB101 is incapable of replication in E. coli (Collins et al., 1985). From the analysis of codon usage (Table 3.3) carried out in this study it would appear that the lack of heterologous gene expression of the other ORFs is most likely due to their extreme base composition; similarly, only one of the 10 ORFs of pIP404 has been reported to be expressed in E. coli (Garnier & Cole, 1988a) and this plasmid exhibits an amino acid utilisation profile

the remaining ORFs of pCB101 is unlikely to be due to poor transcription as all these ORFs possess recognisable promoter sequences upstream of their respective initiation codons (Table 3.5), and E. coli exhibits great promiscuity in its ability to recognise Gram-positive promoter sequences (Moran et al.. 1982). Further, heterologous translational signals are generally recognised in E. coli and do not constitute a barrier to gene expression (Lideman, 1983). Rather, it is proposed here that the extreme codon usage displayed by the ORFs of pCB101 represents a a major translational obstacle because many of the preferred clostridial codons are rarely, if ever, used in E. coli (Ikemura, 1981; Grosjean & Fiers, 1982). On comparison of the frequency of codon utilisation in pCB101 and E. coli a difference of up to 80-fold can be seen for the following preferred codons: Arg, AGA v CGU; Leu, UUA v UUG; Thr, ACA v ACC; Pro, CCA v CCG; Gly, GGA v GGC; Ile, AUA v AUC (Table 3.3).

From the extensive amino acid sequence comparisons carried out (3.2.7.2.) no homologies were identified for any of the pCB101 ORFs other than ORF B and ORF C' thus disenabling the assignment of any functions to any of the remaining putative plasmid encoded products. Of particular interest was the search for a sequence that might encode the 32.5 kd bacteriocin known to be produced by the native host of pCB101.C. butyricum NCIB 7423 (Clarke & Morris, 1976). The only candidate of pCB101 that could possibly encode a protein of this approximate size is ORF D (Table 3.1). However, apart from the complete lack of any nucleotide sequence homology with the bacteriocin encoding ORF of pIP404 (Garnier & Cole, 1986) this putative ORF product does not possess a hydrophobic C-terminus nor a high glycine content which are both characteristics of bacteriocins (Pugsley, 1984; Garnier & Cole, 1986). However, during the

(Pugsley, 1984; Garnier & Cole, 1986). However, during the course of this study the nucleotide sequence of the co-resident plasmid, pCB102 (8.3 kb), has been determined and reported to possess a suitable candidate for the bacteriocin gene (Minton, unpublished data).

In summary, the complete nucleotide sequence of the cryptic clostridial plasmid, pCB101, has been determined and it's minimal replicon characterised. Furthermore, evidence has been supplied which strongly suggests that pCB101 replicates via an asymmetric rolling circle type mechanism similar to some other plasmids of Gram-positive origin and the filamentous and isometric bacteriophages of *Escherichia coli*. Of the eight ORFs identified, ORFs B, C, and most likely C', appear to be essential for replication. The remaining five ORFs, however, remain cryptic. The information gained from these studies will prove valuable to the future *in vitro* construction of *Clostridium/Escherichia coli* shuttle vectors based on the pCB101 minimal replicon.

# CHAPTER FOUR

## THE CONSTRUCTION AND USE OF AN EXPRESSION CARTRIDGE BASED ON THE FERREDOXIN GENE OF *CLOSTRIDIUM PASTEURIANUM*

## 4.1. INTRODUCTION

A particularly useful refinement to a *Clostridium/E. coli* shuttle vector would be the provision of a strong clostridial promoter which could be exploited to elicit the high expression of cloned genes. One popular strategy employed for the isolation of promoter sequences has been to use "promoter-probe" plasmid cloning vehicles. Such vectors commonly possess a structural gene lacking the transcriptional signals necessary for expression. Insertion of a DNA fragment containing suitable regulatory sequences 5' to the promoterless gene results in the expression of the gene product. The genes utilised are such that production of their encoded proteins confers on the host an easily detectable phenotype. Many such vectors have been constructed for *E. coli* (An & Friesen, 1979; Brosius, 1984; Casadaban & Cohen, 1980; West & Rodriguez, 1982), *B. subtilis* (Band *et al*. 1983; Donnelly & Sonenshein, 1984; Yoshimura, 1984; Zukowski *et al*., 1983), and even *Streptococcus sanguis* (Achen *et al*., 1986). However, in the absence of a clostridial transformation procedure, to try to similarly isolate a clostridial promoter would have necessitated the experiments being carried out in an alternative bacterial host. As very little was known about the utilisation of clostridial promoter sequences in other bacteria, an alternative approach was adopted. This was to isolate a promoter from a clostridial gene which was known to be highly expressed in its native environment.

At the time of this study a number of clostridial genes had been cloned (Ishii *et al*., 1983; Karube *et al*., 1983; Daldal & Applebaum, 1985; Beguin *et al*., 1985; Graves *et al*., 1985) but sequence data was only available for two of them. These were the genes encoding the extracellular endoglucanase A (cel A) from *C. thermocellum* (Beguin *et al*., 1985) and the ferredoxin gene (Fd) from *C. pasteurianum* (Graves *et al*., 1985). The Fd gene was chosen for two reasons. Firstly, *C. pasteurianum* is more closely related to the intended host, *C. acetobutylicum*, with similar mol G+C (28 %) as compared to *C. thermocellum* (50%). Secondly,

it had previously been reported that clostridia grown in iron-sufficient media produce yields of the Fd protein representing up to 2% of the total cell protein (Rabinowitz, 1972, cited Graves et al. 1985).

The strategy chosen was to re-clone the Fd gene using an oligonucleotide probe based on the published DNA sequence (Graves et al.. 1985) and utilise the Fd transcriptional control sequences in the construction of an expression cartridge. The proposed cartridge was designed to be "portable", facilitating its incorporation in future plasmid vectors. This cartridge was initially to be evaluated by its ability to direct the expression of a suitable gene in E. coli and B. subtilis. Subsequent analysis in a clostridial host was dependent on the development of a gene transfer system.

As no readily assayable genes of clostridial origin were available, the pseudomonad catechol 2,3-oxygenase gene (xylE) was chosen for the expression studies. This gene seemed a suitable candidate for a number of reasons: (1) it is promoterless, being part of the lower pathway operon on the TOL plasmid for the degradation of aromatic compounds (Franklin et al.. 1981), (2) it possesses a sequence 9 bases upstream of the initiation codon which exhibits significant complementarity to the 3' region of the 16S rRNA of B. subtilis (Moran et al.. 1982; Zukowski et al.. 1983), (3) the xylE gene product is readily assayable as it catalyses the the meta-cleavage of catechol to give the bright yellow coloured pigment 2-hydroxymuconic semialdehyde (Sala-Trepat & Evans, 1971), and (4) it has been reported to be efficiently expressed in both Gram-negative and Gram-positive bacteria when appropriate transcriptional initiation sequences are provided (Zukowski et al.. 1983).

## 4.2. RESULTS

### 4.2.1. CLONING OF THE FERREDOXIN GENE FROM C. pasteurianum

#### 4.2.1.1. ISOLATION OF THE Fd-ENCODING GENOMIC DNA

Genomic DNA was isolated from C. pasteurianum ATCC 6013 and digested to completion with Sau3A. The restricted DNA was size fractionated on a preparative agarose gel (1 % w/v), and 6 gel slices excised, which spanned the 0.5 to 0.7 kb size range with respect to DNA size markers. The DNA electroeluted from these 6 gel fragments was subjected to Southern blot analysis using a $^{32}$P-dATP end-labelled 19 base oligonucleotide probe (5'-AGTTGGAGCACCAGTACAA-3') complementary to the non-coding strand of the published Fd sequence (nucleotide positions 315 to 333, Graves et al., 1985). The hybridisation was carried out at 51$^{o}$C for a period of 2 hours which represents a temperature of Tm-5$^{o}$C and a time duration equivalent to 3 X $CoT_{1/2}$ (Wallace et al., 1981). Subsequent autoradiography revealed that fraction 3 containing the desired 0.6 kb Sau3A fragment DNA.

#### 4.2.1.2. ATTEMPTED CLONING OF THE Fd GENE DIRECTLY INTO M13

Initially it was intended to shot-gun clone the Sau3A fraction DNA (4.2.1.1.) directly into the replicative form (RF) of the E. coli phage vector M13mp7, thereby facilitating DNA sequencing and subsequent site-directed mutagenesis. M13mp7 RF DNA was digested with BamH1. dephosphorylated, and ligated with the isolated DNA. The ligation products were used to transfect competent E. coli JM101. Recombinant phage-infected colonies, derived from 600 white plaques, were subjected to in situ colony hybridisation. The hybridisation conditions and oligonucleotide were the same as employed in the previous southern blot experiment (4.2.1.1.). Autoradiography, however, did not reveal hybridisation-positive clones. Even though a total sum of 600 clones screened should have, in all probability, contained at least one copy of the Fd gene , the possibility that the desired clone was simply "missed" could not be excluded.

In consequence, a further 1400 recombinant phage plaques were similarly screened. Again the results were negative suggesting that the Fd encoding 0.6 kb Sau3A fragment was in someway "unclonable" directly into M13mp7.

4.2.1.3. CLONING OF THE Fd GENE IN pAT153

As an alternative approach it was decided to shot-gun clone the Fd gene into the plasmid cloning vector pAT153. Two different routes were initiated. In the one experiment recombinant plasmids were generated by inserting the previously isolated 0.6 kb Sau3A fragment, shown to include the Fd gene, into pAT153, while in a seperate experiment a gene library was constructed.

Plasmid pAT153 DNA was cleaved with BamHI. dephosphorylated, ligated with the 0.6 kb Sau3A fragments, and the ligation products used to transform competent E. coli W5445. Recombinants were detected by replica plating $Ap^r$ transformants onto L-agar supplemented with Tc. Of the 3000 $Ap^r$ colonies picked, 2300 were also $Tc^s$ owing to the insertion of foreign DNA into the BamHI site of pAT153. Of these recombinants, 2000 were screened for the presence of the Fd gene by in situ colony hybridisation. Autoradiography revealed two positively hybridising clones, FdI and FdII.

Simultaneously a gene library was constructed. Approximately 500 ug of the C. pasteurianum ATCC 6013 genomic DNA was partially digested with Sau3A and the digestion products size fractionated on a preparative agarose gel (0.8% w/v). DNA representing a size range of 7 to 10 kb was excised from the gel DNA extracted by electroelution. The isolated DNA was subsequently ligated with the cloning vector pAT153 which had previously been cleaved with BamHI and dephosphorylated. The ligation products were used to transform competent E. coli W5445. Of the 5000 $Ap^r$ transformants obtained 70 % were $Tc^s$, representing a gene bank of some 3500 clones. Small scale plasmid preparations from 48 randomly chosen recombinants yielded an average insert size of 8.0 kb as deduced from agarose gel electrophoresis of

the products of digestion with HindIII. In situ colony hybridisation was performed on the 3500 clones, revealing a single colony, FdIII, which hybridised strongly to the Fd specific oligonucleotide.

These three clones were employed for the large scale isolation of plasmid DNA. Approximately 1 $\mu$g amounts of pFdI, pFdII , and pFdIII plasmid DNA was digested with Sau3A and the restricted products size fractionated by agarose gel electrophoresis (1% w/v) alongside $^{32}$P end-labelled Lambda/HindIII DNA size markers and blotted to nylon membrane. Subsequent hybridisation with $^{32}$P end-labelled Fd specific oligonucleotide probe revealed, in each case, a strongly hybridising band with an apparent electrophoretic mobility approximating to 0.6 kb with respect to the DNA size markers. Inexplicably, However, the band identified for pFdII appeared to exhibit a slightly faster electrophoretic mobility than that of pFdI and pFdIII (Fig. 4a).

#### 4.2.1.4. DNA SEQUENCE DETERMINATION OF pFdI and pFdIII SAU3A INSERTS

Even though previous attempts to clone the Fd encoding Sau3A fragment DNA directly into M13mp7 had failed, it was decided to make one further attempt, this time with the gel purified Sau3A fragments of pFdI, pFdII, and pFdIII and a freshly prepared M13 cloning vector. The desired 0.6 kb Sau3A fragments were isolated from the three recombinant plasmids and ligated with M13mp7 RF DNA which had been cleaved with BamHI and dephosphorylated. The ligation products were used to transfect E. coli JM101. Thousands of recombinants were obtained with pFdI and pFdIII 0.6 kb Sau3A fragments in comparison with only 50 for the slightly smaller pFdII Sau3A fragment. Template DNA was prepared from 18 recombinant plaques from each of the three cloning experiments. Aliquots (2$\mu$l) of each of the 54 template DNAs were spotted onto NC filter discs and the single stranded template DNA baked onto the filters at 80$^{o}$C in a vacuum for 30 minutes. In an attempt to identify those templates which were "true" recombinants containing the

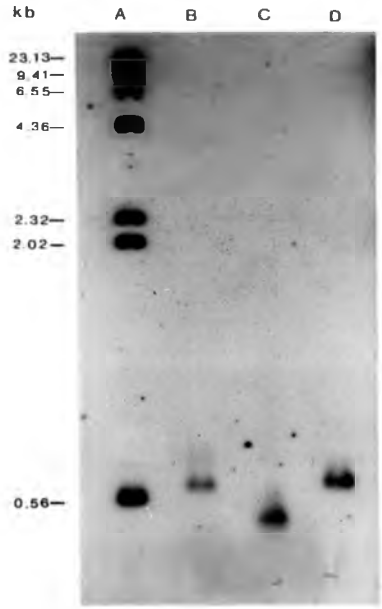
FIGURE 4a AUTORADIOGRAM OF PLASMID DNA DERIVED FROM CLONES FdI, II and III

Lane A λ/HindIII DNA size markers

B pFdI plasmid DNA digested with Sau3A

C pFdII plasmid DNA digested with Sau3a

D pFdIII plasmid DNA digested with Sau3A



non-coding strand of the Fd-encoding sequence and thereby facilitate DNA sequence analysis employing the Fd specific oligonucleotide as the sequencing primer (coding strand), the filter discs were hybridised with $^{32}P$ end-labelled Fd-specific oligonucleotide. As a positive control a colony streak of E. coli JM101 harbouring FdI was grown up overnight on a NC filter disc, processed as for in situ colony hybridisation, and included in the hybridisation. Subsequent autoradiography revealed no positive signals for any of the template DNA samples in the presence of a strong positive signal obtained with the FdI colony streak. This, therefore, implied that if any of the templates contained Fd derived sequence it was present as the coding strand only, suggesting that the Fd encoding fragment DNA was originally cloned in only the one orientation in M13. To verify this all 54 templates were analysed by DNA sequence analysis employing sequence universal primer and the Fd-specific oligonucleotide as primers. Sequence data derived with the sequence universal primer revealed that the majority of the FdI and FdIII derived templates contained the coding strand of Fd. With the FdII derived templates all the sequence data obtained was that of M13 indicating that the FdII Sau3A fragment was possibly unstable in M13mp7. As predicted, no sequence data was obtained when using the Fd specific oligonucleotide as primer.

At this stage it was decided to continue this study with FdI and FdIII clones only. In order to obtain the complete nucleotide sequences of pFdI and pFdIII Sau3A inserts a new oligonucleotide sequence primer was required, one that would enable the generation of sequence data beyond the gel reading limits of the data obtained with sequence universal primer. To achieve this, The reverse complement of the Fd specific oligonucleotide was synthesised (5'-TTGTACTGGTGCTCCAACT-3'). Six templates derived from FdI and FdIII were successfully re-sequenced with the new primer which, in combination with the sequence data already obtained with sequence universal primer, confirmed that both pFdI and pFdIII Sau3A fragment clones were the same giving an identical nucleotide sequence to that of the published

FIGURE 4b NUCLEOTIDE SEQUENCE OF THE FERREDOXIN GENE OF CLOSTRIDIUM PASTEURIANUM

Predicted amino acid sequence is given below the nucleotide sequence.

The initiation codon is underlined.

```
Sau3A
GATCGAGATAGTATATGATGCCATATTCTTTAAATATAGATAAAGTTATAGAAGCAATAGAAGATTTAGCATTTAC 75

TGTAATATAAATTACACTTTTAAAAAGTTTAAAAACATGATACAATAAGTTATGGTAAACTTATGATTAAAATTT 150

TAAGGAGGTGTATTTTTCATGGCATATAAAATCGCTGATTCATGTGTAAGCTGTGGCGCTTGTGCTTCAGAATGT 225
                     m  a  y  k  i  a  d  s  c  v  s  c  g  a  c  a  s  e  c
CCAGTTAATGCTATAAGTCAAGGAGATTCTATATTCGTTATAGATGCCGATACTTGTATCGACTGTGGTAACTGT 300
 p  v  n  a  i  s  q  g  d  s  i  f  v  i  d  a  d  t  c  i  d  c  g  n  c
GCTAACGTTTGTCCAGTTGGAGCACCAGTACAAGAATAATTTATAATTAAAACATAAAAATAAGAAGCCTGCAAA 375
 a  n  v  c  p  v  g  a  p  v  q  e  .
TGCAGGCTTCTTATTTTTATTGTAACTGAATTTTATTATGTTATACTATAAATTGAAGAAAATATTAGGGGGAAA 450

TTGGTATGCTTAATATATTTAGGGACACTTTTCAAGTTATGTCGCCGGTTAATGGAAATATAGTAAATTTAACTA 525

ATGTTCCAGACAGAATGTTTTCAGAGGAAATAGTAGGAAAAGGAATAGCTGTAGACCCATTAGAGGATATAATAA 600

GATC 604
Sau3A
```

Fd sequence (Graves et al.. 1985; Fig. 4b).

### 4.2.2. CONSTRUCTION OF THE FERREDOXIN EXPRESSION CARTRIDGE

The desired expression cartridge should possess two main features. Firstly, there should exist a reasonable number of unique restriction endonuclease sites between the transcriptional initiation and termination signals of the Fd gene, thereby introducing a fair degree of versatility for cloning into those sites. The insertion of heterologous genes into those sites would therefore render them under the control of the Fd transcriptional regulatory signals. Secondly the cartridge should be localised to a DNA fragment which may be generated by a commonly used restriction endonuclease. This latter feature would therefore ensure that the cartridge is "portable", thereby facilitating it's subsequent insertion into commonly used plasmid vectors.

It was envisaged that the cartridge would be rendered "portable" by utilising the vector M13mp7 for sub-cloning as, unlike more commonly used M13mp vectors, the polylinker is symmetric. Therefore, although the insertion of the Fd Sau3A fragment into M13mp7 would not result in the recreation of BamHI sites, the presence of flanking vector EcoRI sites would allow subsequent excision of the Sau3A fragment as a "portable" fragment. As the DNA between the Fd transcriptional control signals carries very few restriction endonuclease sites, it was decided to overcome this paucity by substituting part of the Fd coding sequence with the polylinker region of the versatile cloning vector, pMTL20 (Chambers et al.. 1988). Site-directed mutagenesis was the method chosen to facilitate this substitution.

#### 4.2.2.1. SITE-DIRECTED MUTAGENESIS OF THE FERREDOXIN GENE

Single-stranded template DNA derived from the M13mp7 clone, FdI, was chosen as substrate for site-directed mutagenesis for the creation of unique HindIII and SstI sites in the coding region of the Fd gene; the sites being promoter and

terminator proximal respectively. Two 21 base mutagenic oligonucleotides, Fdmut1 and Fdmut2, were synthesised and analysed for purity by denaturing polyacrylamide gel electrophoresis (20% w/v). Fdmut1 was designed to change the G at position 203 of the Fd sequence to a T, resulting in the creation of a unique HindIII site. In a similar fashion Fdmut2 was designed to change the A at position 323 to a T, resulting in the creation of a unique SstI site (Fig. 4C).

The mutagenesis experiments were carried out as described in methods (2.2.19.). The first mutagenesis employed Fdmut1 oligonucleotide with a calculated $T_{Dm}$ of 62°C in comparison to the $T_{Dwt}$ of 60°C. Successful mutants were identified as the screening washes approached 62°C (Fig. 4d). Four mutants identified were plaque purified. Three isolated plaques from each of the mutants were used for the preparation of template DNA. DNA sequence analysis using the "reverse strand" Fd specific oligonucleotide as primer revealed that all the templates sequenced were authentic mutants. One of the mutant templates was chosen as substrate for the second mutagenesis. The second mutagenic oligonucleotide, Fdmut2 , had a calculated $T_{Dm}$ of 64°C, with the $T_{Dwt}$ being 62°C. Successful mutants were identified at a screen wash of 63 °C (Fig 4d). Again 4 mutants were chosen and plaque purified. Template DNA was prepared from three plaques of each isolate and DNA sequencing undertaken using sequence universal primer. Two successful double mutants were chosen (FdM1 & FdM2) and used to prepare large scale RF DNA. These DNA preparations were subjected to restriction endonuclease analysis. Each contained the newly created single HindIII and SstI sites in addition to the flanking EcoRI sites derived from the symmetrical polylinker of M13mp7 (Fig. 4e).

#### 4.2.2.2. CONSTRUCTION OF THE CARTRIDGE

Fd double mutant RF DNA (approximately 100 μg) was digested with EcoRI and the restricted products size-fractionated by agarose gel electrophoresis (1.2% w/v). The portable 0.6 kb fragment (Fd) released from the M13 RF DNA was

FIGURE 4c

Fd MUTAGENIC OLIGONUCLEOTIDES

| Mutagenic oligonucleotide | Nucleotide position of Fd sequence complement (See Fig. 4b) |
|---|---|
| Fdmut1: 5'-TCAAGCGCCAAAGCTTACACA-3' (* above the second A) | 192 - 213 |
| Fdmut2: 5'-CTTGTACTGGAGCTCCAACTG-3' (* above the second A) | 315 - 334 |

* indicates single base pair mismatch

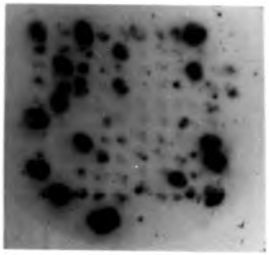
FIGURE 4d HYBRIDISATION SCREENING OF SUCCESSFUL MUTANTS

Autoradiograms of *in situ* colony hybridisations with Fdmut1 and Fdmut2 oligonucleotides.

Dark signals indicate hybridisation-positive colonies.

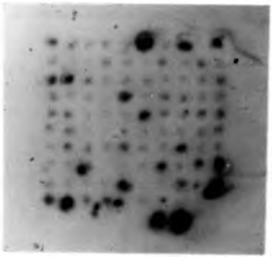
A. Fdmut1 mutagenesis

62°C wash



B. Fdmut2 mutagenesis

63°C wash



excised from the gel and the DNA extracted by electroelution. This fragment was subsequently ligated with pUC7 DNA which had been cleaved with EcoRI and dephosphorylated. The ligation products were used to transform competent E. coli JM83. Small-scale plasmid preparations from 12 presumptive recombinants were digested with EcoRI and several recombinants displaying the expected restriction patterns identified. One of these clones, carrying the desired plasmid chimaera, pUC7Fd, was selected and employed for the large scale isolation of plasmid DNA (Fig. 4f).

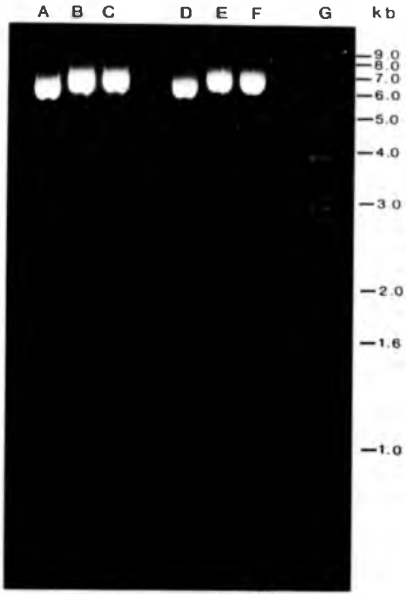
The next stage in the proposed construction was to replace the Fd coding sequence, between the newly created HindIII and SstI sites, with the polylinker of pMTL20, which is contained on a 78 bp HindIII-SstI fragment. In order to circumvent the problems associated with cloning such a small fragment in the absence of a detectable selection system it was decided to "mobilise" the polylinker as two halves flanking a readily selectable marker gene which had previously been cloned into the polylinker of pMTL20. The marker gene chosen was the "in house" pseudomonad carboxypeptidase $G_2$ ($CPG_2$) structural gene which resided on a 2.03 kb BamHI fragment in the vector pNM15 (Minton and Clarke, 1985). The presence of $CPG_2$ is readily detected as it confers, on the host, the ability to utilise folate as a sole carbon source; E. coli cells possessing a functional $CPG_2$ gene produce bright yellow colonies on minimal medium supplemented with folate.

The 2.03kb BamHI fragment carrying $CPG_2$ was isolated from pNM15 (30 $\mu$g) and ligated with pMTL20 which had been cleaved with BamHI and dephosphorylated. The ligation products were used to transform competent E. coli JM83 cells which were subsequently plated out onto minimal folate agar plates supplemented with Ap. Following incubation at 37$^{o}$C many small bright yellow colonies had appeared. Small scale plasmid isolations were prepared from six yellow colonies. All six plasmids carried a c. 2 kb BamHI fragment. A single clone carrying the desired recombinant plasmid was chosen, designated pMTL20/$CPG_2$, and

FIGURE 4e RESTRICTION ENZYME ANALYSIS OF Fd DOUBLE MUTANTS

Agarose gel electrophoresis (1% w/v) of Fd double mutant M13 RF preparations.

Lane A FdDM1 RF DNA digested with EcoRI
B FdDM1 RF DNA digested with HindIII
C FdDM1 RF DNA digested with SstI
D FdDM2 RF DNA digested with EcoRI
E FdDM2 RF DNA digested with HindIII
F FdDM2 RF DNA digested with SstI
G Kilobase ladder DNA size markers



employed for the large scale isolation of plasmid DNA (Fig. 4f). This prepared DNA (30 μg) was doubly digested with HindIII and SstI thus releasing CPG2 as a HindIII-SstI fragment flanked by the two halves of the pMTL20 polylinker. This fragment was ligated with pUC7/Fd DNA which had been appropriately cleaved with HindIII and SstI. The ligation products were used to transform E. coli JM83 and these cells were plated onto minimal folate agar plates containing Ap. Following incubation at 37°C, small bright yellow colonies had appeared. Six of these were chosen and ug quantities of plasmid DNA prepared and digested with HindIII and SstI both singly and jointly. The single digests confirmed, in each case, the presence of single HindIII and SstI sites, with the double digests releasing the expected 2.1 kb fragment insert. One of these clones, carrying the desired recombinant plasmid, pUC7/Fd/$CPG_2$, was selected and employed for the large scale isolation of plasmid DNA (Fig. 4f).

As the final step in the construction, the CPG2 BamHI fragment had to be deleted from pUC7/FdM3/$CPG_2$ to yield pUC7 carrying the polylinker derived from pMTL20 sandwiched between the transcriptional control signals of the Fd gene. This was achieved by transforming E. coli JM83 with self-ligated, BamHI cleaved pUC7/Fd/$CPG_2$ DNA and selecting for non-yellow colonies on minimal/folate agar supplemented with Ap. Four white colonies were picked and small scale plasmid preparations isolated from them. Subsequent restriction endonuclease analysis with BamHI, HindIII, SstI, and EcoRI confirmed the presence of single sites for BamHI, HindIII, and SstI with the flanking EcoRI sites releasing the Fd cartridge as a 573 bp EcoRI fragment. One of the isolates, possessing the desired plasmid construct, pUC7FdCART, was employed for the large scale isolation of plasmid DNA (Fig. 4f). The sequence of the Fd cartridge was confirmed by the subsequent agarose gel isolation of the Fd cartridge as an EcoRI fragment, sub-cloning into EcoRI cleaved M13mp7, and nucleotide sequence determination. The complete nucleotide sequence of the Fd cartridge is given in Fig. 4g.

FIGURE 4f DERIVATION OF THE Fd EXPRESSION CARTRIDGE

Fol +/- denotes the ability to grow on Folate as sole carbon source.

Blue/white denotes colony colour on media supplemented with X-gal.

Restriction sites are as follows: Sua3A, 3A; BamHI, B; EcoRI, E; HindIII, H.

p = transcriptional promoter

t = transciptional terminator

SDM = site-directed mutagenesis

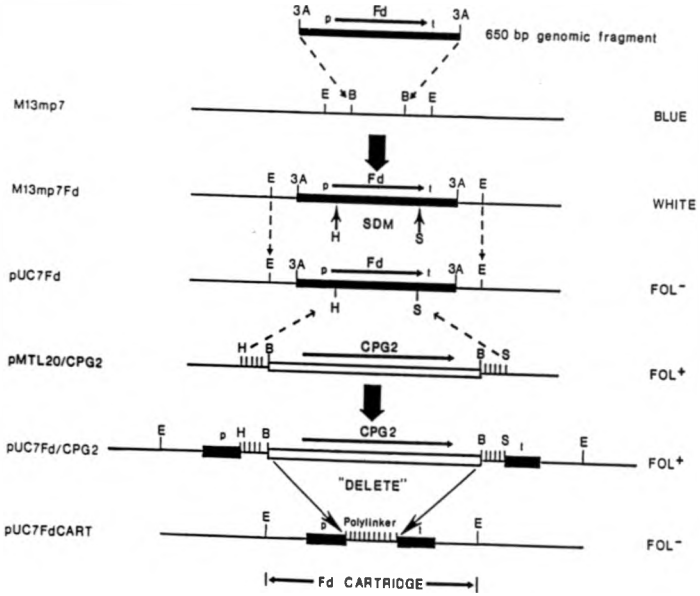

3A
p
Fd
t
3A
650 bp genomic fragment
M13mp7
E
B
B
E
BLUE
M13mp7Fd
E
3A
p
Fd
t
3A
E
WHITE
H
SDM
S
pUC7Fd
E
3A
p
Fd
t
3A
E
FOL⁻
H
S
pMTL20/CPG2
H
B
CPG2
B
S
FOL⁺
pUC7Fd/CPG2
E
p
H
B
CPG2
B
S
t
E
FOL⁺
"DELETE"
pUC7FdCART
E
p
Polylinker
t
E
FOL⁻
Fd CARTRIDGE

FIGURE 4g NUCLEOTIDE SEQUENCE OF THE PORTABLE Fd EXPRESSION CARTRIDGE

Promoters recognised by E. coli, P1 and P2 (Graves and Rabinowitz, 1986) are indicated by dotted lines.

Ribosome binding site (S.D.) is indicated by solid lines.

Arrows indicate the rho-independent transcriptional terminator.

The Sau3A/BamHI junction is indicated by the recreated Sau3A sites (boxed).

```
 EcoRI
GAATTCCCCCGATCGAGATAGTATATGATGCATATTCTTTAAATATAGATAAAGTTATAGAAGCAATAGAAGATTTAGGATTTACTGTAATATAAATTAC
                                                                                                     100
    -35        P1                  -10                                  S.D.                   M   A   Y   K   I
ACTTTTTAAAAAGTTTAAAAACATGATACAATAAGTTATGGTAAACTTATGATTAAAATTTTAACGAGGTGTATTTTTC   ATG GCA TAT AAA ATC
                        -35              P2            -10                                                193
 A   D   S   C   V HindIII  SphI   PstI  StuI  XhoI BglII  NcoI   MluI   AatII SalI   XbaI    BamHI
GCT GAT TCA TGT GTA AGC TTG CAT GCC TGC AGG CCT CGA GAT CTC CAT GGA CGC GTG ACG TCG ACT CTA GAC GAT
                                                                                                     268
  SmaI  KpnI   SstI  P   V   Q   T   *
CCC GGG TAC CGA GCT CCA GTA CAA GAA TAA TTTATAATTAAAACATAAAAATAAGAAGCCTGCAAATGCAGGCTTCTTATTTTTATTGTA
                                                                transcriptional terminator           358

ACTGAATTTTATTATGTTATACTATAAATTGAAGAAAATATTAGGGGGAAATTGGTATGCTTAATATATTTAGGGACACTTTTCAAGTTATGTCGCCGCT
                                                                                                     458

TAATGGAAATATAGTAAATTTAACTAATGTTCCAGACAGAATGTTTTCAGAGGAAATAGTACGAAAAGGAATAGCTGTAGACCCATTACAGGATATAATA
                                                                                                     558

            EcoRI
AGATCCGGGGAATTC
                573
```

### 4.2.3. CONSTRUCTION OF THE SHUTTLE VECTORS pMTL33 AND pMTL34

To facilitate expression studies using the constructed Fd cartridge in both E. coli and B. subtilis genetic backgrounds, a suitable shuttle vector was required. It was decided to construct a vector analogous to pHV33 (Primrose and Ehrlich, 1981) by combining the versatile Gram-negative cloning vector, pMTL21 (Chambers et al.. 1988), and the Gram-positive (B. subtilis) cloning vector, pC194 (Horinouchi & Weisblum, 1982a). This vector, pMTL33, would be superior to pHV33 in that it would carry many more unique cloning sites and enable blue to white selection of recombinant E. coli clones when X-gal is incorporated into the agar medium. In addition to pMTL33, it was elected to construct a deletion derivative of pMTL33, pMTL34, lacking the lac po/lacZ' region of pMTL20 (Fig. 4h). By virtue of a unique PvuII site, this vector would thus facilitate expression studies to evaluate the Fd cartridge in the absence of any interfering transcriptional activity arising from the lac promoter.

#### 4.2.3.1. CONSTRUCTION OF pMTL33

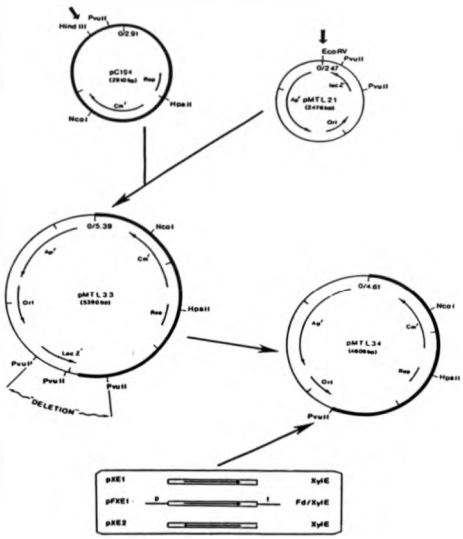
Plasmid pC194 DNA (25 $\mu$g) was linearised by digestion with HindIII and the 2.9 kb fragment blunt-ended with Klenow polymerase. The fragment DNA was further purified by agarose gel electrophoresis (1% w/v). Approximately 100 ng of of linearised pC194 was ligated with an equal quantity of pMTL21 DNA (Chambers et al., 1988) which had been cleaved with EcoRV and remaining restriction enzyme inactivated by heat treatment. The ligation products were used to transform competent E. coli JM83. The desired recombinants were selected for by plating out onto L-agar supplemented with both Ap and Cm. Twelve transformants were chosen for further restriction analysis. Digestion with PvuII of small scale plasmid preparations followed by agarose gel electrophoresis (1% w/v) revealed, in two cases, the desired chimaera with the PvuII site of pC194 proximal to the PvuII sites of pMTL21 (Fig 4h). One of these clones, pMTL33, was

FIGURE 4h

A. Derivation of the shuttle vectors pMTL33 and pMTL34 and the construction of chimeras with the Fd cartridge and the XylE gene

B. Nucleotide sequence of the pMTL21 polylinker

A.




B.

```
                                                                                  AccI
                                                                                  SalI                                  SmaI
                                   XhoI                                                                                 
                 HindIII SphI PstI StuI    BglII NcoI  MluI  AatII    XbaI  BamHI    KpnI  SstI  EcoRI
ATCACCATGATTACGCCAAGCTTGCATGCCTGCAGGCCTCCAGATCTCCATGCACGCCGTCACGTCCACTCTAGAGGATCCCCGGTACCGAGCTCGAATTCACTGRE
 H  T  H  I  T                                                                                          L  R
```

chosen and employed for the large scale isolation of plasmid DNA.

#### 4.2.3.2. CONSTRUCTION OF pMTL34

pMTL33 DNA (25 $\mu$g) was digested with PvuII and the restricted products size fractionated on a preparative agarose gel. The large 4.6 kb PvuII fragment was excised from the gel and the DNA extracted by electroelution, thereby isolating this fragment from the smaller PvuII fragments (345 and 439 bp) which together encompass the lac po/lacZ' region derived from pMTL21. The isolated fragment (4.6 kb) was self ligated and the ligation products used to transform competent E. coli JM83 cells which were subsequently plated out onto L-agar supplemented with Ap and X-gal. Following overnight incubation at 37$^{o}$C, six white colonies were picked and used for small scale plasmid DNA isolations. Digestion of these with PvuII followed by agarose gel electrophoresis (1% w/v) revealed, in each case, the presence of a single PvuII site (Fig. 4h). This single remaining PvuII site was to provide a suitable cloning site for the insertion of blunt-ended DNA fragments, thereby enabling studies of their expression in the absence of the Lac promoter. One of the desired clones, carrying pMTL34 was subsequently employed for the large scale isolation of plasmid DNA.

### 4.2.4. ISOLATION OF THE XYLE GENE

An E. coli JM83 clone was obtained (from The Biotechnology Division, Cranfield Institute), which contained the pseudomonad xylE gene as a 2.2 kb XhoI fragment cloned into the SalI site of pUC18 (Fig. 4i). This fragment is known to carry the majority of the upstream xylL (Inouye et al.. 1986) gene in addition to the xylE gene. As this 900 bp upstream region was surplus to requirements this extra DNA was sequenced in an attempt to identify any useful restriction sites that might help the isolation of the xylE gene as a smaller, self-contained, restriction fragment.

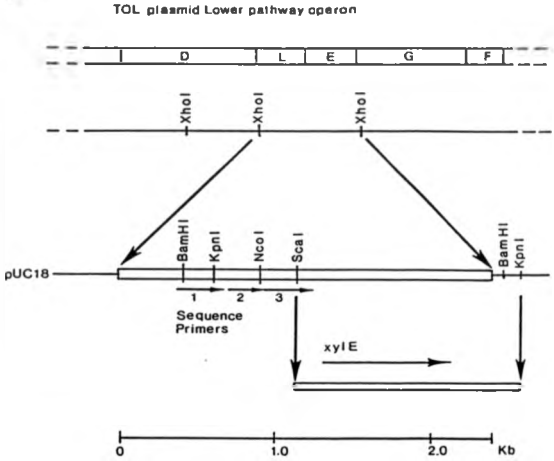
This clone was employed for the large scale isolation of plasmid DNA. Some of this DNA (50μg) was digested with BamHI and the two resultant fragments resolved by agarose gel electrophoresis (1% w/v). This digest released the xylE encoding DNA as a 2.2 kb BamHI fragment by virtue of the previously reported BamHI site in the XhoI fragment (Zukowski et al.. 1983) and the additional BamHI site in the polylinker of pUC18 (Fig. 4i). This fragment was isolated and ligated with M13mp9 which had been cleaved with BamHI and dephosphorylated. The ligation products were used to transfect competent E. coli JM101 cells. Six recombinant plaques were picked and template DNA prepared from them. All six templates were sequenced with universal sequence primer which revealed that all six clones consisted of M13mp9 with the BamHI insert DNA in the same orientation, fortuitously in the desired orientation. Sequencing with universal primer generated some 300 bp of novel DNA in the direction towards xylE. To facilitate further sequencing in that direction a 17 base oligonucleotide primer, xylE primer I (5'-GACGAGCAGGTCGAGGC-3'), was synthesised which was complementary to the sequence between nucleotide positions 240 and 256 of the BamHI fragment sequence (Fig. 4i). Additional Sequencing with xylE primer I generated a further 300 bases of DNA sequence, and as predicted, did not reach as far as the published xylE sequence (Nakai et al, 1983; Zukowski et al.. 1983) necessitating the synthesis of an additional primer. The new primer, xylE primer II (5'-TGCTAGCGGCCAGTCA-3'), also a 17 base oligonucleotide, was synthesised complementary to the the sequence at nucleotide positions 510 to 517 (Fig. 4i). Subsequent sequencing employing xylE primer II generated the remaining DNA sequence, reaching and just overlapping the published sequence. This sequencing revealed the presence of a unique ScaI site 69 bp upstream of the xylE start codon, thereby facilitating the isolation of the xylE gene as a 1.2 kb ScaI-KpnI fragment utilising the KpnI site in the polylinker of pUC18 (Fig. 4i).

FIGURE 41 ISOLATION OF THE XYLE GENE

A. Isolation of the xylE gene from the primary clone

B. Nucleotide sequence upstream of the xylE gene showing unique ScaI site and ribosome binding site

A.




B.

### 4.2.5. EXPRESSION OF XYLE IN E. COLI AND B. SUBTILIS

The expression of the xylE gene in E. coli and B. subtilis when under the control of the Fd transcriptional signals was studied. In order to evaluate the significance of the Fd promoter sequences it also seemed necessary to investigate the levels of expression of xylE when no promoter sequences are provided, thereby providing comparative data. The strategy, therefore, was to insert the xylE gene, with and without the Fd transcriptional control signals, into the unique PvuII restriction site of the shuttle vector, pMTL34, and assay cultures of E. coli and B. subtilis harbouring these chimaeras for the xylE gene product (C23O). In addition, it was decided to obtain comparative data on the orientation of xylE in pMTL34, again with and without the Fd promoter sequences, to ascertain whether or not any transcriptional readthrough from any of the vector sequences might be contributing to levels of expression obtained. It was envisaged that the 1.2 kb xylE ScaI-KpnI fragment would be cloned into pUC7/FdCART and the portable, xylE containing, Fd cartridge sub-cloned as a blunt-ended EcoRI fragment into the PvuII site of pMTL34. The xylE fragment, on its own, would be cloned, again as a blunt-ended fragment, directly into the single PvuII site of pMTL34. In both cases, both orientations would be screened for.

In addition, it was decided to sub-clone the 1.2 kb ScaI-KpnI xylE DNA fragment into pMTL33 where its expression would be under the control of the lac promoter of the pMTL21. Expression studies with E. coli and B. subtilis cells harbouring both orientations of xylE in pMTL33 would therefore provide comparative data to that obtained with the Fd promoter in the pMTL34 constructs.

#### 4.2.5.1. SUB-CLONING OF FdCART/XYLE INTO pMTL34

To isolate the 1.2 kb xylE ScaI-KpnI fragment, the pUC18/xylE clone DNA was digested with HindIII in addition to ScaI and KpnI, thereby eliminating the possibility of

co-isolating a similar sized fragment arising from the ScaI site in pUC7 and the KpnI site upstream of xylE. The restricted products were size fractionated by agarose gel electrophoresis (1% w/v). The 1.2 kb xylE (ScaI-KpnI) fragment was excised from the gel and the DNA extracted by electroelution.

This fragment DNA was subsequently ligated with pUC7/FdCART DNA which had been appropriately cleaved with StuI and KpnI, thereby ensuring that the xylE fragment was cloned in the correct orientation with respect to the Fd transcriptional control signals. The ligation products were used to transform competent E. coli JM83 cells and the transformed cells were plated out onto Ap-containing L-agar. Recombinants containing the xylE fragment were easily identified by the rapid appearance of a bright yellow coloration to the colonies after spraying the resultant transformants with 0.5 M catechol. Four yellow colonies were picked and plasmid DNA prepared from them. The plasmid DNA derived from each isolate was subjected to restriction endonuclease analysis, employing the following restriction enzyme combinations; ClaI-HindIII. ClaI-SstI. ClaI-ScaI. and ScaI-HindIII. The digestion products were size fractionated by agarose gel electrophoresis (1% w/v). The ScaI-HindIII digest confirmed the previously determined orientation of the Fd derived DNA sequences with respect to pUC7 (4.2.2.2.). The remaining digests confirmed that all four isolates were identical with the xylE gene correctly inserted in the Fd cartridge. This clone, carrying the plasmid designated pUC7/FdCART/xylE. was employed for the large scale isolation of plasmid DNA. Some of this DNA (50 µg) was digested with EcoRI. blunt-ended with Klenow polymerase, and the resultant 1.7 kb, xylE containing, blunt-ended fragment separated by agarose gel electrophoresis (1% w/v). The fragment band was excised from the gel and the DNA extracted by electroelution. This fragment DNA was subsequently ligated with pMTL34 which had been cleaved with PvuII and dephosphorylated. The ligation products were used to transform competent E. coli JM83 and the resultant recombinants identified as bright yellow trans-

formants on Ap-containing L-agar plates which had been sprayed with 0.5 M catechol. Small scale plasmid preparations were isolated from twelve yellow transformants and these digested with ClaI to orientate the Fd/xylE cartridge in pMTL34. Size fractionation of the restricted products by agarose gel electrophoresis (1% w/v) identified several clones representing only one orientation. One of these, carrying pFXE1 (Fig. 4h), was chosen for the large scale isolation of plasmid DNA. No further attempts were made to clone the reverse orientation.

4.2.5.2. SUB-CLONING OF xylE INTO pMTL34

The isolated 1.2 kb ScaI-KpnI xylE fragment was blunt-ended with T4 polymerase and ligated with pMTL34 DNA which had previously been cleaved with PvuII and dephosphorylated. The ligation products were used to transform competent E. coli JM83 cells and were plated out onto L-agar supplemented with Ap. Recombinants were identified amongst the resultant transformants again by the appearance of a yellow coloration to the colonies when sprayed with 0.5 M catechol; not exactly the expected observation as the xylE gene had been sub-cloned devoid of any promoter sequences. The yellow coloration observed was noticeably less intense than that observed when the FdCART/xylE fragment was sub-cloned into pMTL34, suggesting that the observed expression of xylE was the result of some transcriptional readthrough from vector derived sequences. This observation has been previously reported (Zukowski et al., 1983). In addition, two slightly differing intensities of yellow coloration were observed suggesting an orientation effect of the xylE fragment with respect to pMTL34. Small scale plasmid DNA preparations were isolated from six of each colour intensity type and subjected to restriction endonuclease analysis to determine the orientation of the inserts. Digestion with ClaI singly, and NcoI-BamHI jointly, followed by agarose gel electrophoresis (1% w/v), confirmed the two orientations and that these were concomitant with the coloration intensities observed. The apparent transcriptional readthrough observed in the more intensely coloured

colonies was due to the pC194 derived sequences of pMTL34. Clones representative of the two orientations, carrying the recombinant plasmids, pXE1 and pXE2, were picked and employed for the large scale isolation of plasmid DNA.

4.2.5.3. SUB-CLONING OF xylE INTO pMTL33

The previously isolated and blunt-ended xylE ScaI-KpnI fragment DNA (4.2.5.2.) was ligated with pMTL33 DNA which had been cleaved with SmaI and dephosphorylated. The ligation products were used to transform competent E. coli JM83 cells and recombinants selected on L-agar supplemented with Ap and X-gal. Further confirmation was obtained by subsequently spraying the transformants with 0.5 M catechol; again two differing intensities in yellow coloration of the colonies was observed suggesting an orientation effect with respect to the lac promoter. Small scale plasmid DNA preparations isolated from six of each colour intensity type were digested jointly with SstI and ClaI and the restricted products size fractionated on an agarose gel (1% w/v). This confirmed the two orientations of the xylE DNA fragment with respect to the lac promoter of pMTL33, again concomitant with the differing coloration intensities observed amongst the respective transformants. Representative clones of each orientation were employed for the large scale production of plasmid DNA, these being the recombinant plasmids pLXE1 and pLXE2.

4.2.5.4. EXPRESSION STUDIES IN E. COLI AND B. SUBTILIS

The multimeric plasmid DNA isolated from the pMTL34 and pMTL33 constructs (i.e., pLXE1, pFXE1, pXE1, and pXE2) was used to transform naturally competent cells of B. subtilis 168. The cells were subsequently plated out onto L-agar supplemented with Cm. $Cm^r$ colonies of each were selected for ensuing expression studies. Overnight shake flask cultures of B. subtilis and E. coli harbouring these constructs were grown in L-broth, supplemented with Cm and Ap respectively, and sonic extracts prepared from them. The levels of the xylE gene product (C23O) were assayed spec-

trophotometrically. Protein levels for the sonic extracts were also determined. From the data obtained specific activities of the C23O enzyme in the samples, and the percentage of total cell soluble protein were calculated. These assays were repeated a further two times and the mean values of the three independent assays taken. This data is given in Table 4.1. Earlier pilot experiments with the C23O assay indicated that maximal levels of the enzyme were obtained after overnight (16 hour) growth.

From the expression data presented in Table 4.1 it is apparent that, in E. coli. the Fd promoter seems to function extremely efficiently, directing the expression of C23O to levels only marginally lower than those achieved when expression was from the E. coli lac promoter (i.e., 13.96% versus 19.86% soluble protein). It was also apparent, however, that expression of xylE occurred even in the absence of either the lac or Fd promoter. In this case the level of expression observed was dependent on the orientation of insertion of the "promoterless" xylE gene in pMTLJ4. The highest activity was seen when the orientation of insertion was such that the transcripts produced would have originated from within pC194 derived DNA (i.e., 0.22% versus 2.12% soluble protein). Even though only one construct representing the one orientation of FdxylE cloned into pMTLJ4 (i.e., pFXE1) was ever used, the high levels of expression of xylE obtained (13.96% soluble protein) were attributable to the Fd promoter as very much reduced levels (0.22% soluble protein) were obtained from the analogous construct, pXE1, which lacked the Fd promoter sequences. In B. subtilis. however, expression of xylE was only observed when the Fd promoter was present albeit at extremely low levels (0.07% soluble protein) in comparison to those levels obtained in E. coli.

TABLE 4.1

EXPRESSION LEVELS OF CATECHOL 2,3-DIOXYGENASE (C230)

1 mU corresponds to the formation at 30°C of one mmole of 2-hydroxymuconic semialdehyde per minute.

| Host | Cloning Vector | Recombinant Plasmid | Promoter | Specific Activity (mU/mg protein) | % Soluble Protein |
|---|---|---|---|---|---|
| E.coli | pMTL33 | pLXE1 | lac | 45, 090 | 19.86 |
| " | pMTL34 | pFXE1 | Fd | 31, 695 | 13.96 |
| " | pMTL34 | pXE2 | none | 4, 806 | 2.12 |
| " | pMTL34 | pXE1 | none | 504 | 0.22 |
| B.subtilis | pMTL33 | pLXE1 | lac | 0 | - |
| " | pMTL34 | pFXE1 | Fd | 165 | 0.07 |
| " | pMTL34 | pXE2 | none | 0 | - |
| " | pMTL34 | pXE1 | none | 0 | - |

4.2.5.5. CONFIRMATION OF THE Fd SEQUENCE

One possible explanation for the low levels of C230 in B. subtilis was that structural rearrangement might have occurred, whether in B. subtilis at this final stage or perhaps at an earlier stage in E. coli. It was therefore decided to re-analyse the various Fd constructs for structural integrity by restriction enzyme analysis and DNA sequencing. The Fd derived DNA of the pUC7/FdCART construct has already been sequenced and shown to be unaltered (4.2.2.2.).

The B. subtilis clone harbouring pFXE1 was used for the small scale isolation of plasmid DNA. This DNA was used to transform competent E. coli JM83 cells and resultant transformants selected on L-agar supplemented with Ap. One transformant was chosen and employed for the large scale isolation of plasmid DNA. This pFXE1 DNA, the previously isolated pFXE1 DNA which had not been passaged through B. subtilis. and the previously sequenced pUC7/FdCART DNA were initially subjected to restriction enzyme analysis singly with HindIII, SphI, PstI. and SstI. and jointly with combinations of EcoRI-PstI and EcoRI-SstI. Subsequent agarose gel (1 % w/v) electrophoresis of the majority of the digestion products confirmed the presence of the polylinker sites, in all three constructs, for HindIII. SphI. PstI. and SstI. In addition, all three constructs released similar sized EcoRI-PstI and EcoRI-SstI fragments (220 and 295 bp respectively) suggesting that all the Fd derived sequences had remained intact showing no obvious signs of structural rearrangement. To further establish this fact some of the remaining EcoRI-PstI and EcoRI-SstI digested DNA, derived from each of the three constructs, was heat treated and ligated with M13mp18 and M13mp19 RF DNA which had been similarly digested and heat treated. The ligation products were used to transfect competent E. coli JM101 cells and template DNA prepared from several resultant recombinants derived from each Fd construct. Subsequent DNA sequencing of the templates revealed that all the Fd derived DNA sequences were identical to that previously

determined (4.2.2.2.).

4.2.5.6. ANALYSIS OF THE SEQUENCE UPSTREAM OF xylE

In consequence of the poor expression of xylE in B. subtilis it was decided to analyse the DNA sequence upstream of the xylE gene for any regions that might contribute to the formation of secondary structure in transcripts initiated at the Fd promoter.

Using the programs of DNASTAR Inc. a 19 bp sequence was identified, lying immediately 3' to the ScaI site of the cloned xylE fragment (4.2.4.), which possesses the potential to form a hairpin loop structure with a $\Delta G$ of -9 kcal $mol^{-1}$ (Fig. 4j) reminiscent of a prokaryotic factor-independent transcriptional terminator. However, this sequence was not followed by a series of thymine residues characteristic of efficient terminators (Platt, 1986). Whether or not this sequence may have been responsible for early termination of transcripts initiated at the Fd promoter in B. subtilis, and not to any significant level in E. coli, was not determined in this study.

FIGURE 4j SECONDARY STRUCTURE PREDICTION OF SEQUENCE UPSTREAM OF XYLE

ΔG value calculated according to Tinoco et al., (1973)

```
      G
   C     U
     A:U
     C:G
     C:G
     G:C
     C:G
     U:G                                                xylE
ScaI U:A                          S.D.                 M   N   K
AGUACU:AAACAAACCUGACAACAUGAACUAUGAAGAGGTGACGTC         ATG AAC AAA
```

$\Delta G = -9 Kcal.mol^{-1}$

## 4.3. DISCUSSION

In this study the ferredoxin (Fd) gene of *C. pasteurianum* was cloned and used to construct a portable expression cartridge comprised of an extensive polylinker region sandwiched by the Fd transcriptional initiation and termination signals. This cartridge was used to direct the expression of the pseudomonad *xylE* gene in *E. coli* and *B. subtilis*, albeit to only low levels in the latter organism.

Previous studies have demonstrated that *E. coli* cells carrying the Fd gene inserted in pBR322 produce a Fd-specific RNA transcript at equivalent levels to *C. pasteurianum* (Graves & Rabinowitz, 1986). This prompted these authors to suggest that the Fd promoter is efficiently utilised by *E. coli* RNA polymerase. The results obtained here support this view demonstrating that the Fd promoter may be used to express the *xylE* gene at levels approaching that obtained with the *lac* promoter. The observed high efficiency of *xylE* expression may, in part, be due to the presence of two promoters (P1 and P2, Fig. 4g) which are recognised by *E. coli* RNA polymerase (Graves & Rabinowitz, 1986).

Although the ability of *B. subtilis* RNA polymerase to transcribe the Fd gene was also examined by Graves and Rabinowitz (1986), the experiments performed were undertaken *in vitro*. The data obtained suggested that the transcription of the Fd gene by *B. subtilis* RNA polymerase was as efficient as with *E. coli* RNA polymerase. No *in vivo* experiments were performed. In our experiments, the Fd promoter only directed the expression of low levels of C23O in *B. subtilis*. Previous studies using the *xylE* gene to detect promoters in *B. subtilis*, and more recently to overexpress C23O in a *B. subtilis* $SacU^h$ pleiotropic mutant, have shown that relatively high levels of C23O may be elicited (Zukowski *et al.*, 1983, 1988; Zukowski & Miller, 1986), indicating that *B. subtilis* is not incapable of efficient expression of this gene. This suggests that, at least *in vivo*, the *C. pasteurianum* Fd promoter is ineffi-

ciently utilised by the transcriptional machinery of B. subtilis.

Since these studies were conducted Leonhardt and Alonso (1988) have suggested that the xylE gene is poorly translated in B. subtilis. These authors obtained levels of expression comparable to those reported in this study when under the transcriptional control of the SP01 promoter, in the concomitant presence of high levels of xylE-specific mRNA as deduced by dot blot analysis. If so, the translational barrier would have to be at either the level of translation initiation or elongation. Restricted initiation of translation appears unlikely as the proposed xylE ribosome binding site exhibits extensive complementarity with the 3' region of the B. subtilis 16S rRNA, with a predicted free energy of base pairing ($\Delta G$) calculated to be $-15$ Kcal $mol^{-1}$ (Zukowski et al.. 1983). In addition, the distance between the ribosome binding site and the initiation codon is calculated to be 9 bases. Both criteria fall well within the ranges of $\Delta G$ and spacer bases calculated from a series of ribosome binding sites for B. subtilis (Moran et al.. 1982). Comparison of the codon usage of xylE. E. coli. and B. subtilis (Table 4.2) reveals that many of the codons rarely used in E. coli are similarly used in xylE. These include the codons for Leu (CUA), Ile (AUA), and Arg (CGA, AGA, and AGG). Also, among the six codons for Leu, CUG is preferentially used, and among six codons for Arg, CGU or CGC is preferentially used. However, the codon usage exhibited by B. subtilis is reported to be remarkedly unbiased (Ogasawara, 1985; Piggot & Hoch, 1985; Table 4.2) leading to the suggestion that B. subtilis might prove to be a good host for the expression of foreign genes (Ogasawara, 1985). Even though it has been suggested that certain codons are more rarely used in B. subtilis (Ogasawara, 1985), none of these are commonplace in xylE with the exception of the codons for Thr (ACA) and Lys (AAG) (Table 4.2).

TABLE 4.2

COMPARISON OF THE CODON UTILISATION OF THE XYLE GENE OF PSEUDOMONAS PUTIDA WITH THAT OF ESCHERICHIA COLI AND BACILLUS SUBTILIS

% Occurrence

| | | xylE | E. coli* | B. subtilis* | | | xylE | E. coli* | B. subtilis* |
|---|---|---|---|---|---|---|---|---|---|
| ARG | CGU | 25.0 | 58.0 | 25.2 | VAL | GUU | 9.5 | 37.5 | 31.4 |
| | C | 40.0 | 35.0 | 17.5 | | C | 19.0 | 12.8 | 29.7 |
| | A | 10.0 | 2.3 | 9.1 | | A | 9.5 | 22.9 | 24.5 |
| | G | 20.0 | 3.2 | 11.1 | | G | 62.0 | 26.8 | 19.4 |
| | AGA | 0.0 | 1.1 | 27.7 | | | | | |
| | G | 5.0 | 0.3 | 9.4 | ILE | AUU | 28.6 | 37.3 | 50.0 |
| | | | | | | C | 71.4 | 62.2 | 39.4 |
| LEU | CUU | 2.9 | 8.6 | 26.1 | | A | 0.0 | 0.5 | 10.6 |
| | C | 23.5 | 6.6 | 9.8 | | | | | |
| | A | 5.8 | 1.8 | 6.3 | LYS | AAA | 35.7 | 76.7 | 75.4 |
| | G | 52.9 | 69.0 | 21.8 | | G | 64.3 | 23.3 | 24.6 |
| | UUA | 0.0 | 5.8 | 22.4 | | | | | |
| | G | 14.7 | 8.1 | 13.6 | ASN | AAU | 33.3 | 24.2 | 53.1 |
| | | | | | | C | 66.6 | 75.8 | 46.9 |
| SER | UCU | 9.1 | 26.6 | 24.5 | | | | | |
| | C | 54.5 | 25.6 | 12.0 | GLN | CAA | 14.3 | 26.6 | 54.2 |
| | A | 0.0 | 8.3 | 18.7 | | G | 85.7 | 73.4 | 45.8 |
| | G | 9.1 | 11.4 | 10.0 | | | | | |
| | AGU | 18.2 | 6.5 | 10.7 | HIS | CAU | 33.3 | 38.9 | 68.6 |
| | C | 9.1 | 21.6 | 24.1 | | C | 66.6 | 61.1 | 31.4 |
| | | | | | | | | | |
| THR | ACU | 11.7 | 23.8 | 14.9 | GLU | GAA | 57.1 | 73.4 | 69.5 |
| | C | 76.5 | 50.6 | 14.1 | | G | 42.9 | 26.6 | 30.5 |
| | A | 5.9 | 5.9 | 43.3 | | | | | |
| | G | 5.9 | 19.7 | 27.9 | ASP | GAU | 29.7 | 51.0 | 63.8 |
| | | | | | | C | 70.3 | 49.0 | 36.2 |
| PRO | CCU | 0.0 | 9.0 | 33.6 | | | | | |
| | C | 25.0 | 6.0 | 9.8 | TYR | UAU | 58.3 | 40.6 | 61.7 |
| | A | 8.3 | 19.9 | 19.1 | | C | 41.7 | 59.4 | 36.2 |
| | G | 66.7 | 65.1 | 37.5 | | | | | |
| | | | | | CYS | UGU | 66.6 | 42.0 | 45.7 |
| ALA | GCU | 21.0 | 27.9 | 27.5 | | C | 33.3 | 58.0 | 54.3 |
| | C | 47.3 | 18.7 | 20.0 | | | | | |
| | A | 21.1 | 22.9 | 27.1 | PHE | UUU | 25.0 | 43.5 | 64.0 |
| | G | 10.5 | 30.5 | 25.4 | | C | 75.0 | 56.5 | 36.0 |
| | | | | | | | | | |
| GLY | GGU | 28.0 | 47.8 | 25.4 | MET | AUG | 100 | 100 | 100 |
| | C | 56.0 | 40.8 | 29.6 | | | | | |
| | A | 8.0 | 4.6 | 31.5 | TRP | UGG | 100 | 100 | 100 |
| | G | 8.0 | 6.8 | 13.5 | | | | | |

* Taken from Garnier and Cole, (1988a)

An alternative explanation for the low levels of C230 in B. subtilis was that structural rearrangement of the Fd sequences had occurred. Previous studies with the cloned Fd gene had indicated that the presence of the 0.6 kb Sau3A fragment carrying the gene in either pBR322, pUC plasmids, or M13 vectors was an unstable situation (Graves & Rabinowitz, 1986). In this study the above fragment was subcloned from the primary clone directly into M13mp7, where the site-directed mutagenesis was undertaken. The mutated gene was then transferred first into pUC7, where the polylinker and the xylE gene were inserted, and then into the shuttle vector pMTL34 prior to transformation of the recombinant plasmid into B. subtilis. It has now become clear that cloned DNA is particularly prone to deletion events and structural rearrangements if the replication of the vector employed involves single stranded intermediates (Janniere & Ehrlich, 1987). It follows that the cloned Fd sequence will be particularly vulnerable to deletions in E. coli when cloned in M13 and in B. subtilis when inserted into vectors based on the pC194 replicon, i.e., pMTL34. During all the manipulations undertaken in M13 no deviations from the expected sequence were observed. However, it should be noted that when the gene was initially sub-cloned into M13mp7 all the recombinants examined had inserted in an identical orientation. This suggests that the insertion of the Sau3A fragment in the opposite orientation may be an unstable situation. Insertion of Fd sequences in pMTL34 was also orientation specific. Thereafter, all subsequent recombinant plasmids generated in E. coli released the expected restriction fragments upon digestion with EcoRI and other restriction enzymes which cleaved within the polylinker employed. As a final check, the Fd sequences were reisolated from the recombinant plasmid cultivated in B. subtilis (pFXE1), cloned into M13, and the presence of the expected sequences confirmed by nucleotide sequence analysis.

The reasons why the xylE gene was poorly expressed in B. subtilis remain unclear, particularly as there is no general data available on the relative ability of B. subtilis

to express genes of a clostridial origin or to translate pseudomonad DNA. In view of the reports that the Fd promoter is efficiently transcribed (in vitro) by the RNA polymerase of B. subtilis (Graves & Rabinowitz, 1986), and that the xylE gene has been expressed in B. subtilis to levels equivalent to 25% total cell soluble protein (Zukowski & Miller, 1986; Zukowski et al., 1988), it is interesting to speculate an alternative reason for the low expression levels obtained. The pseudomonad sequence upstream of the proposed xylE ribosome binding site contains 19 bases capable of forming of a stem-loop structure in transcripts initiated at the Fd promoter. Premature termination of transcripts at this point could account for the low levels of C230 if this putative transcriptional terminator was preferentially recognised by the RNA polymerase of B. subtilis and not the RNA polymerase of E. coli. This hypothesis is supported by (1) the fact that factor-independent transcriptional terminators are utilised to widely varying degrees by E. coli and B. subtilis (Peschke et al., 1985), (2) prematurely terminated transcripts initiated at the SP01 promoter would not have been detected in the studies of Leonhardt and Alonso (1988) as the dot blot analysis performed did not specifically detect xylE transcription as the entire recombinant plasmid was used as the hybridisation probe, and (3) the high levels of xylE expression obtained by Zukowski & Miller (1986) were in a SacU$^{h}$ mutant which is known to stimulate increased expression at the transcriptional level by an anti-termination type mechanism (Zukowski et al., 1988).

In summary, an expression cartridge based on the Fd transcriptional control signals was constructed and used to express the Pseudomonad xylE gene in Escherichia coli and Bacillus subtilis. Although the levels of expression obtained in B. subtilis were somewhat disappointing, this may have simply been the result of poor translation of this marker gene and not poor functioning of the expression cartridge itself. However, further experiments to evaluate the potential usefulness of the expression cartridge in B. subtilis using a different marker gene were beyond the

scope of this study.

# CHAPTER FIVE

# TRANSFER AND EXPRESSION OF THE *XYLE* GENE IN *CLOSTRIDIUM ACETOBUTYLICUM*

## 5.1. INTRODUCTION

At this point in time of the study a clostridial replicon had been characterised (chapter 3) and an expression cartridge, based on the transcriptional control signals of the C. pasteurianum gene, constructed (chapter 4). It was originally intended that the study be concluded with the development of a reliable transformation procedure for the saccharolytic clostridia using an E. coli/Clostridium shuttle vector. Such studies, if successful, would therefore have enabled the evaluation of such a clostridial cloning vector based on the pCB101 replicon, the erythromycin resistance determinant of pAMβ1, and the Fd expression cartridge. However this goal proved to be beyond the scope of this study. As an alternative option it was elected to employ the recently developed conjugal cointegrate transfer system (Oultram et al., 1987) to try and evaluate the usefulness of the Fd cartridge in C. acetobutylicum. Although somewhat cumbersome in nature, this methodology represented the only available means of introducing foreign DNA into any Clostridium.

It was therefore decided to construct a plasmid cloning vehicle analogous to pOD1 (Oultram et al., 1987) that could be transferred to C. acetobutylicum from a B. subtilis donor as a cointegrate molecule with the conjugal mobiliser pAMβ1. The incorporation of a readily detectable gene, under the transcriptional control of the Fd cartridge, in this cloning vehicle would thus test the Fd promoter sequences in C. acetobutylicum following conjugal delivery of the cointegrate plasmid.

As no readily assayable genes of clostridial origin were available at that time, the previously isolated xylE gene of Pseudomonas putida (4.2.4.) was chosen. The strategy

devised was to insert the Fd/xylE containing EcoRI fragment previously isolated (4.2.5.) into the pOD1 analogue vector (pMTL21EC) and assay for the xylE gene product (C230) in C. acetobutylicum. Even though the xylE gene was probably a poor choice as activity of C230 requires the presence of oxygen, there was no reason to suggest that the product could not be assayed for aerobically after the anaerobic growth of the C. acetobutylicum transcipients.

## 5.2. RESULTS

### 5.2.1. CLONING AND SEQUENCING OF THE ERYTHROMYCIN RESISTANCE DETERMINANT OF pAMB1

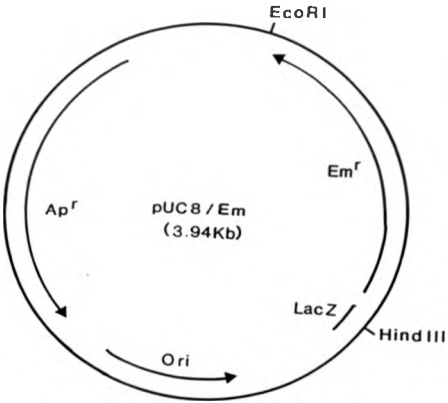
The antibiotic resistance genes chosen as selectable markers, known to function in clostridia (Oultram et al.. 1987), were the erythromycin ($Em^r$) and chloramphenicol resistance ($Cm^r$) determinants of pAM$\beta$1 and pC194 respectively (Leblanc & Lee, 1984; Horinouchi & Weisblum, 1982a). The sequence of the $Cm^r$ determinant has previously been reported (Horinouchi & Weisblum, 1982a) allowing it's isolation as a 1070 bp HpaII fragment from the cloning vector pBD64 which is a deletion derivative of pBD12 which was constructed by combining the pC194 and pUB110 replicons (Gryczan et al., 1980). The $Em^r$ determinant (adenine methylase gene) of pAM$\beta$1 had been previously mapped to a 1.2 kb HhaI fragment (Leblanc & Lee, 1984), and subsequently subcloned as a 3.0 kb EcoRI/HindIII fragment from a deletion derivative of pAM$\beta$1, pVA1, into the streptococcal cloning vector pVA380-1, yielding pVA736 (Macrina et al.. 1980). Further sub-cloning of this resistance determinant as a 1.8 kb AvaI/HindIII from pVA736 into the E.coli cloning vector pRI (Mejean et al.. 1981) yielded pR29 (Vasseghi, and Claverys, 1983).

In order to facilitate the use of the $Em^r$ determinant of pAMβ1, It was decided to sub-clone it as a 1.2 kb HhaI fragment from pR29, check that it was phenotypically expressed, and determine its complete nucleotide sequence.

### 5.2.1.1. SUB-CLONING OF THE $EM^r$ DETERMINANT OF pAMβ1

A recombinant clone of E.coli HB101 harbouring pR29 was obtained from M. Young (University of Aberystwyth) and subsequently employed for the large scale isolation of plasmid DNA. This plasmid DNA was digested with HhaI. the protruding 3' termini blunt-ended with T4 DNA polymerase, and the desired 1.2 kb fragment isolated by agarose gel electrophoresis (1% w/v) and electroelution. This fragment DNA was subsequently ligated with pUC8 DNA which had been cleaved with SmaI and dephosphorylated. The digestion products were used to transform both competent E.coli JM83 and competent E.coli DB11 ($Em^S$) cells. As a positive control, owing to its proven ability to transform E.coli DB11 to $Em^r$, pR29 DNA was also used in the transformation. In addition to the many transformants obtained with E.coli JM83, many $Em^r$ transformants were obtained with E.coli DB11 when plated on Em-containing L-agar, suggesting that the entire functional adenine methylase gene was present. Twelve $Em^r$ transformants were chosen and small scale plasmid DNA preparations isolated from them. Digestion of these DNA preparations with a combination of EcoRI and HindIII followed by size fractionation of the restricted products by agarose gel electrophoresis released, from pUC8, the expected 1.2 kb DNA insert fragment. One of these clones was chosen (pUC8/Em) and employed for the large scale isolation of plasmid DNA (Fig. 5a).

FIGURE 5a THE CHIMAERIC PLASMID pUC8/EM



### 5.2.1.2. DNA SEQUENCE DETERMINATION OF THE $Em^r$ DETERMINANT OF pAMβ1

Plasmid DNA (50 μg) isolated from pUC8/Em was digested jointly with EcoRI and HindIII and the restricted DNA size fractionated by agarose gel electrophoresis (1% w/v). The 1.2 kb HindIII-EcoRI ($Em^r$) fragment DNA was extracted and purified by electroelution. This DNA was subsequently ligated with M13mp8 and M13mp9 DNA which had been similarly digested with HindIII and EcoRI and remaining restriction enzyme activity removed by heat inactivation. The ligation products were used to transform competent E.coli JM101 cells. Thousands of recombinant phage plaques were obtained with the M13mp8 vector in comparison to the few (approximately 50) obtained with the M13mp9 vector. The cloning into M13mp9 was repeated with a newly prepared vector with a similar result. Template DNA was prepared from twelve white plaques from each of the cloning experiments and DNA sequencing undertaken utilising sequence universal primer. DNA sequencing revealed that all the M13mp9 derived templates were simply M13 vector sequences and that all the M13mp8 derived templates gave an identical sequence which was similar to that of the equivalent regions of Tn917 (Shaw & Clewell, 1985) and pAM77 (Horinouchi et al.. 1983). As a result of the observed DNA sequence homology, between the $Em^r$ determinants of pAMβ1 and Tn917, so far obtained, it was decided to synthesise a number of oligonucleotide sequence primers complementary to distal portions of the Tn917 sequence, and use these to determine the remaining nucleotide sequence of the pAM 1 adenine methylase gene. DNA sequencing with the oligonucleotide primers Em1 (5'-TAAGAAGGAGTGATTAC-3'), Em2 (5'-TCACCAAGATATTCTAC-3'), Em3 (5'-AGTAAACAGTGTCTTAA-3'), and Em4 (5'-TAATTCTATGAGTCGCT-3') complementary to the Tn917 sequence at positions 491, 953, 1216, and 1450 respective-

ly, resulted in the elucidation of the remaining nucleotide sequence (Brehm et al.. 1987a). The complete sequence is given in Fig. 5b. A comparison with the equivalent regions of Tn917 and pAM77 is also presented. The major overall difference is the presence of a deletion in the pAMβ1 sequence (nt 254 of the sequence) which removes ORF1, known to be responsible for the induction of Em resistance in the case of Tn917 and pAM77. Its absence explains the observed constitutive expression of Em resistance in cells harbouring pAMβ1 and its derivatives.

### 5.2.2. CONSTRUCTION OF THE pOD1 ANALOG, pMTL21EC

The plasmid vector, pMTL21EC, was designed such that it could be transferred as a cointegrate molecule with pAMβ1 during conjugation between B. subtilis and C. acetobutylicum in an analogous fashion to the published plasmid pOD1 (Oultram et al.. 1987): i.e., cointegration with pAMβ1 could occur by virtue of the homologous $Em^r$ determinant, and subsequent maintenance of the cointegrate molecule in C. acetobutylicum imposed by selection for both Em and Cm. This vector, which is based on pMTL21 (Chambers et al.. 1988), is thus smaller than pOD1 (Oultram et al.. 1987) and contains an extensive polylinker region at the start of the E.coli lacZ' gene. It therefore imparts a blue coloration to colonies of suitable E.coli strains (e.g., JM83) grown on agar medium supplemented with X-gal. Cloning into the polylinker inactivates the lacZ' activity giving colourless colonies, thus providing a simple screen for insertional inactivation.

FIGURE 5b SEQUENCE OF THE ADENINE METHYLASE GENE OF THE STREPTOCOCCUS FAECALIS PLASMID pAMβ1

The complete nucleotide sequence of the adenine methylase gene of the Streptococcus faecalis plamsid pAMβ1. The illustrated sequence extends 1211 bp with the adenine methylase gene labelled as ORF2. A comparison with the equivalent regions from pAM77 (Horinouchi et al., 1983) and Tn917 (Shaw & Clewell, 1985) is also presented. Homology with Tn917 does not begin until nucleotide 132. The previously determined sequence of pAM77 corresponds to nt 151 (*) onwards. Differences at the amino acid and nucleotide sequence level are indicated by upper (Tn917) and lower (pAM77) case letters above and below the sequence, respectively.

```
GCAAAACAAAAACCAAATCATACACCAATCAGTGCAAAAAAAGATATAATGGCAGATAAGACGGTTCGTGTTCGT 75
CATTTTTAAATTGGCACAAACAGGTAACCGTTATTGCAGGTCTATTTCTTATCTATGGGTTTAACATGGATTTTA

GCTGACTTGCACCATATCATAAAAATCCAAACAGCAAAGAATGGCGGAAACCTAAAAGAACTTATGGAAATAAGA 150
TCATTAAAATCATCAGTATTGTCCGAGAGTGATTGGTCTTGCGTATGGTTAACCCTA

                                -35                     -10
CTTAGAAGCAAACTTAAGAGTGTGTTGATAGTGCAGTATCTTAAAATTTTGTATAATAGGAATTGAAGTTAAATT 225
*                                T

                                S.D.   ORF 1  M  L  V  F  Q  M  R  N  V  D  K  T
AGATGCTAAAAATTTGTAATTAAGAAGGAG-------------------------------------------
                                    GGATTCGTCATGTTGGTATTCCAAATGCGTAATGTAGATAAAACA

      i
 S  T  V  L  K  Q  T  K  N  S  D  Y  A  D  K  Y  V  R  L  I  P  T  S  D  .
-----------------------------------------------------------------------------
TCTACTGTTTTGAAACAGACTAAAAACAGTGATTACGCAGATAAATACGTTAGATTAATTCCTACCAGTGACTAA
    a                                   t

-----------------------------------------------------------------------------
TCTTATGACTTTTTAAACAGATAACTAAAATTACAAACAAATCGTTTAACTTCTGTATTTATTTATAGATGTATC
                                                                        c

                                                                N
     S.D.  ORF 2  M  N  K  N  I  K  Y  S  Q  N  F  L  T  S  E  K  V  L  N
----------TGATTACATGAACAAAAATATAAAATATTCTCAAAACTTTTTAACGAGTGAAAAAGTACTCAACC 320
ACTTCAGGAG            a                                 A     G

Q  I  I  K  Q  L  N  L  K  E  T  D  T  V  Y  E  I  G  T  G  K  G  H  L  T
AAATAATAAAACAATTGAATTTAAAAGAAACCGATACCGTTTACGAAATTGGAACAGGTAAAGGGCATTTAACGA 395

T  K  L  A  K  I  S  K  Q  V  T  S  I  E  L  D  S  H  L  F  N  L  S  S  E
CGAAACTGGCTAAAATAAGTAAACAGGTAACGTCTATTGAATTAGACAGTCATCTATTCAACTTATCGTCAGAAA 470
```

```
                N
Y  K  I  V  G  S  I  P  Y  H  L  S  T  Q  I  I  K  K  V  V  F  E  S  H  A
ATAAAATTGTTGGGAGTATTCCTTACCATTTAAGCACACAAATTATTAAAAAAGTGGTTTTTGAAAGCCATGCGT 620
                A

S  D  I  V  L  I  V  E  E  G  F  Y  K  R  T  L  D  I  H  R  T  L  G  L  L
CTGACATCTATCTGATTGTTGAAGAAGGATTCTACAAGCGTACCTTGGATATTCACCGAACACTAGGGTTGCTCT 695
                                                                  tt

L  H  T  Q  V  S  I  Q  Q  L  K  L  P  A  E  C  F  H  P  K  P  K  V  N
TGCACACTCAAGTCTCGATTCAGCAATTGCTTAAGCTGCCAGCGGAATGCTTTCATCCTAAACCAAAAGTAAACA 770

S  V  L  I  K  L  T  R  H  T  T  D  V  P  D  K  Y  W  K  L  Y  T  Y  F  V
GTGTCTTAATAAAACTTACCCGCCATACCACAGATGTTCCAGATAAATATTGGAAGCTATATACCTACTTTGTTT 845

S  K  W  V  N  R  E  Y  R  Q  L  F  T  K  N  Q  F  H  Q  A  N  K  H  A  K
CAAAATGGGTCAATCGAGAATATCGTCAACTGTTTACTAAAAATCAGTTTCATCAAGCAATGAAACACGCCAAAG 920

V  N  N  L  S  T  V  T  Y  E  Q  V  L  S  I  F  N  S  Y  L  L  F  N  G  R
TAAACAATTTAAGTACCGTTACTTATGAGCAAGTATTGTCTATTTTTAATAGTTATCTATTATTTAACGGGAGGA 995

K  .
AATAATTCTATGAGTCGCTTTTGTAAATTTGCAAAGTTACACGTTACTAAAGGGAATGTAGATAAATTATTAGGT 1070


ATACTACTGACAGCTTCCAAGGAGCTAAAGAGGTCCCTAGC 1211
```

### 5.2.2.1. CONSTRUCTION OF pMTL21E

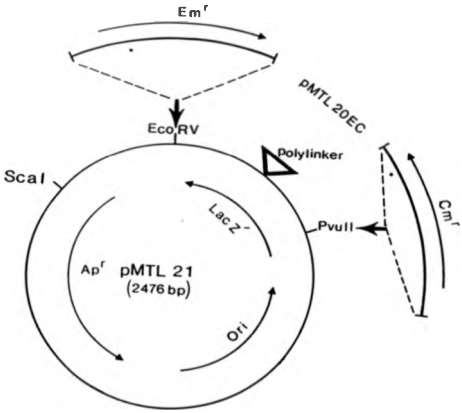
Some of the previously isolated HindIII/EcoRI ($Em^r$) fragment (5.2.1.2.) was blunt-ended with T4 DNA polymerase prior to ligation with pMTL21 DNA which had been cleaved with EcoRV and dephosphorylated. Subsequent transformation of competent E.coli DB11 cells ($Em^s$) yielded $Em^r$ transformants, six of which were chosen for the isolation of small scale plasmid preparations. Digestion of these preparations with ScaI followed by size-fractionation of the products by agarose gel electrophoresis (1% w/v) revealed that the $Em^r$ fragment had been cloned in both orientations. One transformant was chosen (pMTL21E), with the orientation of insertion of the $Em^r$ determinant as shown in Fig. 5c, and employed for the large scale isolation of plasmid DNA.

### 5.2.2.2. CONSTRUCTION OF pMTL21EC

Approximately 50 $\mu$g pBD64 DNA was digested with HpaII and the products resolved on a preparative 1% (w/v) agarose gel. The 1070 bp ($Cm^r$) fragment was extracted from the gel by electroelution and blunt-ended with T4 DNA polymerase. This fragment DNA was then ligated with pMTL21E DNA which had been cleaved with PvuII and dephosphorylated. The ligation products were used to transform competent E.coli DB11 cells and subsequent plating onto L-agar supplemented with both Em and Cm yielded transformants, six of which were employed for the small scale isolation of plasmid DNA. Digestion of these preparations jointly with ScaI and StuI and subsequent resolution of the products by agarose gel electrophoresis (1% w/v) revealed that the $Cm^r$ determinant had been cloned in both orientations. One transformant, pMTL21EC, with the orientations of both resistance genes as shown in Fig. 5c, was chosen for large scale isolation of plasmid DNA.

FIGURE 5c DERIVATION OF pMTL21EC

* Indicates the positions of asymmetric ScaI and Stu I restriction sites in the $Em^r$ and $Cm^r$ resistance genes respectively.



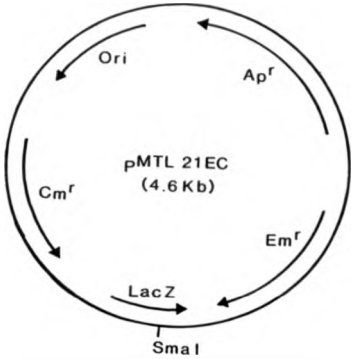
### 5.2.3. CONJUGAL COINTEGRATE TRANSFER OF THE XYLE GENE TO C.ACETOBUTYLICUM

The strategy devised was to sub-clone the xylE gene both with and without the Fd promoter sequences into the vector pMTL21EC. Subsequent cointegration of these chimaeras with the conjugal mobiliser, pAMβ1, and delivery to C. acetobutylicum via conjugation would thus evaluate the usefulness of the Fd promoter in directing the expression of xylE in this host (Brehm et al., 1987b, 1988).

#### 5.2.3.1. SUB-CLONING OF THE XYLE GENE INTO pMTL21EC

The blunt-ended Fd/xylE EcoRI and xylE ScaI/KpnI fragments previously isolated (4.2.5.) were ligated with pMTL21EC DNA which had been digested with SmaI and dephosphorylated. The ligation products were used to transform competent E.coli JM83 cells and recombinants identified as colourless colonies on L-agar supplemented with X-gal and Ap. Twelve isolates were chosen from each cloning and small scale plasmid isolations prepared from them. Digestion of the preparations with a combination of ScaI and ClaI and size-fractionation of the products by agarose gel electrophoresis (1% w/v) revealed that each fragment had been cloned in both orientations. Isolates representing each orientation of xylE in pMTL21EC, both with and without the Fd promoter sequences, were chosen and employed for the large scale isolation of plasmid DNA. The derivation of the chimaeric plasmids p21ECFX1, p21ECFX2, and p21ECX1, and p21ECX2 is summarised in Fig 5d.

FIGURE 5d DERIVATION OF THE CHIMAERIC PLASMIDS p21ECX1, pECX2 p21ECFX1, AND p21ECFX2





P = Promoter t = terminater

### 5.2.3.2. CONJUGAL COINTEGRATE TRANSFER OF XYLE TO C.ACETOBUTYLICUM

Plasmids p21ECFX1, p21ECFX2, p21ECX1, and p21ECFX2 were used to transform competent whole cells of B. subtilis 168 strains, either carrying or not carrying pAMβ1, and selection imposed for $Cm^r$. B. subtilis transformants were only obtained when pAMβ1 was present. This observation was consistent with the notion that pMTL21EC chimaeras only become established within the B. subtilis cell by cointegration with resident pAMβ1 DNA by virtue of the homologous $Em^r$ sequences (Oultram et al.. 1987). B. subtilis transformants carrying all four pMTL21EC chimaeras were subsequently used as donors in conjugation experiments with C. acetobutylicum as outlined in methods. Transconjugants were obtained from each experiment, albeit at low frequencies of around $10^{-8}$ transconjugants per recipient colony. After restreaking and anaerobic incubation of the transconjugants on RCM agar medium they were sprayed with a 0.5 M catechol solution in an aerobic environment. A yellow coloration was imparted to the clostridial cells only in the cases where the Fd promoter sequences were present, i.e., in the case of p21ECFX1 and p21ECFX2 (Fig. 5e, 5f).

### 5.2.4. EXPRESSION OF XYLE IN C. ACETOBUTYLICUM

Overnight broth cultures of C. acetobutylicum possessing the pAMB1::p21ECFX1 and pAMB1::p21ECFX2 cointegrates were grown anaerobically in the presence of Em and Cm. The cells were harvested and sonic extracts prepared from them. Levels of the xylE gene product (C23O) were assayed spectrophotometrically.

The level of protein present in the extracts was also determined. The levels of C23O obtained were disappointing-

ly low, representing only 0.001% total cell soluble protein. The experiments were repeated a further two times with no change in result, suggesting that the C230 levels observed in C. acetobutylicum were almost a hundredfold lower than those observed for B. subtilis and greater than ten thousandfold lower than in E.coli (4.2.5.4.).

FIGURE 5e STRATEGY FOR COINTEGRATE CONJUGAL TRANSFER OF XYLE INTO C. ACETOBUTYLICUM

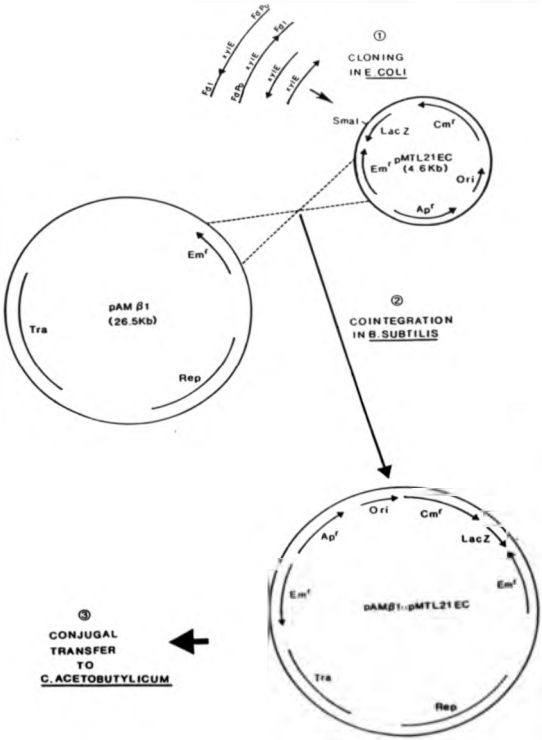


Figure 5f Catechol plate assays for the XYLE gene product (C230)

A. E.coli JM83 containing p21ECFX1

B. B. subtilis containing p21ECFX1::pAMB1 cointegrate

C. C. acetobutylicum containing p21ECFX1::pAMB1 cointegrate (+)
in addition to a negative control of C. acetobutylicum (-)

## 5.3. DISCUSSION

In this study the erythromycin resistance determinant (adenine methylase gene) of pAMβ1 was sub-cloned and its nucleotide sequence determined. In conjunction with the chloramphenicol resistance determinant of pBD64, and the plasmid cloning vector, pMTL21, the pOD1 analogue vector pMTL21EC was constructed. This vector was used to clone the xylE gene and facilitate its transfer to C. acetobutylicum as a cointegrate with the conjugal mobilisor, pAMβ1, where it was inefficiently expressed.

The reason(s) why the xylE gene was expressed at such low levels in C. acetobutylicum are unknown, but may be, in part at least, due to the same reason(s) for the low levels of C23O recorded in B. subtilis. In C. acetobutylicum the levels of C23O were almost one hundredfold lower than those obtained in B. subtilis. a large difference that is unlikely to be attributable to the lower gene dosage in C. acetobutylicum resultant of the pAMβ1 replicon. In view of the recent report by Leonhardt & Alonso (1988) that the xylE gene is poorly translated in B. subtilis it follows that the same might also apply to clostridia, particularly as the difference in mol G + C is even greater (57% v 28%). For all the clostridial genes sequenced to date there is a strong preference for dA or dT in the third codon position ("wobble position") (Beguin et al. 1985; Chen et al.. 1986; Fairweather & Lyness, 1986; Garnier & Cole, 1988a; Graves et al.. 1985; Grepinet & Beguin, 1986; Hinton & Freyer, 1986; Janssen et al.. 1988; Joliff et al.. 1986; Wang et al., 1987). This is not the case for xylE. However, nothing is known about the relative abundances of the various iso-accepting tRNA species of clostridia, nor the genetic organisation of their encoding sequences, which might otherwise help to understand whether or not codon bias

might play a significant role in gene expression in clostridia.

Poor translation of the xylE gene in C. acetobutylicum could also be affected at the level of initiation. As no sequence data is available for the 3' end of any clostridial 16S rRNA, it is not known whether or not conventional ribosome binding sites are recognised by these organisms. However, of the clostridial genes sequenced to date, the majority possess sequences 5' to the initiation codons which exhibit significant complementarity with the sequence of the 3' end of B. subtilis 16S rRNA. This therefore suggests a common consensus between the ribosome binding sites of these Gram-positive organisms.

A final consideration should be given to the promoter sequences used in this study. Even though the Fd promoter appears to be typical of Gram-positive promoters in general (Graves & Rabinowitz, 1986) it cannot, however, be assumed that there is no inter-species transcriptional barriers between different clostridial species. It would therefore be of interest to transfer the p21ECFX::pAMβ1 cointegrates to C. pasteurianum by conjugation and study the level of expression of C23O. Unfortunately, such experiments were beyond the scope of this study.

In summary, the adenine methylase gene ($Em^r$) of the streptococcal R-factor pAMβ1 was isolated, it's complete nucleotide sequence determined, and used in the in vitro construction of a shuttle vector for conjugal cointegrate transfer into C. acetobutylicum. This vector, pMTL21EC, was successfully used in conjunction with pAMβ1 to transfer and express the pseudomonad xylE gene in C. acetobutylicum, albeit at only a very low level.

# CHAPTER SIX

# GENERAL DISCUSSION

At the outset of the work detailed in this thesis, host/vector methodology for the saccharolytic clostridia such as C. acetobutylicum did not exist. The two major paucities apparent at the time were: (1), the complete lack of any suitable identifiably marked plasmid from a saccharolytic Clostridium that might serve as cloning vector, and (2), a reliable means of delivering such a vector to the intracellular environment of the host. In addition, very little was known about the mechanisms involved in the expression of clostridial genes, reflected by only two reports of cloned clostridial genes (Graves et al.. 1985; Karube et al.. 1983). Despite these problems, in the climate of awakening awareness of the great biotechnological importance of the saccharolytic clostridia, it seemed prudent to initiate studies towards the development of such an invaluable facility.

Since the outset of this work there have been numerous reports of cloned genes from saccharolytic clostridia (see Table 1.1) and to a much lesser extent, isolation of cryptic plasmids from these organisms (see 1.3.1.7.). However, none of which have been characterised to any appreciable degree. Perhaps the most useful contribution to clostridial genetics, to emerge during this period, has been the development of conjugal cointegrate transfer methodology.

During the course of this study a cryptic plasmid indigenous to C. butyricum (pCB101) has been completely sequenced and it's minimal replicon characterised to the extent of elucidating some of the mechanics of it's replication machinery (Chapter 3). Also, an expression cartridge based on the transcriptional control signals of the C. pasteurianum ferredoxin gene was constructed and used to express the pseudomonad xylE gene in both B. subtilis (Chapter 4; Brehm et al. submitted) and C. acetobutylicum (Chapter 5). In

addition, the adenine methylase gene ($Em^r$) of the broad host range streptococcal plasmid pAMβ1 has been completely sequenced (Brehm et al., 1987a) and used in the construction of E. coli/B. subtilis/C. acetobutylicum shuttle vectors to enable conjugal transfer of the xylE gene into C. acetobutylicum (Chapter 5). However, as originally intended, studies towards the development of alternative transformation methodology using E. coli/Clostridium shuttle vectors based on the pCB101 minimal replicon proved beyond the scope of this study.

The DNA sequence of pCB101 represents the first reported sequence for a plasmid derived from a saccharolytic Clostridium, although the second reported sequence of a clostridial plasmid to date. In addition, evidence has been supplied supporting the notion that pCB101 replicates via an asymmetric rolling circle type mechanism in a similar way to some other staphylococcal and streptococcal plasmids. However, in contrast to these other plasmids, the minimal replicon of pCB101 is comprised of two or even three essential ORFs and a "plus" origin of replication. It follows that control of replication of pCB101 might similarily be effected by short countertranscribed RNA molecules which block the synthesis of the trans-active replication protein. Future work should identify these RNA species.

In view of the disappointing levels of expression of the xylE gene in both B. subtilis and C. acetobutylicum, when under the transcriptional control of the Fd promoter, further work is required to ascertain the usefulness of the expression cartridge itself. Initially, studies should be directed towards the expression of different marker genes whose products are readily assayable. Currently availalable promoterless candidates include the chloramphenicol acetyl

transferase gene (CAT) of pC194 (Horinouchi & Weisblum, 1982a) the alpha-amylase gene of B. licheniformis (Stephens et al. 1984) and the superoxide dismutase gene from B. stearothermophilus (Brehm et al. unpublished data). Should the Fd promoter prove useful, it could be regulated by incorporating the E. coli lac operator sequence 3' to the Fd promoter sequences in the cartridge. With the concomitant expression of the lac repressor gene on the same or other compatible plasmid the Fd promoter could be modulated by IPTG, an inducer of the lac operon in E. coli. This system has previously been shown to work in B. subtilis (Yansura & Henner, 1984).

Subsequent to the work in this thesis, transformation of C. acetobutylicum whole cells with a bifunctional vector by electroporation has been reported (Oultram et al., 1988b). Using this methodology chimaeric E. coli/C. acetobutylicum shuttle vectors based on the pCB101 minimal replicon have been introduced into C. acetobutylicum where they are stably maintained at a sufficiently high copy number as determined by visualisation of the plasmid in agarose gels (Oultram et al., unpublished data).

In conclusion, Since the initiation of this study and additional work since, the basics for a host/vector system for C. acetobutylicum now exist: i.e., selectable bifunctional shuttle vectors and reliable transformation methodology. Should the Fd promoter cartridge prove functional and perhaps controlable, it should facilitate the overexpression of commercially important heterologous genes in C. acetobutylicum in addition to extending the substrate utilisation range of this industrially important host.

## BIBLIOGRAPHY


Abraham, L.J., and Rood, J.I. (1985a). Molecular analysis of transferable tetracycline resistance plasmids from Clostridium perfringens. J. Bact. 61: 636-640.

Abraham, L.J., and Rood, J.I. (1985b). Cloning and analysis of the Clostridium perfringens tetracycline resistance plasmid, pCW3. Plasmid 13: 155-162.

Abraham, L.J., and Rood, J.I. (1987). Identification of Tn4451 and Tn4452, chloramphenicol resistance transposons of Clostridium perfringens. J. Bact. 169: 1579-1584.

Abraham, L.J., Berryman, D.I., and Rood, J.I. (1988). Hybridization analysis of the class P tetracycline resistance determinant from the Clostridium perfringens R-plasmid, pCW3. Plasmid 19: 113-120.

Achen, M.G., Davidson, B.E., and Hillier, A.J. (1986).Construction of plasmid vectors for the detection of streptococcal promoters. Gene 45: 45-49.

Allcock, E.R., Reid, S.J., Jones, D.T., and Woods, D.R. (1982). Clostridium acetobutylicum protoplast formation and regeneration. Appl. Environ. Microbiol. 43: 719-721.

Alonso, J.C. and Taylor, R.H. (1987). Initiation of plasmid pC194 replication and its control in Bacillus subtilis. Mol. Gen. Genet. 210: 476-484.

An, G., and Friesen, J.D. (1979). Plasmid vehicles for direct cloning of Escherichia coli promoters. J. Bact. 140: 400-407.

Anagostopolous, C., and Spizizen, J. (1961). Requirements for transformation in Bacillus subtilis. J. Bact. 81: 741-746.

Antranikian, G., Zablowski, P., and Gottschalk, G. (1987). Conditions for the overproduction and excretion of thermostable alpha-amylase and pullulanase from Clostridium thermohydrosulphuricum DSM 567. Appl. Microbiol. Biotechnol. 27: 75-81.

Aoyama, T., and Takanami, M. (1985). Essential structure of Escherichia coli promoters II. Effect of the sequences around the RNA start point on promoter function. Nucleic Acids Res. 13: 4085-4096.

Arai, T., Kusakabe, A., Nakashio, S., and Nakamura, M. (1984). A survey of plasmids in Clostridium difficile strains. Kitasato Arch. Exp. Med. 57: 285-288.

Atkins, J.F., Gesteland, R.F., Reid, B.R., and Anderson, C.W. (1979). Normal tRNAs promote ribosomal frameshifting. Cell 18: 1119-1131.

Baas, P.D. (1985). DNA replication of single-stranded Escherichia coli DNA phages. Biochim. Biophys. Acta. 825: 111-139.

Band, L., Yansura, D.G., and Henner, D.J. (1983). Construction of a vector for cloning promoters in Bacillus subtilis. Gene 26: 313-315.

Banner, C.D.B., Moran, C.P.Jr., and Losick, R. (1983). Deletion analysis of a complex promoter for a developmentally regulated gene from Bacillus subtilis. J. Mol. Biol. 168: 351-365.

Barker, H.A. (1961). Fermentation of nitrogenous organic compounds. In, I.C. Gunsalus, and R.Y. Stanier (eds.), The Bacteria. 2: 151-207.

Bartlett, J.G., Taylor, N.S., Chang, T., and Dzink, J. (1980). Clinical and laboratory observations in Clostridium difficile colitis. Am. J. Clin. Nutr. 33: 2521-2526.

Beesch, S.C. (1952). Acetone-butanol fermentation of sugars. Eng. Proc. Dev. 44: 1677-1682.

Beguin, P., Cornet, P., and Aubert, J.P. (1985). Sequence of a cellulase gene of a thermophilic bacterium Clostridium thermocellum. J. Bact. 162: 102-105.

Beguin, P., Rocancourt, M., Chebrou, M.C., and Aubert, J.P. (1986). Mapping of mRNA encoding endoglucanase A from Clostridium thermocellum. Mol. Gen. Genet. 202: 251-254.

Bibb, M.J., Ward, J.M., and Hopwood, D.A. (1978). Transformation of plasmid DNA into Streptomyces at high frequency. Nature 274: 398-400.

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

Bisaria, V.S., and Goze, T.K. (1981). Biodegradation of cellulosic materials: substrates, microorganisms, enzymes and products. Enzyme. Microbiol. Technol. 3: 90-104.

Blaschek, H.P., and Klacik, M.A. (1984). Role of DNAase in the recovery of plasmid DNA from Clostridium perfringens. Appl. Environ. Microbiol. 48: 178-181.

Blaschek, H.P., and Solberg, M. (1981). Isolation of a plasmid responsible for caseinase activity in Clostridium perfringens ATCC 3626B. J. Bact. 147: 262-266.

Bowring, S.N., and Morris, J.G. (1985). Mutagenesis of Clostridium acetobutylicum. J. Appl. Bacteriol. 58: 577-584.

Brefort, G., Magot, M., Ionesco, H., and Sebald, M. (1977). Characterization and transferability of Clostridium perfringens plasmids. Plasmid 1: 52-66.

Brehm, J.K., Barstow, D.A., Pennock, A., Young, M., and Minton, N.P. (1986). Characterization of a replication origin from a cryptic plasmid of Clostridium butyricum. Abstracts of XIV Int. Cong. Microbiol., p. 132.

Brehm, J.K., Oultram, J.D., Thompson, D.E., Swinfield, T.J., and Minton, N.P. (1987b). Transfer of homologous and heterologous genes into Clostridium acetobutylicum. Abstract CEP-BAP Meeting, Ioannina, p. 51

Brehm, J.K., Oultram, J.D., Thompson, D.E., Swinfield, T.J., Peck, H., Young, M., and Minton, N.P. (1988). Construction of plasmid vector systems for Clostridium acetobutylicum. p.409-414. In, A.T. Ganeson and J.A. Hoch (eds.), Genetics and Biotechnology of Bacilli, Volume II. Academic Press Inc. New York.

Brehm, J.K., Thompson, D.E., Swinfield, T.J., Oultram, J.D., Salmond, G.P.C., and Minton, N.P. Construction of an expression cartridge based on the ferredoxin gene of Clostridium pasteurianum. Submitted to Gene.

Brehm, J.K., Salmond, G., and Minton, N.P. (1987a). Sequence of the adenine methylase gene of the Streptococcus faecalis plasmid pAMβ1. Nucleic Acids Res. 15: 3177.

Brenner, D.G. and Shaw, W.V. (1985). The use of synthetic oligonucleotides with universal templates for rapid DNA sequencing: results with staphylococcal replicon pC221. EMBO J. 4: 561-568.

Brosius, J. (1984). Plasmid vectors for the selection of promoters. Gene 27: 151-160.

Brown, B.J., and Carton, B.C. (1980). Plasmid-mediated transformation in Bacillus megaterium. J. Bact. 142: 508-512.

Burgess, R.R. (1969). Separation and characterization of the subunits of ribonucleic acid polymerase. J. Biol. Chem. 244: 6168-6176.

Burgess, R.R., Travers, A.A., Dunn, J.J., and Bautz, E.K.F. (1969). Factor stimulating transcription by RNA polymerase. Nature 221: 43-46.

Burton, Z., Burgess, R.R., Lin, J., Moore, D., Holder, S., and Gross, C.A. (1981). The nucleotide sequence of the cloned rpoD gene for the RNA polymerase sigma subunit from E.coli K12. Nucleic Acids Res. 9: 2889-2903.

Caillaud, F., and Courvalin, P. (1987). Nucleotide sequence of the ends of the conjugative shuttle transposon Tn1545. Mol. Gen. Genet. 209: 110-115.

Caillaud, F., Carlier, C., and Courvalin, P. (1987). Physical analysis of the conjugative shuttle transposon Tn1545. Plasmid 17: 58-60.

Carter, P., Bedouelle, H., and Winter, G. (1985). Improved oligonucleotide site-directed mutagenesis using M13 vectors. Nucleic Acids Res. 13: 4431-4443.

Cary, J.W., Petersen, D.J., and Bennett, G.N. (1988). Cloning and expression of Clostridium acetobutylicum phosphotrans-butyrylase and butyrate kinase in Escherichia coli. J. Bact. 170: 4613-4618.

Casadaban, M.J., and Cohen, S.N. (1980). Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. J. Mol. Biol. 138: 179-207.

Cato, E.P., George, W.L., and Finegold, S.M. (1986). Genus Clostridium. p. 1141-1200. In, P.H.A. Sneath, N.S. Mair, M.E. Sharpe, and J.G. Holt (eds.), Bergy's Manual of Systematic Bacteriology. Williams and Wilkins, Baltimore.

Cato, E.P., and Stackebrandt, E . (1989). Taxonomy and phylogeny. In, N.P. Minton and D.J. Clarke (eds.), Clostridia. Plenum Publishing Corp, New York, in press.

Chambers, S.P., Prior, S.P., Barstow, D.A., and Minton, N.P. (1988). The pMTL nic$^-$ cloning vectors. I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. Gene 68: 139-149.

Chang, S., and Cohen, S.N. (1979). High frequency transformation of Bacillus subtilis protoplasts by plasmid DNA. Mol. Gen. Genet. 168: 111-115.

Chassy, B.M., and Flickinger, J.L. (1987). Transformation of Lactobacillus casei by electroporation. FEMS Microbiol. Lett. 44: 173-177.

Chater, K.F., and Carter, A.T. (1979). A new, wide host range, temperate bacteriophage (R4) of Streptomyces and its interaction with some restriction-modification systems. J. Gen. Micro. 115: 431-442.

Chen, K.C.K., Chen, J.S., and Johnson, J.L. (1986). Structural features of multiple nifH-like sequences and very biased codon usage of nitrogenase genes of Clostridium pasteurianum. J. Bact. 166: 162-172.

Chou, P.Y., and Fasman, G.D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. 47: 45-148.

Clarke, D.J., and Morris, J.G. (1976). Butyricin 7423; a bacteriocin produced by Clostridium butyricum NCIB 7423. J. Gen. Microbiol. 95: 67-77.

Clewell, D.B., and Gawron-Burke, C. (1986). Conjugative transposons and the dissemination of antibiotic resistance in streptococci. Ann. Rev. Microbiol. 40: 635-659.

Clewell, D.B., and Helinski, D.R. (1969). Supercoiled circular DNA-protein complex in Escherichia coli: purification and induced conversion to an open circular DNA form. Proc. Natl. Acad. Sci. U.S.A. 62: 1159-1166.

Clewell, D.B., Yagi, Y., Dunny, M., and Schultz, S.K. (1974). Characterization of three plasmid deoxyribonucleic acid molecules in a strain of Streptococcus faecalis: Identification of a plasmid determining erythromycin resistance. J. Bact. 117: 283-289.

Cohen, S.N., Chang, A.C.Y., and Hsu, L. (1972). Nonchromosomal antibiotic resistance in bacteria: genetic transformation of Escherichia coli by R-factor DNA. Proc. Natl. Acad. Sci. U.S.A. 69: 2110-2114.

Collins, M.E., Oultram, J.D., and Young, M. (1985). Identification of restriction fragments from two cryptic Clostridium butyricum plasmids that promote the establishment of a replication-defective plasmid in Bacillus subtilis. J. Gen. Microbiol. 131: 2097-2105.

Contag, P.R., and Rogers, P. (1988). The cloning and expression of the Clostridium acetobutylicum B643 butyraldehyde dehydrogenase by complementation of an Escherichia coli mutant. ASM Annual Meeting Abstracts: 168. American Society for Microbiology, Washington D.C.

Cornet, P., Tronik, D., Millet, J., and Aubert, J.P. (1983). Cloning and expression in Escherichia coli of Clostridium thermocellum genes coding for amino acid synthesis and cellulose hydrolysis. FEMS Microbiol. Lett. 16: 137-141.

Courvalin, P., and Carlier, C. (1987). Tn1545: a conjugative shuttle transposon. Mol. Gen. Genet. 206: 259-264.

Cowles, P.B. (1934). A bacteriophage for Cl. tetani. J. Bact. 27: 163-164.

Cummins, C.S., and Johnson, J.L. (1971). Taxonomy of the Clostridia: wall composition and DNA homologies in Clostridium butyricum and other butyric acid producing bacteria. J. Gen. Microbiol. 67: 33-46.

Cutting, S., and Mandelstam, J. (1986). The nucleotide sequence and transcription during sporulation of the gerE gene of Bacillus subtilis. J. Gen. Microbiol. 132: 3013-3024.

Daldal, F., and Applebaum, J. (1985). Cloning and expression of Clostridium pasteurianum galactokinase gene in Escherichia coli K-12 and nucleotide sequence analysis of a region affecting the amount of enzyme. J. Mol. Biol. 186: 533-545.

Davies, A., Oultram, J.D., Pennock, A., Williams, D.R., Richards, D.F., Minton, N.P., and Young, M. (1988). Conjugal gene transfer in Clostridium acetobutylicum. p. 391-395. In, A.T. Ganeson, and J.A. Hoch (ed.), Genetics and biotechnology of Bacilli, Volume II. Academic Press Inc. New York.

Deininger, P.L. (1983). Random subcloning of sonicated DNA: application to shotgun DNA sequence analysis. Anal. Biochem. 129: 216-223.

del Solar, G., Puyet, A., and Espinosa, M. (1987). Initiation signals for the conversion of single-stranded to double-stranded DNA forms in the streptococcal plasmid pLS1. Nucleic Acids Res. 15: 5561-5580.

Doi, R.H., and Wang, L-F. (1986). Multiple procaryotic ribonucleic acid polymerase sigma factors. Microbiol. Rev. 50: 227-243.

Donnelly, C.E., and Sonenshein, A.L. (1984). Promoter-probe plasmid for Bacillus subtilis. J. Bact. 157: 965-967.

Dotto, G.P., Horiuchi, K.S., and Zinder, N.D. (1982). Initiation and termination of phage f1 plus-strand synthesis. Proc. Natl. Acad. Sci. U.S.A. 79: 7122-7126.

Drlica, K. (1984). Biology of bacterial deoxyribonucleic acid topoisomerases. Microbiol. Rev. 48: 273-289.

Dubbert, W., Luczak, H., and Staudenbauer, W.L. (1988). Cloning of two chloramphenicol acetyltransferase genes from Clostridium butyricum and their expression in Escherichia coli and Bacillus subtilis. Mol. Gen. Genet. In press.

Dubnau, D. (1976). Genetic transformation of Bacillus subtilis: a review with emphasis on the recombination mechanism, p.14-27. In, D. Schlesinger (ed.), Microbiology 1976. American Society for Microbiology, Washington. D.C.

Duncan, C.L., Rokos, E.A., Christensen, C.M., and Rood, J.I. (1978). Multiple plasmids in different toxigenic types of Clostridium perfringens: possible control of beta-toxin production, p. 246-248. In, D. Schlesinger (ed.), Microbiology 1978. American Society for Microbiology. Washington. D.C.

Efstathiou, I., Reysset, G., and Truffaut, N. (1986). A study of inulase activity in the Clostridium acetobutylicum strain ABKN8. Appl. Microbiol. Biotechnol. 25: 143-149.

Efstathiou, I., and Truffaut, N. (1986). Cloning of Clostridium acetobutylicum genes and their expression in Escherichia coli and Bacillus subtilis. Mol. Gen. Genet. 204: 317-321.

Ehrlich, S.D. (1977). Replication and expression of plasmids from Staphylococcus aureus in Bacillus subtilis. Proc. Natl. Acad. Sci. U.S.A. 74: 1680-1682.

Ehrlich, S.D. (1978). DNA cloning in Bacillus subtilis. Proc. Natl. Acad. Sci. U.S.A. 75: 1433-1436.

Eisel, U., Jarausch, W., Goretzki, K., Henschen, A., Engels, J., Weller, U., Hudel, M., Habermann, E., and Niemann, H. (1986). Tetanus toxin: primary structure, expression in E.coli, and homology with botulinum toxins. EMBO J. 5: 2495-2502.

Eklund, M.W., and Poysky, F.T. (1974). Interconversion of type C and D strains of Clostridium botulinum by specific bacteriophages. Appl. Microbiol. 27: 251-258.

Eklund, M.W., Poysky, F.T., Boatman, E.S., and Pelroy, G.S. (1974). Interspecies conversion of Clostridium botulinum type C to Clostridium novyi type A by bacteriophage. Science 186: 456-458.

Eklund, M.W., Poysky, F.T. and Reed, S.M. (1972). Bacteriophage and the toxigenicity of Clostridium botulinum type D. Nature New Biol. 235: 16-17.

Eklund, M.W., Poysky, F.T., Reed, S.M., and Smith, C.A. (1971). Bacteriophage and the toxigenicity of Clostridium botulinum type C. Science 172: 480-482.

Elton, R.A. (1983). The relationship of DNA base composition and individual protein composition in micro-organisms. J. Mol. Evol. 2: 263-276.

Engel, H.W.B., Soedirman, N., Rost, J.A., van Leeuwen, W.J., and van Embden, D.A. (1980). Transferability of macrolide, lincomycin, and streptogramin resistances between group A, B, and D streptococci, Streptococcus pneumoniae, and Staphylococcus aureus. J. Bact. 142: 407-413.

Espinosa, M., Lopez, P., Urena, P., and Lacks, S.A. (1982). Interspecific plasmid transfer between Streptococcus pneumoniae and Bacillus subtilis. Mol. Gen. Genet. 188: 195-201.

Evans, R.P., and Macrina, F.L. (1983). Streptococcal R plasmid pIP501: endonuclease site map, resistance determinant location and construction of novel derivatives. J. Bact. 154: 1347-1355.

Fairweather, N.F., and Lyness, V.A. (1986). The complete nucleotide sequence of tetanus toxin. Nucleic Acids Res. 14: 7809-7812.

Fairweather, N.F., Lyness, V.A., Pickard, D.J., Allen, G., and Thomson, R.O. (1986). Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in Escherichia coli. J. Bact. 165: 21-27.

Faure, E., Bagnara, C., Belaich, A., and Belaich, J.P. (1988). Cloning and expression of two cellulase genes of Clostridium cellulolyticum in Escherichia coli. Gene 65: 51-58.

Finn, C. W. Jr., Silver, R.P., Habig, W.H., Hardegree, M.C., Zon, G., and Garon, C.F. (1984). The structural gene for tetanus neurotoxin is on a plasmid. Science 224: 881-884.

Fitzgerald, G.F., and Clewell, D.B. (1985). A conjugative transposon (Tn919) in Streptococcus sanguis. J. Bact. 47: 415-420.

Fodor, K., and Alfoldi, L. (1976). Fusion of protoplasts of Bacillus megaterium. Proc. Natl. Acad. Sci. U.S.A. 73: 2147-2150.

Fornari, G.S., and Kaplan, S. (1982). Genetic transformation of Rhodopseudomonas sphaeroides by plasmid DNA. J. Bact. 152: 89-97.

Fox, T.D., and Weis-Brummer, B. (1980). Leaky +1 amd -1 frameshift mutations at the same site in a yeast mitochondrial gene. Nature 288: 60-63.

Franke, A.E., and Clewell, D.B. (1981). Evidence for a chromosome-borne resistance transposon (Tn916) in Streptococcus faecalis that is capable of "conjugal"transfer in the absence of a conjugative plasmid. J. Bacteriol. 145: 494-502.

Franklin, F.C.H., Bagdasarian, M., Bagdasarian, M.M., and Timmis, K.N. (1981). Molecular and functional analysis of the TOL plasmid pWWO from Pseudomonas putida and cloning of genes for the entire regulated aromatic ring meta cleavage pathway. Proc. Natl. Acad. Sci. U.S.A. 78: 7458-7462.

Fromm, M., Taylor, L.P., and Walbot, V. (1985). Expression of genes transferred into monocot and dicot plant cells by electroporation. Proc. Natl. Acad. Sci. U.S.A. 82: 5824-5828.

Gait, M.J., Matthes, H.W.D., Singh, M., Sproat, B.S., and Titmas, R.C. (1982). Rapid synthesis of oligodeoxyribonucleotides VII. Solid phase synthesis of oligodeoxyribonucleotides by a continuous flow phosphotriester method on a kieselguhr-polyamide support. Nucleic Acids Res. 10: 6243-6254.

Garnier, T. and Cole, S.T. (1986). Characterization of a bacteriocinogenic plasmid from Clostridium perfringens and molecular genetic analysis of the bacteriocin-encoding gene. J. Bact. 168: 1189-1196.

Garnier, T., and Cole, S.T. (1988a). Complete nucleotide sequence and genetic organization of the bacteriocinogenic plasmid, pIP404, from Clostridium perfringens. Plasmid 19: 134-150.

Garnier, T., and Cole, S.T. (1988b). Identification and molecular genetic analysis of replication functions of the bacteriocinogenic plasmid pIP404 from Clostridium perfringens. Plasmid 19: 151-160.

Garnier, T., and Cole, S.T. (1988c). Studies of UV-inducible promoters from Clostridium perfringens in vivo and in vitro. Mol. Microbiol. 2: 607-614.

Garnier, T., Saurin, W., and Cole, S.T. (1987). Molecular characterization of the resolvase gene, res, carried by a multicopy plasmid from Clostridium perfringens: common evolutionary origin for prokaryotic site-specific recombinases. Mol. Microbiol. 1: 371-376.

Gasson, M.J., and Davies, F.L. (1980). Conjugal transfer of the drug resistance plasmid pAMβ1 in the lactic streptococci. FEMS Microbiol. Lett. 7: 51-53.

Gawron-Burke, C., and Clewell, D.B. (1984). Regeneration of insertionally inactivated streptococcal DNA fragments after excision of transposition of transposon Tn916 in Escherichia coli: strategy for targeting and cloning of genes from Gram-positive bacteria. J. Bact. 159: 214-221.

Geis, A. (1982). Transfection of protoplasts of Streptococcus lactis subsp. diacetylactis. FEMS Microbiol. Lett. 15: 119-122.

George, H.A., and Chen, J.S. (1983). Aciditory conditions are not obligatory for onset of butanol formation by Clostridium beijerinkii (synonym, C. butylicum). Appl. Environ. Microbiol. 46: 321-327.

Gibbs, D.F. (1983). The rise and fall (...and rise?) of the acetone-butanol fermentation. Trends in Biotechnol. 1: 12-15.

Gibson, E.M., Chace, N.M., London, S.B., and London, J. (1979). Transfer of plasmid-mediated antibiotic resistance from Streptococci to Lactobacilli. J. Bact. 137: 614-619.

Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B.S., and Stormo, G. (1981). Translation initiation in prokaryotes. Ann. Rev. Microbiol. 35: 365-403.

Gottschalk, G., Andreesen, J.R., and Hippe, H. (1981). The genus Clostridium (non-medical aspects), Chapter 138. In, M.P. Starr, H. Stolp, H.G. Truper, A. Balows, and H.G. Schlegel (eds.), The prokaryotes II. Springer Verlag.

Gottwald, M., Hippe, H., and Gottschalk, G. (1984). Formation of n-butanol from D-glucose by strains of Clostridium tetanomorphum group. Appl. Environ. Microbiol. 48: 573-576.

Gotz, F., Ahrne, S., and Lindberg, M. (1981). Plasmid transfer and genetic recombination by protoplast fusion in staphylococci. J. Bact. 145: 74-81.

Gouy, M., and Gautier, C. (1982). Codon usage in bacteria: correlation with expressivity. Nucleic Acids Res. 10: 7055-7074.

Goze, A. and Ehrlich, S.D. (1980). Replication of plasmids from Staphylococcus aureus in Escherichia coli. Proc. Natl. Acad. Sci. U.S.A. 77: 7333-7337.

Grabnitz, F., and Staudenbauer, W.L. (1988). Characterization of two alpha-glucosidase genes from Clostridium thermocellum. Biotechnol. Lett. 10: 73-78.

Graves, M.C., Mullenbach, G.T., and Rabinowitz, J.C. (1985). Cloning and nucleotide sequence determination of the Clostridium pasteurianum ferredoxin gene. Proc. Natl. Acad. Sci. U.S.A. 82: 1653-1657.

Graves, M.C., and Rabinowitz, J.C. (1986). In vivo and in vitro transcription of the Clostridium pasteurianum ferredoxin gene. J. Biol. Chem. 261: 11409-11415.

Gren, E.J. (1984). Recognition of messenger RNA during translation initiation in Escherichia coli. Biochimie 66: 1-29.

Grepinet, O., and Beguin, P. (1986). Sequence of the cellulase gene of Clostridium thermocellum coding for endoglucanase B. (1986). Nucleic Acids Res. 14: 1791-1799.

Gros, M.F., te Riele, H., and Ehrlich, S.D. (1987). Rolling circle replication of single-stranded DNA plasmid pC194. EMBO J. 6: 3863-3869.

Grosjean, H., and Fiers, W. (1982). Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. Gene 18: 199-209.

Grosschedl, R., and Schwartz, E. (1979). Nucleotide sequence of the cro-cII-oop region of bacteriophage 434 DNA. Nucleic Acids Res. 6: 867-881.

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

Gruss, D.A., Ross, H.F., and Novick, R.P. (1987). Functional analysis of a palindromic sequence required for normal replication of several staphylococcal plasmids. Proc. Natl. Acad. Sci. U.S.A. 84: 2165-2169.

Gruss, D.A., and Ehrlich, S.D. (1989). The family of highly inter-related single-stranded DNA plasmids. Microbiol. Rev. In press.

Gryczan, T., Shivakumar, A.G., and Dubnau, D. (1980). Characterization of chimaeric plasmid cloning vehicles in Bacillus subtilis. J. Bact: 246-253.

Gunsalus, R.P., and Yanofsky, C. (1980). Nucleotide sequence and expression of E. coli trpR: the structural gene for the trp aporepressor. Pcoc. Natl. Acad. Sci. U.S.A. 77: 7118-7121.

Hachler, H., Berger-Bachi, B., and Kayser, F.H. (1987a). Genetic characterization of a Clostridium difficle erythromycin-clindamycin resistance determinant that is transferable to Staphylococcus aureus. Antimicrob. Agents and Chemother. 31: 1039-1045.

Hachler, H., Kayser, F.H., and Berger-Bachi, B. (1987b).Homology of a transferable tetracycline resistance determinant of Clostridium difficile with Streptococcus (Enterococcus) faecalis transposon Tn916. Antimicrob. Agents and Chemother. 31: 1033-1038.

Hager, P.W., and Rabinowitz, J.C. (1985a). Translational specificity in Bacillus subtilis. p.1-29. In, D. Dubnau (ed.), The molecular biology of the bacilli, Vol II. Academic Press, New York.

Hager, P.W., and Rabinowitz, J.C. (1985b). Inefficient translation of T7 late mRNA by Bacillus subtilis ribosomes. J. Biol. Chem. 260: 15163-15167.

Hall, J., Hazlewood, G.P., Barker, P.J., and Gilbert, H.J. (1988). Conserved reiterated domains in Clostridium thermocellum endoglucanases are not nessential for catalytic activity. Gene 69: 29-38.

Hall, M.N., Gabay, J., Debarbouille, M., and Schwartz, M. (1982). A role for mRNA secondary structure in the control of translation initiation. Nature 295: 616-618.

Hambleton, P., Shone, C.C., and Melling, J. (1987). Botulinum toxin: structure, action, and clinical use, p. 233-260. In, P. Jenner (ed.), Neurotoxins and their pharmocological application. Raven Press, New York.

Harlander, S.K. (1987). Transformation of Streptococcus lactis by electroporation, p. 229-233. In, J.J. Ferretti, and R. Curtis (eds.), Streptococcal Genetics. American Society for Microbiology, Washington, D.C.

Hashimoto, H., Morikawa, H., Yamada, Y., and Kimura, A. (1985). A novel method for transformation of intact yeast cells by electroinjection of plasmid DNA. Appl. Microbiol. Biotechnol. 21: 336-339.

Hawley, D.K., and McClure, W.R. (1983). Compilation and analysis of Escherichia coli promoter DNA sequences. Nucleic Acids Res. 11: 2237-2255.

Hayter, P.M., and Dale, J.W. (1984). Detection of plasmids in clinical isolates of Clostridium difficile. Microbios Lett. 27: 151-156.

Hazlewood, G.P., Romaniec, M.P.M., Davidson, K., Grepinet, O., Beguin, P., Millet, J., Raynaud, O., and Aubert, J.P. (1988). A catalogue of Clostridium thermocellum endoglucanase, alpha-glucosidase, and xylanase genes cloned in Escherichia coli. FEMS Microbiol. Lett. 51: 231-236.

Heefner, D.L., Squires, C.H., Evans, R.J., Kopp, B.J., and Yarus, M.J. (1984). Transformation of Clostridium perfringens. J. Bact. 159: 460-464.

Hinton, S.M., and Freyer, G. (1986). Cloning, expression and sequencing the molybdenum-pterin binding protein (mop) gene of Clostridium pasteurianum in Escherichia coli. Nucleic Acids Res. 14: 9371-9380.

Holmes, D.S., and Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. Analytical Biochem. 114: 193-197.

Holland, S.K., Cutting, S., and Mandelstam, J. (1987). The possible DNA-binding nature of the regulatory proteins, encoded by spoIID and gerE, involved in the sporulation of Bacillus subtilis. J. Gen. Microbiol. (1987). 133: 2381-2391.

Holm, L. (1986). Codon usage and gene expression. Nucleic Acids Res. 14: 3075-3087.

Horii, T., Ogawa, T., and Ogawa, H. (1981). Nucleotide sequence of the lexA gene of E.coli. Cell 44: 925-933.

Horinouchi, S. and Weisblum, B. (1982a). Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. J. Bact. 150: 815-825.

Horinouchi, S., and Weisblum, B. (1982b). Nucleotide sequence and functional map of pE194, a plasmid that specifies inducible resistance to macrolide, lincosamide, and streptogramin type B antibiotics. J. Bact. 150: 804-814.

Horinouchi, S., Woo-Hyeon, B., and Weisblum, B. (1983) A complex attenuator regulates inducible resistance to macrolides, lincosamides, and streptogramin type antibiotics in Streptococcus sanguis. J. Bact. 154: 1252-1262.

Hotchkiss R.D., and Gabor, M.H. (1985). Protoplast fusion in Bacillus and its consequences, p. 109-149. In, D.A. Dubnau (ed.), The molecular biology of the bacilli, volume II. Academic Press, New York.

Hsiang, M.W., Cole, R.D., Takeda, Y., and Echols, H. (1977). The amino acid sequence of cro regulatory protein from bacteriophage lambda. Nature 270: 275-277.

Hu, N., and Messing, J. (1982). The making of strand-specific M13 probes. Gene 17: 271-277.

Hyun, H.H., and Zeikus, J.G. (1985). General biochemical characterization of the thermostable extracellular beta-amylase from Clostridium thermosulphurogenes. Appl. Environ. Microbiol. 49: 1162-1167.

Iemura, T. (1981). Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes. J. Mol. Biol. 146: 1-21.

Imanaka, T., Tanaka,T., Tsunekawa, H., and Aiba, S. (1981). Cloning of the genes for penicillinase, penP and penI, of Bacillus licheniformis in some vector plasmids and their expression in Escherichia coli. Bacillus subtilis. and Bacillus licheniformis. J. Bact. 147: 776-786.

Imanaka, T., Fujii, M., Aramori, I., and Aiba, S. (1982). Transformation of Bacillus stearothermophilus with plasmid DNA and characterization of shuttle vector plasmids between Bacillus stearothermophilus and Bacillus subtilis. J. Bact. 149: 824-830.

Inouye, S., Nakazawa, A., and Nakazawa, T. (1986). Expression of the regulatory gene xylS controlled by the xylR gene product. Proc. Natl. Acad. Sci. U.S.A. 84: 5182-5186.

Ionesco, H. (1980). Transfert de la resistance a la tetracycline chez Clostridium difficile. Ann. Inst. Pasteur Paris 131A: 171-179.

Ionesco, H., and Bauanchaud, D.H. (1973). Production de bacteriocine a la presence d'un plasmide chez Clostridium perfringens type A. C.R. Acad. Sci. (Paris) Serie D. 276: 2855-2857.

Ionesco, H., Wolff, A., and Sebald, M. (1974). The induced production of bacteriocin and bacteriophage by BP6K-N strain of Clostridium perfringens. Ann. Microbiol. (Inst. Pasteur) 125B: 335-346.

Iordanescu, S., Surdeanu, M., Dell Latta, P., and Novick, R. (1978). Incompatability and molecular relationships between small staphylococcal plasmids carrying the same resistance marker. Plasmid 1: 468-479.

Isarentant, D., and Fiers, W. (1980). Secondary structure of mRNA and efficiency of translation initiation. Gene 9: 1-12.

Ishii, K., Kudo, T., Honda, H., and Horikoshi, H. (1983). Molecular cloning of beta-isopropylmalate dehydrogenase gene from Clostridium butyricum M588. Agric. Biol. Chem. 47: 2313-2317.

Janniere, L., and Ehrlich, S.D. (1987). Recombination between short repeated sequences is more frequent in plasmids than in the chromosome of Bacillus subtilis. Mol. Gen. Genet. 210: 116-121.

Janssen, P.J., Jones, W.A., Jones, D.T., and Woods, D.R. (1988). Molecular analysis and regulation of the glnA gene of the Gram-positive anaerobe Clostridium acetobutylicum. J. Bact. 170: 400-408.

Joliff, G., Beguin, P., and Aubert, J.P. (1986). Nucleotide sequence of the cellulase gene celD encoding endoglucanase D of Clostridium thermocellum. Nucleic Acids Res. 14: 8605-8613.

Jones, D.T., Jones, W.A., and Woods, D.R. (1985). Production of recombinants after protoplast fusion in Clostridium acetobutylicum P262. J. Gen. Microbiol. 131: 1213-1216.

Jones, D.T., van der Westhuizen, A., Long, S., Allcock, E.R., Reid, S.J., and Woods, D.R. (1982). Solvent production and morphological changes in Clostridium acetobutylicum. Appl. Environ. Microbiol. 43: 1434-1439.

Jones, D.T., and Woods, D.R. (1986). Acetone-butanol fermentation revisited. Microbiol. Rev. 50: 484-524.

Kaczorek, M., Zettlmeissl, G., del Peyroux, F., and Streeck, R.E. (1985). Diptheria toxin promoter function in Corynebacterium diptheriae and Escherichia coli. Nucleic Acids Res. 13: 147-159.

Kadam, S., Demain, A.L., Millet, J., Beguin, P., and Aubert, J.P. (1988). Molecular cloning of a gene for a thermostable alpha-glucosidase from Clostridium thermocellum into Escherichia coli. Enzyme Microb. Technol. 10: 9-13.

Karube, I., Matsunga, T., Otsuka, T., Kayano, H., and Suzuki, S. (1981). Hydrogen evolution by co-immobilised chloroplasts and Clostridium butyricum. Biochim. Biophys. Acta. 673: 490-495.

Karube, I., Urano, N., Yamada, T., Hirochika, H., and Sakaguchi, K. (1983). Cloning and expression of the hydrogenase gene from Clostridium butyricum in Escherichia coli. FEBS Lett. 158: 119-122.

Kastelein, R.A., Berkhout, B., Overbeek, G.P., and van Quin, J. (1983). Effect of the sequences upstream from the RBS on the yield of protein from the cloned gene for phage MS2 coat protein. Gene 23: 245-254.

Kastelein, R.A., Remaut, E., Fiers, W., and van Duin, J. (1982). Lysis gene expression of RNA phage MS2 depends on a frameshift during translation of the overlapping coat protein gene. Nature 295: 35-41.

Kazutoshi, I., Nishida, T., and Izaki, K. (1988). Application of electroporation for transformation in Erwinia carotovora. Agric. Biol. Chem. 52: 293-294.

Khan, S.A., Adler, G.K., and Novick, R.P. (1982). Functional origin of replication of pT181 plasmid DNA is contained within a 168-base-pair segment. Proc. Natl. Acad. Sci. U.S.A. 79: 4580-4584.

Khan, S.A. and Novick, R.P. (1983). Complete nucleotide sequence of pT181, a tetracycline-resistance plasmid from Staphylococcus aureus. Plasmid 10: 251-259.

Knight, D.E., and Scrutton, M.C. (1986). Gaining access to the cytosol: the technique and some applications of electropermeabilization. Biochem. J. 234: 497-506.

Knowlton, S., Ferchak, J.D., and Alexander, J.K. (1984). Protoplast regeneration in Clostridium tertium: isolation of derivatives with high frequency regeneration. Appl. Environ. Microbiol. 48: 1246-1247.

Koepsel, S.A., Murray, R.W., Rosenblum, W.D., and Khan, S.A. (1985). The replication initiator protein of plasmid pT181 has sequence-specific endonuclease and topoisomerase-like activities. Proc. Natl. Acad. Sci. U.S.A. 82: 6845-6849.

Koepsel, R.R., Murray, R.W., and Khan, S.A. (1986). Sequence-specific interaction between the replication initiator protein of plasmid pT181 and its origin of replication. Proc. Natl. Acad. Sci. U.S.A. 83: 5484-5488.

Koepsel, R.R. and Khan, S.A. (1987). Cleavage of single-stranded DNA by plasmid pT181-encoded RepC protein. Nucleic Acids Res. 15: 4085-4097.

Kok, J., van der Vossen, J.M.B., and Venema, G. (1984). Construction of plasmid cloning vectors for lactic streptococci which also replicate in Bacillus subtilis and Escherichia coli. Appl. Environ. Microbiol. 43: 726-731.

Kondo, J.K., and McKay, L.L. (1982). Transformation of Streptococcus lactis protoplasts by plasmid DNA. Appl. Environ. Microbiol. 43: 1213-1215.

Konisberg, W., and Godson, N. (1983). Evidence for use of rare codons in the dnaG gene and other regulatory genes of Escherichia coli. Proc. Natl. Acad. Sci. U.S.A. 80: 687-691.

Koths, K. and Dressler, D. (1978). Analysis of the ØX174 replication cycle by electron microscopy. Proc. Natl. Acad. Sci. U.S.A. 75: 605-609.

Kreft, J., Bernhard, K., and Goebel, W. (1978). Recombinant plasmids capable of replication in Bacillus subtilis. Mol. Gen. Genet. 162: 59-67.

Kyte, J., and Doolittle, F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157: 105-132.

Lacks, S.A., Lopez, P., Greenberg, B., and Espinosa, M. (1986). Identification and analysis of genes for tetracycline resistance and replication functions in the broad-host-range plasmid pLS1. J. Mol. Biol. 192: 753-765.

Laird, W.J., Aaronson, W., Silver, R.P., Habig, W.H., and Hardegree, M.C. (1980). Plasmid associated toxigenicity in Clostridium tetani. J. Infect. Dis. 142: 623.

Lampson, B.C. and Parisi, J.T. (1986). Nucleotide sequence of the constitutive macrolide-lincosamide-streptogramin B resistance plasmid pNE131 from Staphylococcus epidermidis and homologies with Staphylococcus aureus plasmids pE194 and pSN2. J. Bact. 167: 888-892.

Landman, O.E., and Pepin R.A. (1982). Bacillus subtilis mating system: dicaryotic recombinants and suppression of the prototrophic phenotype, p. 25-40. In, A.T. Ganesan, S. Chang, and J.A Hoch (eds.), Molecular cloning and gene regulation in Bacilli. Academic Press, Inc., New York.

Langeveld, S.A., van Mansfeld, A.D.M., Baas, P.D., Jansz, H.S., van Arkel, G.A., and Weisbeek, P.J. (1978). Nucleotide sequence of the origin of replication in bacteriophage ØX174 RF DNA. Nature 271: 417-419.

Laskey, R.A., and Mills, A.D. (1977). Enhanced autoradiographic detection of $^{32}P$ and $^{125}I$ using intensifying screens and hypersensitised films. FEBS Lett. 82: 314-316.

Lathe, R., Kieny, M.P., Skory, S., and Lecocq, J.P. Linker tailing: unphosphorylated linker oligonucleotides for joining DNA termini. DNA. 3: 173-182.

Leblanc, D.J., Hawley, R.J., Lee, N., and St. Martin, E.J. (1978). "Conjugal" transfer of plasmid DNA among oral streptococci. Proc. Natl. Acad. Sci. U.S.A. 75: 3484-3487.

Leblanc, D.J., and Lee, L.N. (1984). Physical and genetic analysis of streptococcal plasmid pAMβ1 and cloning of its replication region. J. Bact. 157: 445-453.

Lee, C.K., Durre, P., Hippe, H., and Gottschalk, G. (1987). Screening for plasmids in the genus Clostridium. Arch. Microbiol. 148: 107-114.

Lee, G., Talkington, C., and Pero, J. (1980). Nucleotide sequence of a promoter recognized by Bacillus subtilis RNA polymerase. Mol. Gen. Genet. 180: 57-65.

Lee, S.F. and Forsberg, C.W. (1987). Isolation and some properties of a beta-D-xylosidase from Clostridium acetobutylicum ATCC 824. Appl. Environ. Microbiol. 53: 651-654.

Lee-Wickner, L. Ja., and Chassy, B.M. (1984). Production and regeneration of Lactobacillus casei protoplasts. Appl. Environ. Microbiol. 48: 994-1000.

Legault-Demare, L., and Chambliss, G.H. (1975). Selective messenger translation of Bacillus subtilis ribosomes. Mol. Gen. Genet. 142: 277-287.

Leonhardt, H., and Alonso, J.C. (1988). Construction of a shuttle vector for inducible gene expression in Escherichia coli and Bacillus subtilis. J. Gen. Microbiol. 134: 605-609.

Lereclus, D., Menou, G., and Lecadet, M.M. (1983). Isolation of a DNA sequence related to several plasmids from Bacillus thuringensis after a mating involving the Streptococcus faecalis plasmid pAMβ1. Mol. Gen. Genet. 191: 307-313.

Li, A.W., Krell, P.J., and Mahony, D.E. (1980). Plasmid detection in a bacteriocinogenic strain of Clostridium perfringens. Can. J. Microbiol. 26: 1018-1022.

Lideman, L.F. (1983). Factors limiting the expression of Escherichia coli genes in cells of Bacilli. Genetika 19: 693-699.

Lin, J.H.C., and Savage, D.C. (1986). Genetic transformation of rifampicin resistance in Lactobacillus acidophilus. J. Gen. Microbiol. 132: 2107-2111.

Lin, Y., L., and Blaschek, H.P. (1984). Transformation of heat-treated Clostridium acetobutylicum protoplasts with pUB110 plasmid DNA. Appl. Environ. Microbiol. 48: 737-742.

Long , S., Jones, D.T., and Woods, D.R. (1984). The relationship between sporulation and solvent production in Clostridium acetobutylicum P262. Biotechnol. Lett. 6: 529-534.

Lopez-Diaz, I., Clarke, S., and Mandelstam, J. (1986). spoIID operon of Bacillus subtilis: cloning and sequence. J. Gen. Microbiol. 132: 341-354.

Losick, R., and Pero, J. (1981). Cascades of sigma factors. Cell 25: 582-584.

Lovell, C.R., Przybyla, A., and Ljungdahl, L.G. (1988). Cloning and expression in Escherichia coli of the Clostridium thermoaceticum gene encoding thermostable formyl-tetrahydrofolate synthetase. Arch. Microbiol. 149: 280-285.

Luczak, H., Schwarzmoser, H., and Staudenbauer, W.L. (1985). Construction of Clostridium butyricum hybrid plasmids and transfer to Bacillus subtilis. Appl. Microbiol. Biotechnol. 23: 114-122.

Macrina, F.L., Jones, K.E., and Wood, P.H. (1980). Chimeric streptococcal plasmids and their use as molecular cloning vehicles in Streptococcus sanguis (Challis). J. Bact. 143: 1425-1435.

Magot, M. (1983). Transfer of antibiotic resistances from Clostridium innocuum to Clostridium perfringens in the absence of detectable DNA. FEMS Microbiol.Lett. 18: 149-151.

Magot, M. (1984). Physical characterization of the Clostridium perfringens tetracycline-chloramphenicol resistance plasmid pIP401. Ann. Microbiol. (Inst. Pasteur) 135B: 269-282.

Makaroff, C.A., Zalkin, H., Switzer, R.L., and Vollmer, S.J. (1983). Cloning of the Bacillus subtilis glutamine phosphoribosylpyrophosphate amidotransferase gene in Escherichia coli. J. Biol. Chem. 258: 10586-10593.

Manch-Citron, J.N., Gennaro, M.L., Majumder, S., and Novick, R.P. (1986). RepC is rate limiting for pT181 plasmid replication. Plasmid 16: 108-115.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). Molecular cloning. A laboratory manual. Cold Spring Harbour, New York.

Marmur, J. (1961). A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J. Mol. Biol. 3: 208-218.

Martin, P.A.W., Lohr, J.R., and Dean, D.H. (1981). Transformation of Bacillus thuringensis protoplasts by plasmid deoxyribonucleic acid. J. Bact. 145. 980-983.

Maxam, A.M., and Gilbert, W. (1977). A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. 74: 560-564.

McCoy, E., Fred, E.B., Paterson, W.H., and Hastings, E.G. (1926). A cultural study of the acetone butyl alcohol organism. J. Infect. Dis. 39: 457-485.

McCutchan, W.N.,and Hickey, R.J. (1954). The butanol-acetone fermentations. Ind. Ferment. 1: 347-388.

McDonnell, M.W., Simon, M.N., and Studier, F.W. (1977). Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J. Mol. Biol. 110: 119-146.

McKenzie, T., Hoshino, T., Tanaka, T., and Sueoka, N. (1986). The nucleotide sequence of pUB110: some salient features in relation to replication and its regulation. Plasmid 15: 93-103.

McKenzie, T, Hoshino, T., Tanaka, T., and Sueoka, N. (1987). A revision of the nucleotide sequence and functional map of pUB110. Plasmid 17: 83-85.

McLaughlin, J.R., Murray, C.L., and Rabinowitz, J.C. (1981a). Plasmid-directed expression of Staphylococcus aureus beta-Lactamase by Bacillus subtilis in vitro. J. Biol. Chem. 256: 11273-11282.

McLaughlin, J.R., Murray, C.L., and Rabinowitz, J.C. (1981b). Unique features in the ribosome binding site of the Gram-positive Staphyloccocus aureus beta-Lactamase gene. J. Biol. Chem. 256: 11283-11291.

McLaughlin, J.R., Murray, C.L., and Rabinowitz, J.C. (1981c). Initiation factor-independent translation of mRNAs from Gram-positive bacteria. Proc. Natl. Acad. Sci. U.S.A. 78: 49-4916.

Mejean, V., Claverys, J.P., Vasseghi, H., and Sicard, A.M. (1981). Rapid cloning of specific DNA fragments of Streptococcus pneumoniae by vector integration into the chromosome followed by endonucleolytic excision. Gene 15: 289-293.

Mercenier, A., Robert, C., Romero, D.A., Castellino, I., Slos, P., and Lemoine, Y. (1988). Development of an efficient spheroplast transformation procedure for Streptococcus thermophilus: the use of transfection to define a regeneration medium. Biochimie 70: 567-577.

Messing, J., Gronenborn, B., Muller-Hill, B., and Hofschneider, P.H. (1977). Filamentous coliphage M13 as a cloning vehicle: insertion of a HindIII fragment of the lac regulatory region in the M13 replicative form in vitro. Proc. Natl. Acad. Sci. U.S.A. 74: 3642-3646.

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

Meyer, T.F., Geider, K., Kurz, C., and Schaller, K. (1979). Cleavage site of bacteriophage fd gene II protein in the origin of viral strand replication. Nature 278: 365-367.

Michel, B., and Ehrlich, S.D. (1986a). Illegitimate recombination occurs between the replication origin of the plasmid pC194 and a progressing replication fork. EMBO J. 5: 3691-3696.

Michel, B., and Ehrlich, S.D. (1986b). Illegitimate recombination at the replication origin of bacteriophage M13. Proc. Natl. Acad. Sci. U.S.A. 83: 3386-3390.

Mihlec, V.A., Duncan, C.L., and Chambliss, G.H. (1978). Characterization of a bacteriocinogenic plasmid in Clostridium perfringens CW55. Antimicrob. Agents and Chemother. 14: 771-779.

Millet, J., Petre, D., Beguin, P., Raynaud, O., and Aubert, J.P. (1985). Cloning of ten distinct fragments of Clostridium thermocellum coding for cellulases. FEMS Microbiol. Lett. 29: 145-149.

Minton, N.P., and Clarke, L.E. (1985). Identification of the promoter of the Pseudomonas gene coding for carboxypeptidase G2. J. Mol. Appl. Genet. 3: 26-35.

Minton, N. and Morris, G. (1981). Isolation and partial characterization of three cryptic plasmids from strains of Clostridium butyricum. J. Gen. Microbiol. 127: 325-331.

Minton, N.P., and Morris, J.G. (1983). Regeneration of protoplasts of Clostridium pasteurianum ATCC 6013. J. Bact. 155: 432-434.

Minton, N.P., Oultram, J.D., Brehm, J.K., and Atkinson, T. (1988). The replication proteins of plasmids pE194 and pLS1 have N-terminal homology. Nucleic Acids. Res. 16: 3101.

Minton, N.P., and Thompson, D.E. (1989). Genetics of anaerobes. In, B.I. Duerden and B.S. Drasser, (eds.), Anaerobes in human disease. Edward Arnold LTD, London, in press.

Monod, M., Denoya, C., and Dubnau, D. (1986). Sequence and properties of pIM13, a macrolide-lincosamide-streptogramin B resistance plasmid from Bacillus subtilis. J. Bact. 167: 138-147.

Moran, C.P.Jr., Lang, N., LeGrice, S.F.J., Lee, G., Stephens, M., Sonenshein, A.L., Pero, J., and Losick, R. (1982). Nucleotide sequences that signal the initiation of transcription and translation in Bacillus subtilis. Mol. Gen. Genet. 186: 339-346.

Morelli, L., Cocconcelli, P.S., Bottazzi, V., Damiani, G., Ferretti, L., and Sgaramella, V. (1987). Lactobacillus protoplast transformation. Plasmid 17: 73-75.

Morris, J.G. (1989). Bioconversions. In, N.P. Minton and D.J. Clarke (eds.), Clostridia. Plenum Publishing Corp., New York, in press.

Muldrow, L.L., Archibold, E.R., Nunez-Montiel, O.L., and Sheehy, R.J. (1982). Survey of the extrachromosomal gene pool of Clostridium difficile. J. Clin. Microbiol. 16: 637-640.

Muller, R.E., Ano, T., Imanaka, T., and Aiba, S. (1986). Complete nucleotide sequences of Bacillus plasmids pUB110dB, pRBH1 and its low copy number mutants. Mol. Gen. Genet. 202: 169-171.

Murai, M., Miyashita, H., Araki, H., Seki, T., and Oshima, Y. (1987). Molecular structure of the replication origin of a Bacillus amyloliquefaciens plasmid pFTB14. Mol. Gen. Genet. 210: 92-100.

Murray, C.L., and Rabinowitz, J.C. (1982). Nucleotide sequences of transcription and translation initiation regions in Bacillus phage Ø29 early genes. J. Biol. Chem. 257: 1053-1062.

Nakahama, K., Miyazaki, T., and Kikuchi, M. (1986). A promoter for B. subtilis expression vector. Gene 36: 179-182.

Nakai, C., Kagamiyama, H., Nozaki, M., Nakazawa, T., Inouye, S., Ebina, Y., and Nakazawa, A. (1983). Complete nucleotide sequence of the metapyrocatechase gene on the TOL plasmid of Pseudomonas putida mt-2. J. Biol. Chem. 258: 2923-2928.

Nakamura, S., Kimura, I., Yamakawa, K., and Nishida, S. (1983). Taxanomic relationships among Clostridium novyi types A and B, Clostridium haemolyticum and Clostridium botulinum type C. J. Gen. Microbiol. 129: 1473-1479.

Noguchi, N., Aoki, T., Sasatsu, M., Kono, M., Shishido, K., and Ando, T. (1986). Determination of the complete nucleotide sequence of pNS1, a staphylococcal tetracycline-resistance plasmid propagated in Bacillus subtilis. FEMS Microbiol. Lett. 37: 283-288.

Novick, R.P. (1987). Plasmid incompatibility. Microbiol. Rev. 51: 381-395.

Novick, R.P., Adler, G.K., Projan, S.J., Carleton, S., Highlander, S., Gruss, A., Khan, S.A., and Iordanescu, S. (1984a). Control of pT181 replication, I. The pT181 copy control function acts by inhibiting the synthesis of a replication protein. EMBO J. 3: 2399-2405.

Novick, R.P., Projan, S.J., Rosenblum, W., and Edelman, I. (1984b). Staphylococcal plasmid cointegrates are formed by host-and phage-mediated rec systems that act on short regions of homology. Mol. Gen. Genet. 195: 374-377.

Nozaki, M. (1970). Metapyrochatechase (Pseudomonas). Methods Enzymol. 7A: 522-525.

O'Brien, R.W., and Morris, J.G. (1971). Oxygen and the growth and metabolism of Clostridium acetobutylicum. J. Gen. Microbiol. 68: 307-318.

O'Farrell, P. (1981). Replacement synthesis method of labelling DNA fragments. Bethesda Research Labs. Focus 3: 1.

Ogasawara, N. (1985). Markedly unbiased codon usage in Bacillus subtilis. Gene 40: 145-150.

Ogata, S., and Hongo, M. (1979). Bacteriophages of the genus Clostridium. Adv. Appl. Microbiol. 25: 241-273.

Ogata, S.K., Choi, H., Yoshino, S., and Hayashida, S. (1981). Studies on sucrose-induced autolysis of clostridial cells, part 3. Autolytic formation of clostridial protoplasts (autoplasts) as revealed by microscopy and some of their properties. J. Fac. Agric. Kyushu Univ. 25: 201-222.

Orzech, K.A., and Burke, W.F. Jr. (1984). Conjugal transfer of pAMβ1 in Bacillus sphaericus 1593. FEMS Microbiol. Lett. 25: 91-95.

Oultram, J.D. (1986). Genetic studies with species of Clostridium. PhD dissertation, U.C.W. Aberystwyth.

Oultram, J.D., and Young, M. (1985). Conjugal transfer of plasmid pAMβ1 from Streptococcus lactis and Bacillus subtilis to Clostridium acetobutylicum. FEMS Microbiol. Lett. 27: 129-134.

Oultram, J.D., Davies, A., and Young, M. (1987). Conjugal transfer of a small plasmid from Bacillus subtilis to Clostridium acetobutylicum by cointegrate formation with plasmid pAMβ1. FEMS Microbiol. Lett. 42: 113-119.

Oultram, J.D., Peck, H., Brehm, J.K., Thompson, D.E., Swinfield,T.J.,and Minton, N.P. (1988a). Introduction of genes for leucine biosynthesis from Clostridium pasteurianum to C. acetobutylicum by cointegrate conjugal transfer. Mol. Gen. Genet. 214: 177-179.

Oultram, J.D., Loughlin, M., Swinfield, T.J., Brehm, J.K., Thompson, D.E., and Minton, N.P. (1988b). Introduction of plasmids into whole cells of Clostridium acetobutylicum by electroporation. FEMS Microbiol. Lett. 56: 83-88.

Pabo, C.O., and Sauer, R.T. (1984). Protein-DNA recognition. Ann. Rev. Biochem. 53: 293-321.

Pan-Hou, H.S.K., Hosono, M., and Imura, N. (1980). Plasmid-controlled mercury biotransformation by Clostridium cochlearium T-2. Appl. Environ. Microbiol. 40: 1007-1011.

Peeters, B.P.H., Schoenmakers, J.G.G., and Konings, R.N.H.(1987). Comparison of the DNA sequences involved in replication and packaging of the filamentous phages IKe and Ff (M13, fd, and f1). DNA 6: 139-147

Peschke, U., Bauck, V., Bujard, H., Gentz, R., and Le Grice, S. (1985). Efficient utilization of Escherichia coli transcriptional signals in Bacillus subtilis. J. Mol. Biol. 186: 547-555.

Petre, D., Millet, J., Longin, R., Beguin, P., Girard, H., and Aubert, J.P. (1986). Purification and properties of the endoglucanase C of Clostridium thermocellum produced in Escherichia coli. Biochimie 68: 687-695.

Piggot, P.J., and Hoch, J.A. (1985). Revised genetic linkage map of Bacillus subtilis. Microbiol. Rev. 49: 158-179.

Platt, T. (1986). Transcription termination and the regulation of gene expression. Ann. Rev. Biochem. 55: 339-372.

Popoff, M.R., and Truffaut, N. (1985). Survey of plasmids in Clostridium butyricum and Clostridium beijerinckii strains from different origins and different phenotypes. Current Microbiol. 2: 151-156.

Posno, M., Leer, R.J., van Rijn, J.M.M., Lokman, B.C., and Pouwels, P.H. (1988). Transformation of Lactobacillus plantarum by plasmid DNA, p. 397-401. In, A.T. Ganeson and Hoch, J.A. (ed.), Genetics and biotechnology of Bacilli, Volume II. Academic Press Inc. New York.

Potter, H., Weir, L., and Leder, P. (1984). Enhancer-dependent expression of human k immunoglobulin genes introduced into mouse pre-B lymphocytes by electroporation. Proc. Natl. Acad. Sci. U.S.A. 81: 7161-7165.

Powell, I.B., Achen, M.G., Hillier, A.J., and Davidson, B.E. (1988). A simple and rapid method for genetic transformation of lactic streptococci by electroporation. Appl. Environ. Microbiol. 54: 655-660.

Prazmowski, A. (1880). Untersuchung uber die Entwicklungsgeschichte und Fermentwirkung einiger Bakterien-Arten. Inaug. Diss. Hugo Voigt, Leipzig, p. 1-58.

Primrose, S., and Ehrlich, S.D. (1981). Isolation of plasmid deletion mutants and study of their instability. Plasmid 6: 193-201.

Projan, S.J., Kornblum, J., Moghazeh, S.L., Edelman, I., Gennaro, M.L., and Novick, R.P. (1985). Comparative sequence and functional analysis of pT181 and pC221, cognate plasmid replicons from Staphylococcus aureus. Mol. Gen. Genet. 199: 452-464.

Projan, S.J., Moghazeh, S., and Novick, R.P. (1988). Nucleotide sequence of pS194, a streptomycin-resistance plasmid from Staphylococcus aureus. Nucleic Acids Res. 16: 2179-2187.

Pugsley, A.P. (1984). The ins and outs of colicins. I. Production and translocation accross membranes. Microbiol. Sci. 1: 168-175.

Puyet, A., del Solar, G.H., and Espinosa, M. (1988). Identification of the origin and direction of replication of the broad-host-range plasmid pLS1. Nucleic Acids Res. 16: 115-133.

Raibaud, O., and Schwartz, M. (1984). Positive control of transcription initiation in bacteria. Ann. Rev. Genet. 18: 173-206.

Reed, K.C., and Mann D.A. (1985). Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res. 13: 7207-7221.

Reid, S.J., Errol, R., Jones, D.T., and Woods, D.R. (1983). Transformation of Clostridium acetobutylicum protoplasts with bacteriophage DNA. Appl. Environ. Microbiol. 45: 305-307.

Reysset, G., Hubert, J., Podvin, L., and Sebald, M. (1987). Protoplast formation and regeneration of Clostridium acetobutylicum strain N1-4080. J. Gen. Microbiol. 133: 2595-2600.

Reysset, G., and Sebald, M. (1985). Conjugal transfer of plasmid mediated antibiotic resistance from streptococci to Clostridium acetobutylicum. Ann. Inst. Pasteur Paris. Microbiol. 136B: 275-282.

Rigby, P.W.J., Dieckmann, M., Rhodes, C., and Berg, P. (1977). Labelling deoxyribonucleic acid to high specific activity in vitro by nick-translation with DNA polymerase I. J. Mol. Biol. 113: 237-251.

Roberts, R.J. (1987). Restriction enzymes and their isochizomers. Nucleic Acids Res. 15 (suppl.): r189-r217.

Robson, B. (1974). Analysis of the code relating sequence to conformation in globular proteins. Biochem. J. 141: 853-867.

Rogers, P. (1986). Genetics and biochemistry of Clostridium relevant to development of fermentation processes. Adv. Appl. Microbiol. 31: 1-60.

Roggentin, P., Roth, B., Lottspeich, F., and Schauer, R. (1988). Cloning and sequencing of a Clostridium perfringens sialidase gene. FEBS Lett. 238: 31-34.

Rokos, E.A., Rood, J.I., and Duncan, C.L. (1978). Multiple plasmids in different toxigenic types of Clostridium perfringens. FEMS Microbiol. Lett. 4: 323-326.

Romaniec, M.P.M., Davidson, K., and Hazlewood, G.P. (1987). Cloning and expression in Escherichia coli of Clostridium thermocellum DNA encoding beta-glucosidase activity. Enzyme Microb. Technol. 9: 474-478.

Romero, D.A., Slos, P., Robert, C., Castellino, I., and Mercenier, A. (1987). Conjugative mobilization of an alternative vector delivery system for lactic streptococci. Appl. Environ. Microbiol. 53: 2405-2413.

Rood, J.I. (1983). Transferable tetracycline resistance in Clostridium perfringens strains of porcine origin. Can. J. Microbiol. 29: 1241-1246.

Rood, J.I., Scott, V.N., and Duncan, C.L. (1978). Identification of a transferable tetracycline resistance plasmid (pCW3) from Clostridium perfringens. Plasmid 1: 563-570.

Rosenberg, M., and Court, D. (1979). Regulatory sequences involved in the promotion and termination of RNA transcription. Ann. Rev. Genet. 13: 319-53.

Ross, D. (1961). The acetone-butanol fermentation. Prog. Ind. Microbiol. 3: 73-85.

Ruhfel, R.E., Robillard, N.J., and Thorne, C.B. (1984). Interspecies transduction of plasmids among Bacillus anthracis, B. cereus, and B. thuringensis. J. Bact. 157: 708-711.

Sacco, M., Millet, J., and Aubert, J.P. (1984). Cloning and expression in Saccharomyces cerevisiae of a cellulase gene from Clostridium thermocellum. Ann. Microbiol. Inst. Pasteur Paris. 155A: 485-488.

Sala-Trepat, J.M., and Evans, W.C. (1971). The meta cleavage of catechol by Azotobacter species. Eur. J. Biochem. 20: 400-413.

Sanders, M.E., and Nicholson, M.A. (1987). A method for genetic transformation of nonprotoplasted Streptococcus lactis. Appl. Environ. Microbiol. 53: 1730-1736.

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

Sanger, F., and Coulson, A.R. (1978). The use of thin acrylamide gels for DNA sequencing. FEBS Lett. 87: 107-110.

Sauer, R.T., and Anderegg, R. (1978). Primary structure of lambda repressor. Biochemistry 17: 1092-1100.

Schaberg, D.R., Clewell, D.B., and Glatzer, L. (1982). Conjugative transfer of R-plasmids from Streptococcus faecalis to Staphylococcus aureus. Antimicrob. Agents and Chemother. 22: 204-207.

Schaeffer, P., Cami, B., and Hotchkiss, R.D. (1976). Fusion of bacterial protoplasts. Proc. Natl. Acad. Sci. U.S.A. 73: 2151-2155.

Schallehn, G., and Kramer, J. (1981). Detection of plasmids in alpha toxin-producing and non-producing strains of Clostridium novyi type A. FEMS. Microbiol. Lett. 11: 313-316.

Schaller, H. (1978). The intergenic region and the origins for filamentous phage DNA replication. Cold Sprind Harbour Symp. Quant. Biol. 43: 401-408.

Schallehn, G., and Eklund, M.W. (1980). Conversion of Clostridium novi type D (C. haemolyticum) to alpha toxin production by phages of C.novvi type A. FEMS Microbiol. Lett. 7: 83-86.

Schlechte, H., Schwabe, K., Mehnert, W.H., and Brauniger, H. (1982). Chemotherapy for tumours using clostridial oncolysis, antibiotics and cyclophosphamide: model trial on the UVT 15264 tumour. Arch. Gesch-wulstforsch. 52: 41-48.

Schwarz, W., Bronnenmeier, K., and Staudenbauer, W.L. (1985). Molecular cloning of Clostridium thermocellum genes involved in beta-glucan degradation in bacteriophage lambda. Biotechnol. Lett. 7: 859-864.

Schwarz, W.H., Grabnitz, F., and Staudenbauer, W.L. (1986). Properties of a Clostridium thermocellum endoglucanase produced in Escherichia coli. Appl. Environ. Microbiol. 51: 1293-1299.

Schwarz, W.H., Schimming, S., Rcknagel, K.P., Burgschwaiger, S., Kreil, G., and Staudenbauer, W.L. (1988a). Nucleotide sequence of the cel C gene encoding endoglucanase C of Clostridium thermocellum. Gene 63: 23-30.

Schwarz, W.H., Schimming, S., and Staudenbauer, W.L. (1988b). Isolation of a Clostridium thermocellum gene encoding a thermostable alpha-1,3-glucanase (laminarinase). Biotechnol Lett. 10: 225-230.

Schwarz, W.H., Jarvis, S., Kouba, M., Adelsberger, H., and Staudenbauer, W.L. (1988c). Molecular cloning and expression in Escherichia coli of Clostridium stercorarium genes involved in cellulose and xylan degradation. Mol. Gen. Genet. Submitted.

Scott, J.R. (1984). Regulation of plasmid replication. Microbiol. Rev. 48: 1-23

Scott, V.N., and Duncan, C.L. (1978). Cryptic plasmids in Clostridium botulinum and C.botulinum-like organisms. FEMS Microbiol. Lett. 4: 55-58.

Sebald, M., Bouanchaud, D., and Bieth, G. (1975). Nature plasmidique de la resistance plusiers antibiotiques chez Clostridium perfringens typeA, souche 659. C. R. Acad. Sci. Paris ser. D 280: 2401-2404.

Selker, E., and Yanofsky, C. (1979). Nucleotide sequence of the trpB-trpC intercistronic region from Salmonella typhimurium. J. Mol. Biol. 130: 135-143.

Senghas, E., Jones, J.M., Yamamoto, M., Gawron-Burke, C., and Clewell, D.B. (1988). Genetic organization of the bacterial conjugative transposon Tn916. J. Bact. 170: 245-249.

Shaw, J.H., and Clewell, D.B. (1985). Complete nucleotide sequence of macrolide-lincosamide-streptogramin B-resistance transposon Tn917 in Streptococcus faecalis. J. Bact. 164: 782-796.

Shimoi, H., Nagata, S., Esaki, N., Tanaka, H., and Soda, K. (1987). Leucine dehydrogenase of a thermophilic anaerobe, Clostridium thermoaceticum: gene cloning, purification, and characterization. Agric. Biol. Chem. 51: 3375.

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

Shishido, K. (1988). An occurrence of a noticeable alternating pyrimidine-purine run in the replication origins of tetracycline-resistance plasmids pNS1 and pT181. Nucleic Acids Res. 16: 1640.

Shlomai, J., and Kornberg, A. (1980). An Escherichia coli replication protein that recognizes a unique sequence within a hairpin region in ØX174 DNA. Proc. Natl. Acad. Sci. U.S.A. 77: 799-803.

Shoemaker, N.B., Smith, M.D., and Guild, W.R. (1980). DNAase-resistant transfer of chromosomal cat and tet insertions by filter mating in Pneumococcus. Plasmid 3: 80-87.

Shorenstein, R.G., and Losick, R. (1973). Comparative size and proteins of the sigma subunits of RNA polymerase from Bacillus subtilis and Escherichia coli. J. Biol. Chem. 248: 6170-6173.

Shpaer, E.G. (1986). Constraints on codon context in Escherichia coli genes. Their possible role in modulating the efficiency of translation. J. Mol. Biol. 188: 555-564.

Shrago, A.W., Chassy, B.M., and Dobrogosz, W.J. (1986). Conjugal plasmid transfer (pAMβ1) in Lactobacillus plantarum. Appl. Environ. Microbiol. 52: 574-576.

Shuttleworth, H.L., Duggleby, C.J., Jones, S.A., Atkinson, T., and Minton, N.P. (1987). Nucleotide sequence analysis of the gene for protein A from Staphylococcus aureus Cowan I (NCTC 8530) and its enhanced expression in Escherichia coli. Gene 58: 283-295.

Simon, D., Rouault, A., and Chopin, M.C. (1985). Protoplast transformation of group N streptococci with cryptic plasmids. FEMS Microbiol. Latt. 26: 239-241.

Simon, H., Bader, J., Gunther, H., Neumann, S., and Thanos, J. (1985). Biohydrogenation and electromicrobial and electroenzymatic reduction methods for the preparation of chiral compounds. Ann. N.Y. Acad. Sci. 434: 171-181.

Sims, J., and Dressler, D. (1978). Site-specific initiation of a DNA fragment: nucleotide sequence of the bacteriophage G4 negative-strand initiation site. Proc. Natl. Acad. Sci. U.S.A. 75: 3094-3098

Singer, B.W., Gold, L., Shinedling, S.T., Colkitt, M., Hunter, L.R.,Pribnow, D., and Nelson, M.A. (1981). Analysis in vivo of translational mutants of the rII cistron of bacteriophage T4. J. Mol. Biol. 149: 405-432.

Smith, C.J., Markowitz, S.M., and Macrina, F. (1981). Transferable tetracycline resistance in Clostridium difficile. Antimicrob. Agents and Chemother. 19: 997-1003.

Smith, H.O. (1980). Recovery of DNA from gels. Methods Enzymol. 65: 371-380.

Smith, M.D. (1985). Transformation and fusion of Streptococcus faecalis protoplasts. J. Bact. 162: 92-97.

Snedecor, B.R., and Gomez, R.F. (1983). Prospects for the use of genetic engineering to produce industrial chemicals with strict naerobes. In, D.L. Wise (ed.), Organic chemicals from biomass. Benjamin Cummings Publishing Co. Inc., London.

Solberg, M., Blaschek, H.P., and Khan, P. (1981). Plasmids in Clostridium perfringens. J. Food Safety 3: 267-289.

Southern, E. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503-507.

Soutschek-Bauer, E., Hartl, L., and Staudenbauer, W.L. (1985). Transformation of Clostridium thermohydrosulphuricum DSM-568 with plasmid DNA. Biotechnol. Lett. 7: 705-710.

Soutschek-Bauer, E., and Staudenbauer, W.L. (1987). Synthesis and secretion of a heat-stable carboxymethylcellulase from Clostridium thermocellum in Bacillus subtilis and Bacillus stearothermophilus. Mol. Gen. Genet. 208: 537-541.

Sowers, A.E., and Lieber, M.R. (1986). Electropore diameters, lifetimes, numbers and locations in individual erythrocyte ghosts. FEBS Lett. 205: 179-184.

Squires, C.H., Heefner, D.L., Evans, R.J, Kopp, B.J, and Yarus, M.J. (1984). Shuttle plasmids for Escherichia coli and Clostridium perfringens. J. Bact. 159: 465-471.

Stallcup, M.R., and Rabinowitz, J.C. (1973). Specificity in the translation of natural messenger ribonucleic acids. J. Biol. Chem. 248: 3209-3215; 3216-3219.

Stanssens, P., Remaut, E., and Fiers, W. (1985). Alterations upstream from the Shine-Dalgarno region and their effect on bacterial gene expression. Gene 36: 211-223.

Stewart, G.J., and Carlson, C.A. (1986). The biology of natural transformation. Ann. Rev. Microbiol. 40: 211-236.

Stephens, M.A., Ortlepp, S.A., Ollington, J.F., and McConnell, D.J. (1984). Nucleotide sequence of the 5' region of the Bacillus licheniformis alpha-amylase gene: comparison with the B. amyloliquefasciens gene. J. Bact. 158: 369-372.

Strom, M.S., Eklund, M.W., and Poysky, F.T. (1984). Plasmids in Clostridium botulinum and related Clostridium species. Appl. Environ. Microbiol. 48: 956-963.

Suzuki, S., and Karube, I. (1980). Biochemical energy conversion using immobilized whole cells of Clostridium butyricum. Biochimie 62: 353-358.

Swanstrom, R. and Shank, P.R. (1978). X-ray intensifying screens greatly enhance the detection by autoradiography of the radioactive isotopes $^{32}P$ and $^{125}I$. Anal. Biochem. 86: 184-192.

Takahashi, W., Yamagata, H., Yamaguchi, K., Tsukagoshi, N., and Udaka, S. (1983). Genetic transformation of Bacillus brevis 47, a protein-secreting bacterium, by plasmid DNA. J. Bact. 156: 1130-1134.

Tannock, G.W. (1987). Conjugal transfer of plasmid pAMβ1 in Lactobacillus reuteri and between lactobacilli and Enterococcus faecalis. Appl. Environ. Microbiol. 53: 2693-2695.

te Riele, H., Michel, B., and Ehrlich, S.D. (1986a). Single-stranded plasmid DNA in Bacillus subtilis and Staphylococcus aureus. Proc. Natl. Acad. Sci. U.S.A. 83: 2541-2545.

te Riele, H., Michel, B., and Ehrlich, S.D. (1986b). Are single-stranded circles intermediates in plasmid DNA replication? EMBO J. 5: 631-637.

Thummel, C.S., Burgess, T.L., and Tjian, R. (1981). Properties of simian virus 40 small t antigen overproduced in bacteria. J. Virol. 37: 683-697.

Tinoco, I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O. C., Crothers, D.M., and Gralla, J. (1973). Improved estimation of secondary structure in ribonucleic acids. Nature New Biol. 246: 40-41.

Truffaut, N., and Sebald, M. (1983). Plasmid detection and isolation in strains of Clostridium acetobutylicum and related species. Mol. Gen. Genet. 189: 178-180.

Twigg, A.J., and Sherratt, D. (1980). Trans-complementable copy-number mutants of plasmid ColE1. Nature 283: 216-218.

Urano, N., Karube, I., Suzuki, S., Yamada, T., Hirochika, H., and Sakaguchi, K. (1983). Isolation and partial characterization of large plasmids in hydrogen-evolving bacterium Clostridium butyricum. Eur. J. Appl. Microbiol. Biotechnol. 17: 349-354.

Usdin, K.P., Zappe, H., Jones, D.T., and Woods, D.R. (1986). Cloning, expression, and purification of glutamine synthetase from Clostridium acetobutylicum. Appl. Environ. Microbiol. 52: 413-419.

van Mansfeld, A.D.M., van Teeffelen, H.A.A.M., Baas, P.D., and Janz, H.S. (1986). Two juxtapositioned tyrosyl-OH groups participate in ØX174 gene A protein catalysed cleavage and ligation of DNA. Nucleic Acids Res. 14: 4229-4238.

Vasseghi, H., and Claverys, J.P. (1983). Amplification of a chimaeric plasmid carrying an erythromycin-resistance determinant introduced into the genome of Streptococcus pneumoniae. Gene. 21: 285-292.

Vescco, M., Morelli, L., Botazzi, V., and Gasson, M.J. (1983). Conjugal transfer of broad-host-range plasmid pAMβ1 into enteric species of lactic acid bacteria. Appl. Environ. Microbiol. 46: 753-755.

Vieira, J., and Messing, J. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. (1982). Gene 19: 259-268.

Villafana, R., Bechhofer, D.H., Narayanan, C.S., and Dubnau, D. (1987). Replication control genes of plasmid pE194. J. Bact. 169: 4822-4829.

Viret, J. and Alonso, J.C. (1987), Generation of linear multigenome-length plasmid molecules in Bacillus subtilis. Nucleic Acid. Res. 16: 6349-6367.

Vold, B.S. (1973). Analysis of isoaccepting transfer RNA species in Bacillus subtilis. J. Bact. 113: 825-833.

Vold, B.S. (1985). Structure and organization of genes for transfer ribonucleic acid in Bacillus subtilis. Microbiol. Rev. 49: 71-80.

Volk, W.A., Bizzini, B., Jones, K.R., and Macrina, F.L. (1988). Inter-and intragenic transfer of Tn916 between Streptococcus faecalis and Clostridium tetani. Plasmid 19: 255-259.

Wahl, G.M., Stern, N., and Stark, G.R. (1979). Efficient transfer of large DNA fragments from agarose gels to diazobenzylmethyl-paper and rapid hybridization by using dextran sulphate. Proc. Natl. Acad. Sci. U.S.A. 76: 3683-3687.

Walker, G.C. (1983). Genetic strategies in strain design for fermentations, p. 349-376. In,. A. Hollaender, A.I. Laskin, and P. Rogers (ed.), Basic biology of new developments in biotechnology. Plenum Publishing Corp., New York.

Wallace, R.B., Johnson, N.J., Hirose, T., Miyake, M., Kawashima, E.H., and Itakura, K. (1981). The use of synthetic oligonucleotides as hybridization probes. II. Hybridization of oligonucleotides of mixed sequence to rabbit beta-globin DNA. Nucleic Acids Res. 9: 879-894.

Wallace, R.B., Shaffer, J., Murphy, R.F., Bonner, J., Hirose, T., and Itakura, K. (1979). Hybridization of synthetic oligodeoxyribonucleotides to ØX174 DNA: the effect of single base pair mismatch. Nucleic Acids Res. 11: 3543-3557.

Wang. S.Z., Chen, J.S., and Johnson, J.L. (1987). Nucleotide and deduced amino acid sequences of nifD encoding the alpha-subunit of nitrogenase MoFe protein of Clostridium pasteurianum. Nucleic Acids Res. 15: 3935.

Wang, S.Z., Chen, J.S., and Johnson, J.L. (1988). The presence of five nifH-like sequences in Clostridium pasteurianum: sequence divergence and transcriptional properties. Nucleic Acids Res. 16: 439-454.

Weickert, M.J., Chambliss, G.H., and Sugiyama, H. (1986). Production of toxin by Clostridium botulinum type A strains cured of plasmids. Appl. Environ. Microbiol. 51: 52-56.

West, R.W.Jr., and Rodriguez, R.W. (1982). Construction and characterization of E. coli promoter-probe plasmid vectors, III. pBR322 derivatives with deletions in the tetracycline resistance promoter region. Gene 20: 291-304.

Whitehead, T.R., and Rabinowitz, J.C. (1986). Cloning and expression in Escherichia coli of a gene for 10-formyltetrahydrofolate synthetase from Clostridium acidiurici ("Clostridium acidi-urici"). J. Bact. 167: 205-209.

Wiegel, J., and Ljungdahl, L.G. (1986). The importance of thermophilic bacteria in biotechnology. CRC Crit. Rev. Biotechnol. 3: 39-108

Wiggs, J.L., Bush, J.W., Chamberlin, M.J. (1979). Utilization of promoter and terminator sites on bacteriophage T7 DNA by RNA polymerases from a variety of bacterial orders. Cell 16: 97-109.

Wren, B.W., Clayton, C.L., Mullany, P.P., and Tabaqchali, S. (1987). Molecular cloning and expression of Clostridium difficile toxin A in Escherichia coli. FEBS Lett. 225: 82-86.

Wren, B.W., Mullany, P.P., Clayton, C.L., and Tabaqchali, S. (1988). Molecular cloning and genetic analysis of a chloramphenicol acetyltransferase determinant from Clostridium difficile. Antimicrob. Agents and Chemother. 23: 784-786.

Wust, J., and Hardegger, U. (1983). Transferable resistance to clindamycin, erythromycin, and tetracycline in Clostridium difficile. Antimicrob. Agents and Chemother. 23: 784-786.

Wyrick, P.B., and Rogers, H.J. (1973). Isolation and characterization of cell wall-defective variants of Bacillus subtilis and Bacillus licheniformis. J. Bact. 116: 456-465.

Yagi, Y., Mclellan, T., Frez, T., and Clewell, D. (1978). Characterization of a small plasmid determining resistance to erythromycin, lincomycin, and veramycin Ba in a strain of Streptococcus sanguis isolated from dental plaque. Antimicrob. Agents and Chemother. 13: 884-887.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33: 103-119.

Yansura, D.G., and Henner, D.J. (1984). Use of the Escherichia coli lac repressor and operator to control gene expression in Bacillus subtilis. Proc. Natl. Acad. Sci. U.S.A. 81: 439-443.

Yoshimura, K., Uemura, J., Seki, T., and Oshima, Y. (1984). Construction of a promoter-probe vector for a Bacillus subtilis host by using the trpD$^+$ gene of Bacillus amyloliquefaciens. J. Bact. 159: 905-912.

Yoshino, S., Ogata, S., and Hayashida, S. (1984). Regeneration of protoplasts of Clostridium saccharo-perbutylacetonicum. Agric. Biol. Chem. 48: 249-250.

Youngleson, J.S., Santangelo, J.D., Jones, D.T., and Woods, D.R. (1988). Cloning and expression of a Clostridium acetobutylicum alcohol dehydrogenase gene in Escherichia coli. Appl. Environ. Microbiol. 54: 676-682.

Yu, P.L., and Pearce, L.E. (1986). Conjugal transfer of streptococcal antibiotic resistance plasmids into Clostridium acetobutylicum. Biotechnol. Lett. 8: 469-474.

Zappe, H., Jones, D.T., and Woods, D.R. (1986). Cloning and expression of Clostridium acetobutylicum endoglucanase, cellobiase and amino acid biosynthesis genes in Escherichia coli. J. Gen. Microbiol. 132: 1367-1372.

Zappe, H., Jones, D.T., and Woods, D.R. (1987). Cloning and expression of a xylanase gene from Clostridium acetobutylicum P262 in Escherichia coli. Appl. Microbiol. Biotechnol. 27: 57-63.

Zappe, H., Jones, W.A., Jones, D.T., and Woods, D.R. (1988). Structure of an endo-a-1,4-glucanase gene from Clostridium acetobutylicum P262 showing homology with endoglucanase genes from Bacillus spp. Appl. Environ. Microbiol. 54: 1289-1292.

Zeikus, J.G. (1980). Chemical and fuel production by anaerobic bacteria. Ann. Rev. Microbiol. 34: 423-464.

Zukowski, M.M., Gaffney, D.F., Speck, D., Kauffmann, M., Findeli, A., Wisecup, A., and Lecocq, J.P. (1983). Chromogenic identification of genetic regulatory signals in Bacillus subtilis based on expression of a cloned Pseudomonas gene. Proc. Natl. Acad. Sci. U.S.A. 50: 1101-1105.

Zukowski, M., and Miller, L. (1986). Hyperproduction of an intracellular heterologous protein in a $sacU^h$ mutant of Bacillus subtilis. Gene 46: 247-255.

Zukowski, M., Miller, L., Cogswell, P., Chen, K. (1988). Inducible expression system based on sucrose metabolism genes of Bacillus subtilis. p. 17-22. In. A.T. Ganeson and J.A. Hoch (eds.), Genetics and Biotechnology of Bacilli, volume II. Academic Press Inc., New York.