# DEVELOPMENT AND USE OF A VECTOR SYSTEM FOR *METHYLOPHILUS METHYLOTROPHUS*.

GEOFFREY S. SHARPE

Ph.D. THESIS

LEICESTER UNIVERSITY

1984.





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

FOR

MY FAMILY

FOR THEIR LOVE, PATIENCE AND ENCOURAGEMENT.

ACKNOWLEDGEMENTS

I am indebted to Dr. E.C.Dart and Professor W.J.Brammar for their initial encouragement and help, and to Imperial Chemical Industries plc for paying my salary.

I wish to thank Professors W.J.Brammar and I.B.Holland and Dr. J.D.Windass for their critical and helpful dicussions throughout the duration of this work, with special thanks to Professor W.J.Brammar for critically reading the manuscript. I also wish to thank Drs. P.T.Barth and B.Wilkins for valuable informal discussions, Pam M$^{C}$Turk for her careful and painstaking work on the electron microscope and all my collegues in the ICI/University Joint Laboratory for their help, advice and friendship.

**CONTENTS**

# CHAPTER 1

## Introduction

### 1. The "host" organism, *Methylophilus methylotrophus*.

Single cell micro-organisms, yeasts in particular, have long been known to be rich sources of high-quality protein. Many micro-organisms during growth convert inexpensive non-proteinaceous materials, such as carbohydrates and hydrocarbons, and simple nitrogen compounds, such as nitrates and ammonia, into protein. This conversion can be extremely fast, (mass doubling times of 20-120 minutes are not uncommon), and has led to the development of viable processes involving single-cell micro-organisms to produce cheap, high-quality protein products as supplements for animal feeds.

In the late 1960's ICI commenced research on a bacterial "single cell protein" (SCP) process using methane gas as sole carbon source for direct feeding to a micro-organism. Methane was the obvious choice for ICI as it was cheap, plentiful, readily available from the North Sea, and micro-organisms known to metabolize methane had been studied in detail. However there were serious drawbacks to using methane gas, since being gaseous and of limited solubility in water, energy had to be used to supply the organism with a soluble substrate, and the microbial yield of the organism in pure culture from methane was poor. (The yield as a percentage of carbon-source carbon incorporated into cellular organic carbon is referred to as "carbon conversion"). Coincidently, ICI developed an interest in methanol synthesis from natural gas and in the early 1970's methane was abandoned and a methanol-based route to SCP initiated.

Once methanol had been identified as the substrate of choice, the appropriate micro-organism had to be found. The characteristics necessary for successful development were that the micro-organism

should be stable and capable of high carbon conversion, be safe in use and produce a non-toxic, nutritious end-product. A bacterium was seen as likely to be capable of matching these characteristics, since, generally, bacterial growth rates and yields are high and the final crude protein-content of the end-product of bacterial fermentation is usually higher than that of other organisms (MacLennan *et al.*, 1973). In the initial search, ICI examined a large number of methylotrophs and those that appeared promising after first evaluation were subjected to preliminary screening for pathogenicity and toxicity. The outcome of this screening was the selection of a single obligate methanol-utiliser, *Methylophilus methylotrophus*.

The organism was found to grow readily, with high doubling rates, in laboratory fermenters, using methanol as carbon-source and ammonia as nitrogen-source. In order to increase the fermenter size to achieve commercial viability, the novel technology of pressure-cycle fermentation was developed, with a $37m^3$ pilot plant becoming operational in 1973 and a $1500m^3$ production plant in 1980.

> [It may help to appreciate the scale of the operation to review some statistics of this production fermenter: it was workshop-built in one piece weighing over 600t and standing 60m high. It has been run continuously for periods in excess of one hundred days, and has the capacity to produce 50-70kt of SCP per year].

With several years of fermentation research having been devoted to optimizing the growth of *Methylophilus methylotrophus*, further significant improvements to growth efficiency were unlikely to be achieved by manipulation of the fermentation environment alone. Biochemical studies revealed a potential source of methanol-wastage, the organism's mode of ammonia-assimilation using the low efficiency,

high affinity, two-stage pathway dependent on glutamine synthetase and glutamate synthase (GOGAT). Other micro-organisms, including *E.coli*, possess a low affinity, high efficiency pathway using glutamate dehydrogenase, which requires one less molecule of ATP per molecule of glutamine formed (Tyler,1978).

In what was probably the first example of commercially significant genetic engineering, the glutamate dehydrogenase gene from *E.coli* was cloned on a plasmid and transferred to a temperature-sensitive GOGAT$^-$ *M.methylotrophus* (Windass *et al.*,1980), allowing more efficient conversion of growth substrate, methanol, into cellular carbon. However, the project highlighted several problems relating to the delivery of the cloned gene into the methylotroph, as no transformation procedure for this organism has yet been identified. Thus only self-transmissible or mobilizable plasmids capable of replication in the methylotroph could be considered as vectors, ruling out the common *E.coli* plasmid vectors, e.g. pBR322 (Bolivar *et al.*,1977) and pACYC184 (Chang and Cohen,1978). Instead, two other plasmid derivatives were used, pRP301 a 54.7kb, low copy number, conjugative derivative of RP4 (Barth,1979) and pTB70, a derivative of the wide host-range, multicopy, IncQ plasmid R300B (Barth and Grinter,1974; Grinter and Barth,1976), that carried the transposon Tn*5* in order to introduce a site for the restriction endonuclease *Sal*I. pRP301 was relatively large, making it difficult to handle for vector purposes and producing only a low gene-dosage of the cloned gene, while pTB70 contained the cloned gene on an active transposon together with an extra piece of DNA of uncertain origin.

Clearly, a more convenient vector system capable of readily accepting fragments sub-cloned from other vectors, or of directly replacing the *E.coli* cloning vectors altogether, would facilitate the use of *M.methylotrophus* as a host organism. Further, if the

production of cloned gene-products could be demonstrated in *M.methylotrophus* similar to that in *E.coli*, then a methylotrophic host would begin to look attractive with its advantages of safety and potential for massive scale-up.

## 1.2. The plasmid vectors.

The fundamental requirement in DNA cloning is to be able to join a segment of DNA of interest to a "vector" molecule which will then enable propagation of that segment to take place in bacteria. Most of the early recombinant DNA work was based on *E.coli*, and the vectors used were derived from its natural phages or plasmids. Plasmid vectors were rapidly developed to represent an increasingly wide-ranging and versatile set of cloning "tools". Interestingly, DNA cloning techniques applied to the plasmids themselves have also contributed to the understanding of important plasmid functions such as replication, partition and copy number control. Of special interest in this thesis are the plasmids which are capable of replication in organisms other than *E.coli*, particularly in *M.methylotrophus*.

Plasmids are extrachromosomal, autonomously replicating, double-stranded, circular nucleic acid molecules which are stably inherited by their (bacterial) hosts. They can be highly purified away from any host chromosomal DNA and re-introduced into a bacterial cell, either directly by transformation or indirectly by conjugation/mobilization, and it is this property that has made them such excellent candidates for cloning vectors.

An ideal plasmid vector should have a small size, making it easier to handle, small molecules being less prone to breakage by shearing. Being small is often associated with having a high copy number, which makes isolation of plasmid DNA easier and also means a higher gene-dosage of a cloned gene-product. Since the efficiency of transformation is reduced as the plasmid size increases, having a small molecule to start with maximizes the ratio of passenger to vector DNA and also simplifies restriction mapping.

A second feature of the ideal plasmid vector would be that it should be readily detectable in a host cell, distinguishable from cells which did not contain plasmids, by conferring some phenotypic property such as antibiotic resistance. Further, if two resistance genes were present on a vector, one could be inactivated by insertion of a cloned fragment of DNA while the other provided the selection for transformants. Cells carrying vector plus insert would then easily be detected by the associated change in phenotype.

To allow insertional inactivation would require that a single restriction enzyme site (or multiple sites clustered in a small region) be present within the resistance determinant; this would be a third feature of the ideal vector, that there be several such unique sites. Often, possession of unique restriction sites would be a consequence of the plasmid molecule being small, since the statistical chance of multiple sites would thereby be reduced.

Approaching the requirements of the "ideal" cloning vector are the pBR- series plasmids of which pBR322 (Bolivar *et al.*,1977) is the best known and most widely used. pBR322 was derived from a ColE1-like plasmid, pMB1, which in a complicated series of *in vitro* insertions and deletions, received tetracyline (Tc) and ampicillin (Ap) reistance genes from pSC101 and RSF2124, repectively. Eventually a 4363bp plasmid was produced encoding resistance to Ap and Tc, and containing many unique recognition sequences for restriction enzymes including: *Bam*HI; *Cla*I; *Eco*RI; *Hind*III; *Pst*I; *Sal*I and *Sph*I, with all except the *Eco*RI target situated within an antibiotic resistance gene sequence [Fig.1]. This latter defect was remedied by the insertion into pBR322 of a chloramphenicol (Cm) resistance gene from phage P1-$Cm^r$ (Bolivar,1978), carrying a unique *Eco*RI site. A further *in vitro* deletion to remove mobilization functions was made to produce pBR328 (Soberon *et al.*,1980) [Fig.2].

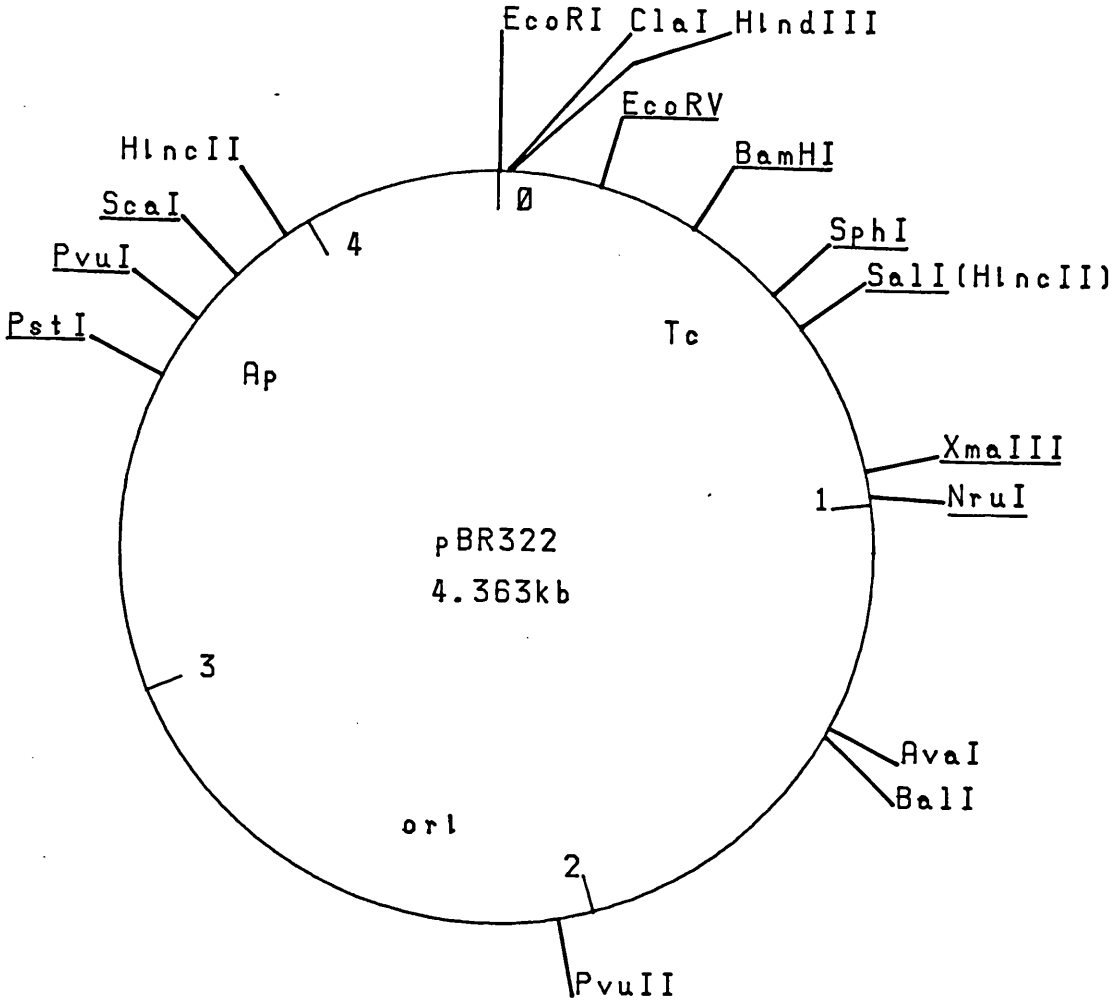



**Figure 1**

A restriction map of pBR322 (Bolivar *et al.*,1977). The enzymes which cleave the plasmid within the coding sequences of the antibiotic-resistance genes are shown underlined.

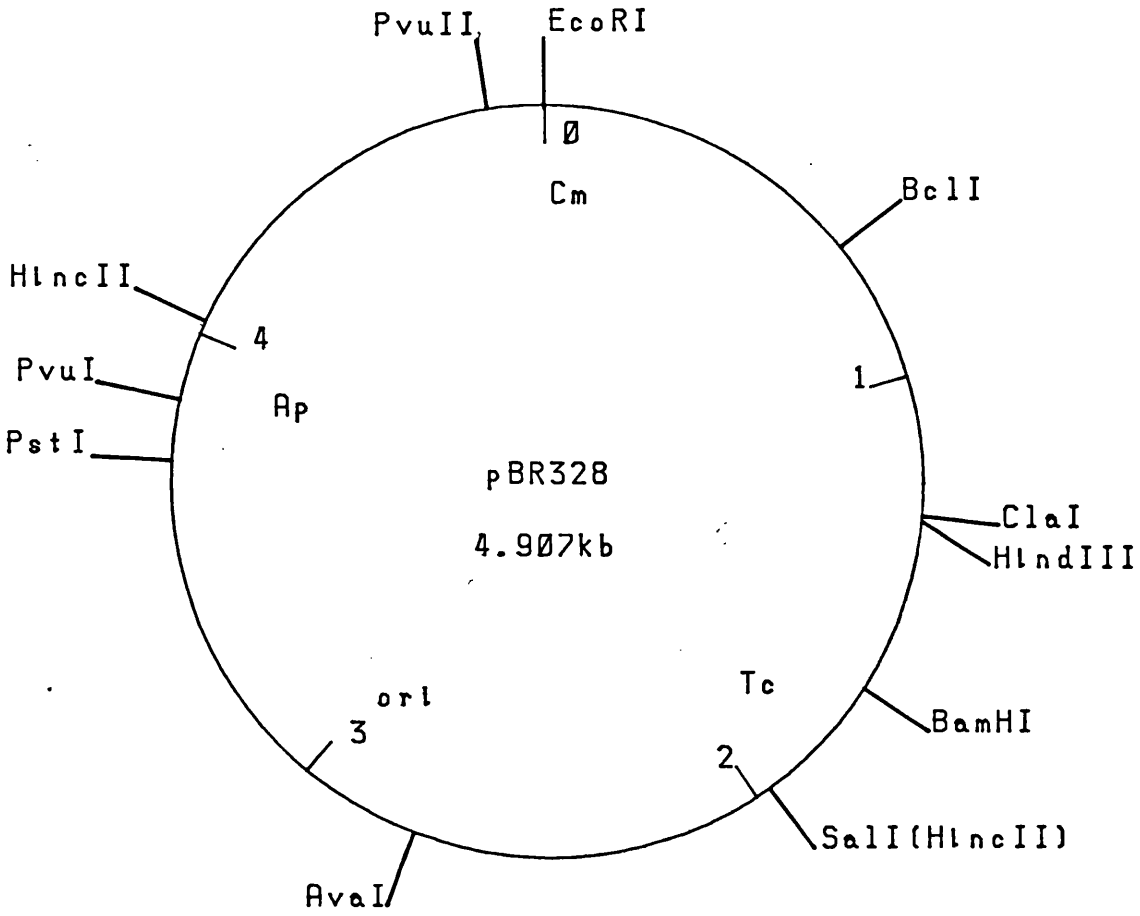



Figure 2.

A restriction map of pBR328 (Soberon *et al.*,1980). The plasmid was derived from pBR322, after insertion of the $Cm^r$ gene between the $Ap^r$ and $Tc^r$ genes, but has a 1089-bp deletion of sequence from between the end of the $Tc^r$ gene and *ori*. This plasmid now has a unique *Eco*RI site within the $Cm^r$ gene.

These plasmids share with ColE1 the ability for a "relaxed" mode of replication in which their DNA accumulates to make up about one-third of the total cellular DNA when protein synthesis is inhibited (Chang and Cohen,1978). This feature can be used to obtain high yields of a cloned fragment of DNA.

An important advantage with pBR322 and pBR328 is that complete restriction data can be predicted and updated, as their total DNA sequences are known (Sutcliffe,1978a,b; Prentki *et al.*,1981; Peden, 1983). This may become a further advantage when obtaining sequence data of a cloned fragment by allowing identification of the fragment/ plasmid junctions.

The *Pst*I site in the ampicillin-resistance (β-lactamase) gene has further advantages. If dG homopolymer tails are added to *Pst*I-cleaved pBR322 DNA and dC homopolymer tails to the DNA to be inserted, the *Pst*I sites are reconstituted in many of the resulting recombinant plasmids (Villa-Komaroff *et al.*,1978; Bolivar *et al.*,1977; Mann *et al.*, 1978). The inserted DNA segment, together with its short flanking dG:dC oligomeric segments, may then be cleaved from the vector by *Pst*I digestion. Furthermore, the segment of DNA inserted into the $Ap^r$ gene is transcribed from the B-lactamase promoter and this has permitted the transcription and translation of e.g. mouse dihydrofolate reductase in *E.coli* (Chang *et al.*,1978) and in *M. methylotrophus* (this thesis and Hennam *et al.*,1982).

The plasmids referred to above are, however, restricted to use in *E.coli* and closely related enteric bacteria such as *Salmonella*, *Proteus* and *Serratia* species. The pBR322, pACYC184, and pSC101 replicons, for example, cannot be maintained in *Pseudomonas* strains (Franklin *et al.*,1981).

In contrast, plasmids from the incompatibility groups P and Q can be maintained in almost any Gram negative bacterial species and must

be among the first to be considered in development of vectors for organisms other than *E.coli*. Availability of such vectors would widen the applicability of gene-cloning techniques to include, for example, planned development of industrially useful strains (Windass *et al.*,1980), and would allow the study of the versatile degradative enzymes of *Pseudomonas* in their normal metabolic background. This is an important consideration, as the genes coding for some enzymes e.g. those for toluene-degradation, are expressed poorly when cloned in *E.coli* (Jacoby *et al.*,1978).

The P-group plasmids as exemplified by RP4 (Datta *et al.*,1971) are generally large, low copy number, conjugative plasmids with a wide host-range, and although they have been used occasionally as vectors in their own right (Barth,1979), they tend to be too large for easy use. Smaller derivatives have been made, although these are not without their problems. A non-conjugative, non-mobilizable 11kb derivative of RK2, which is identical to RP4 (Burkhardt *et al.*,1979), has been isolated (Kahn *et al.*,1979), but can only be introduced into strains for which a transformation procedure has been defined. This problem has been circumvented by another RK2 broad host-range vector system in which the 20kb cloning vector can be mobilized by providing the transfer functions in *trans*, cloned on a hybrid ColE1 replicon (Ditta *et al.*,1980). A similar two-component system based on the IncW plasmid pSa has also been reported (Tait *et al.*,1983).

The other class of wide host-range plasmids that has been studied recently is that derived from the incompatibility group Q [or IncP4 group of *Pseudomonas* (Jacoby and Shapiro,1977)] and includes R300B (Barth and Grinter,1974), NTP2 (Smith *et al.*,1974), RSF1010 (Guerry *et al.*,1974) which is identical to NTP2 (deGraffe *et al.*,1978) and R1162 (Meyer *et al.*,1979), which is reported to be identical to RSF1010 (Gautier and Bonewald,1980). All these plasmids seem

indistinguishable at present (Barth and Grinter,1974; Heffron *et al.*, 1975), and in this thesis I have used R300B exclusively.

R300B was originally isolated from a *Salmonella typhimurium* strain that contained two plasmids, R300A, a conjugative 87kb plasmid, and R300B, a non-conjugative 8.7kb plasmid encoding resistance to sulphonamides (Su) and streptomycin (Sm) (Barth and Grinter,1974). Whilst in the wild R300A provided R300B with the ability to transfer, several other conjugative plasmids from a variety of incompatibility groups were also capable of promoting R300B mobilization, (Grinter and Barth,1976) [Table I], although RP4 was the only one that was effective in all crosses (Barth *et al.*,1981).

The host-range of the IncQ group plasmids, as represented by R300B, includes the majority of Gram negative organisms (Barth *et al.*, 1981), so that this group of plasmids has recently been the focus of interest in developing cloning vectors for use in organisms other than *E.coli*. Apart from its host-range characteristics, several other properties of R300B make it suitable as the starting point for the development of a potential cloning vector; it is relatively small at 8.68kb (Meyer *et al.*,1982) and it is multicopy, 9-12 copies per chromosome in *E.coli*, (Barth and Grinter,1974). In common with most other broad host-range plasmids, there are relatively few sites for cleavage by restriction enzymes [Fig.3] (Barth *et al.*,1981; Meyer *et al.*,1982), with only *Sst*I and *Sst*II sites in the $Sm^r$ gene and the *Pst*I sites in the $Su^r$ gene giving insertional inactivation.

Since pBR322 and pBR328 have many more useful restriction sites, several workers have made hybrid plasmids from various combinations of pBR322 and an IncQ plasmid (Bagdasarian *et al.*,1979,1981; Gautier and Bonewald,1980; Wood *et al.*,1981). However there have been

**Table I**

| PLASMIDS ABLE TO MOBILIZE R300B | |
|---|---|
| IncFI | (e.g. F104) |
| IncFII | (e.g. R1*drd*19, JR72) |
| IncIα | (e.g. R144*drd*19, R483, JR66a, R64*drd*11) |
| IncN | (e.g. N3T) |
| IncP | (e.g. RP4*, R751, R702) |
| IncW | (e.g. R7K, R388, S-a) |

* Only RP4 works in all crosses.

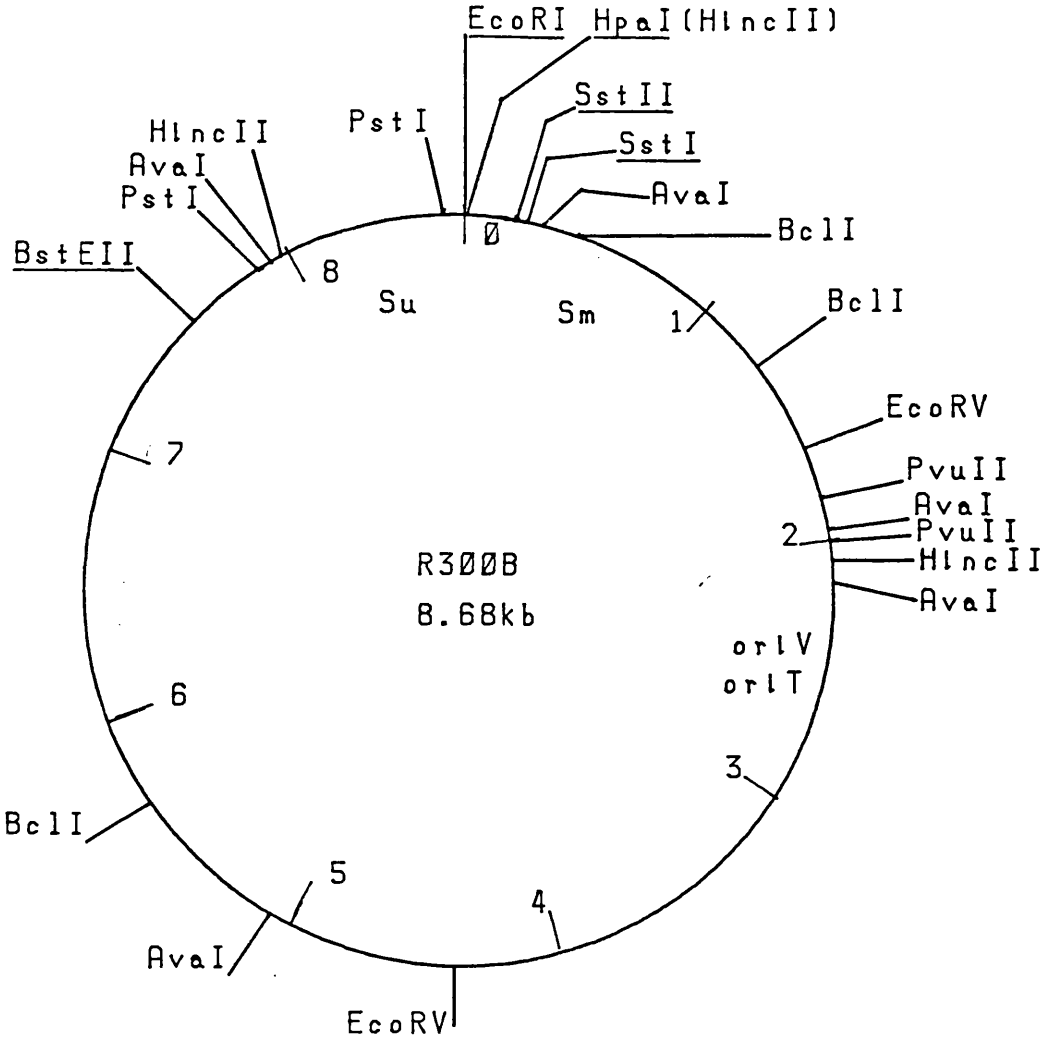



Figure 3.

A restriction map of R300B. Unique restriction sites are shown underlined. There are no sites for *Bam*HI, *Bgl*II, *Hind*III, *Kpn*I, *Sal*I, *Sma*I, *Stu*I, *Xho*I.

reports of instability (Gautier and Bonewald,1980) or reduced host-range (Bagdasarian *et al.*,1979), of such chimeric plasmids.

Here I describe the development of a series of plasmid cloning vectors that combines the broad host-range characteristics of R300B with increased cloning capability in genes derived from pBR322 and pBR328, and use some of these vectors to demonstrate the successful expression of heterologous coding sequences in *M.methylotrophus*.

## 1.3. The strategy for broad host-range vector development

From the preceding dicussions about pBR322, pBR328 and R300B, a combination of the properties of these two plasmid classes should provide a useful cloning vector for applications involving most Gram negative bacteria. The valuable range of unique cloning sites within the antibiotic-resistance genes could be coupled to a broad host-range, mobilizable replicon. The potential of such a combination has not gone unnoticed; as early as 1974 ColE1 and RSF1010 were combined via their *Eco*RI sites (Tanaka and Weisblum,1975), although this hybrid was never considered as a useful cloning vector. Since then, other workers have joined pBR322 and an IncQ plasmid (RSF1010 or R1162) via their *Pst*I sites (Bagdasarian *et al.*,1979) or via their *Eco*RI sites (Gautier and Bonewald,1980; Wood *et al.*,1981). The two latter constructions differed only in the relative orientations of the parent plasmids, though their reported properties were quite different. However, none of these arrangements has been totally satisfactory with regard to stability or host-range, and none could make use of the *Pst*I site in the B-lactamase gene.

As joining the two plasmid types via a single restriction site was so limited, I decided to use partial digests of the parental molecules with an enzyme that normally cut each plasmid at many sites. This approach has the potential to produce many different chimeric structures, that will have different properties, from which the most useful constructions could be selected.

The enzyme chosen, *Hae*II, which has eleven sites in pBR322 and about thirty in R300B, produces a staggered break and leaves a single -stranded 3'-extension having the sequence $^{5'}\_\_GCGC^{3'}$ which should anneal readily to produce a good substrate for T4-DNA ligase. Preliminary experiments revealed the possibility of producing partial

digests containing relatively high proportions of linear molecules. Thus, since it was not known at this stage how much of each plasmid was going to be required to produce a stable broad host-range cloning vector, it was thought desirable to start from a total co-integrate of both parents. This would be a similar situation to the co-integrates formed by using *Eco*RI, for example, except that here was the potential to form many different co-integrates which would then be transformed into host cells where they would undergo a crude *in vivo* selection for the most stable arrangements. Once obtained, the co-integrates could be mapped then genetically manipulated *in vitro* to give a series of derivative plasmids, with different properties, to suit different applications.

## 1.4. The delivery system for the vectors.

Isolation and amplification (cloning) of a desired arrangement of fragments from the mixture of recombinants produced in a joining reaction can be accomplished if the relevant recombinant can replicate in a host organism. A feature of *E.coli* that makes it the preferred organism for plasmid cloning is the ability of specially treated cells to take up naked vector DNA, be it plasmid or phage, by transformation or transfection (Mandel and Higa,1970), so that some cells receive recombinant molecules. This separates the various molecules in a manner that allows them to be recognized in individual clones of transformed cells selected for plasmid-specified phenotypes.

Under ideal conditions, transformation (or transfection) will typically produce $10^5$ plasmid-containing colonies (or plaques) per microgram of DNA. After a cutting and re-joining experiment the efficiency will drop by approximately two orders of magnitude. For many applications this is acceptable, but in some cases e.g. construction of gene libraries from entire complex genomes, the efficiency of obtaining recombinants, and the size of the inserted fragment would have to be increased considerably to keep the scale of the experiments within reasonable limits (less than 100$\mu$g of vector DNA). Bacteriophage $\lambda$ DNA can be packaged *in vitro* into phage capsids and inserted into cells by a normal phage infection such that approximately $10^8$ plaques per microgram of DNA are formed. Again this efficiency is reduced by approximately two orders of magnitude if the phage DNA is cut and re-ligated with donor DNA, but still represents a significant improvement over transfection.

The region on the bacteriophage $\lambda$ genome required for the packaging of DNA in phage capsids, the *cos* region, is quite small

(Hohn,1983) and has been cloned into plasmids (Collins and Hohn,1978) which then become known as "cosmids". In addition to the λ *cos* site, cosmids carry antibiotic-resistance genes and restriction sites for cloning. The *cos* site is recognized by the λ packaging system and concatemeric DNA forms, produced by ligating linear cosmid and passenger DNA at high concentrations, mimic the natural packaging substrate. Acquisition of the cosmid by the cells can be selected using the antibiotic-resistance markers, and once in the cells, the cosmids replicate as plasmids and DNA can be isolated using standard techniques. Construction of broad host-range cosmids is reported in this thesis, although attempts at packaging them *in vitro* proved unsuccessful.

Transformation and transfection procedures are not, however, universal amongst all bacterial species, although modified techniques have permitted transformation of, for example, *Pseudomonas putida* (Chakrabarty *et al.*,1975; Nagahari and Sakaguchi,1978), *Pseudomonas aeruginosa* (Nagahari and Sakaguchi,1978; Mercer and Loutit,1979), *Agrobacterium tumefaciens* (Hille and Schilperoort,1981), *Azotobacter vinelandii* (David *et al.*,1981) and *Rhodopseudomonas sphaeroides* (Fornari and Kaplan,1982). No such transformation system exists for *Methylophilus methylotrophus* and introduction of plasmids into this organism is only achieved by conjugal transfer from a suitable donor, usually *E.coli*. This then requires that any plasmid used as a vector to carry a cloned coding sequence into *M.methylotrophus*, be self-transmissible (conjugative) or mobilizable.

Conjugative plasmids, characterised by their ability to promote their own transfer, have been isolated from many Gram negative bacteria and include members of more than twenty incompatibility groups (Bukhari *et al.*,1977). Extensive regions of the plasmids seem to required for transfer (e.g. 33kb in the case of F), making these

plasmids very large and consequently quite difficult to manipulate.

Transfer of plasmid DNA is initiated at a specific site on the plasmid molecule, called the "origin of transfer" or *oriT*. Cloning techniques have been used to isolate *oriT* regions from a number of conjugative plasmids e.g. from F, on a 385bp fragment (Everett and Willets,1982), and seem to be all that is required in *cis* for conjugative transfer when the necessary *trans*-acting functions are provided on a compatible conjugative plasmid. Similarly, non-conjugative, mobilizable plasmids have been shown to have *oriT* regions, several of which have been cloned, e.g. ColE1 (Warren *et al.*,1978), pMB1 (Covarrubias *et al.*,1981; Finnegan and Sherratt, 1982) and RSF1010 (K.Derbyshire and N.Willetts, unpublished data). *oriT* also seems to be the region at which protein binds to form a relaxation complex after introduction of a single strand nick. In the best studied example, ColE1, the relaxation complex is implicated in conjugation, since mutants deficient in forming the complex are mobilized inefficiently (Collins *et al.*,1978; Dougan and Sherratt, 1977; Inselburg,1974). Further, the site at which the mobility proteins act has been mapped, using *cis*-acting mutations that prevent mobilization, and shown to lie at, or near to, the *nic* site (Warren *et al.*,1978). Hence *nic* is presumably the same as *oriT* and may be involved in the initiation of plasmid transfer.

There appears to be little DNA sequence homology amongst the *oriT* regions of non-conjugative plasmids, and even of the conjugative plasmids able to mobilize them. Also, non-conjugative plasmids are not mobilized with the same efficiency by conjugative plasmids from different incompatibility groups (Willetts and Crowther,1981). R300B, however, is mobilized very efficiently by IncP plasmids in *E.coli* and *P.aeruginosa* (Willetts and Crowther,1981) and is also mobilized into a number of different organisms by IncFI, IncFII,

IncIα, IncN and IncW plasmids, although the IncP plasmid was the only one effective in all crosses, (Barth *et al.*,1981). The broad spectrum of plasmids capable of mobilizing R300B may be a reflection of its evolutionary development to become a widely distributed entity.

While the efficiency of mobilization of R300B and its derivatives by IncIα plasmids was not as high as that of IncP plasmids, it was observed that, particularly for mobilization to *M.methylotrophus*, the IncIα plasmids held an advantage in that they could never be detected in the recipient organism. Clearly this was a significant step in the development of a host-vector system based on *M.methylotrophus*, by providing a clean delivery of the plasmid vector.

Thus R300B and derivatives which carry its *ori*T sequence are readily mobilized to a number of different Gram-negative organisms by by several different conjugative plasmids. For the specialized application of using *M.methylotrophus* as a host in which to study expression of cloned coding sequences, IncIα plasmids provide a delivery system uncomplicated by concomitant introduction of a second plasmid.

# CHAPTER 2

## Materials and methods

### 2.1 Bacteria and their growth conditions

**_E.coli_ Strains**

BE6; F$^-$,*hsd*R,*met*,*trp*E. (B.Ely).

BW85; ColI$^r$,*deo*B,*leu*B6,*rps*L,*thi*1,*thy*A. (B.Wilkins).

BW162; F$^-$,ColI$^r$,*gln*V,*lac*Y1,*leu*B6,*thi*1,*thr*1,*ton*A21,*sup*E44. (B.Wilkins).

C600; F$^-$,*gln*V,*lac*Y1,*leu*B6,*thi*1,*thr*1,*ton*A21,*sup*E44. (B.Wilkins).

DS903; *ara*14,*arg*E3,*his*4,*lac*Y1,*leu*B6,*mtl*1,*pro*A2,*rec*F,*rps*L,*sup*37, *thi*1,*thr*1,*tsx*33,*xyl*5. (D. Sherratt).

ED8654; *hsd*R,*met*,*sup*E,*sup*F. (W.J.Brammar).

HB101; F$^-$,*hsd*M,*hsd*R,*hsd*S20,*pro*A2,*ara*14,*gal*K2,*lac*Y1,*mtl*1,*xyl*5, *rec*A13,*rps*L,*sup*E44. (D.W.Burt).

JA221; *hsd*R,*lac*Y1,*leu*B6,*rec*A,*trp*E5. (J.D.Windass).

KM4196; *rec*A.(I.Hickson)

NM36; *pol*A1,*sup*$^o$. (W.J.Brammar).

PA340; F$^-$,*ara*14,*arg*H1,*gal*6,*gdh*1,*glt*B31,*his*1,*lac*Y1,*leu*B6,*mal*A1,*mtl*2, *rps*L,*sup*E44,*thi*1,*thr*1,*ton*A2,*xyl*7,$\lambda^r$,$\lambda^-$. (D.Pioli)

Other bacteria

*Alcaligenes eutrophus*

5M-2; NCIB N$^o$. 10442 (D.Byrom).

*Methylomonas methylovora* NACC N$^o$.2731 (D.Byrom).

*Methylophilus methylotrophus*

AS1; w.t. NCIB N$^o$. 10515 (D.Byrom).

CMB13; restrictionless AS1. (A.C.Boyd).

*Pseudomonas aeruginosa*

PAO1 w.t. (W.J.Brammar)

All bacterial strains, with the exception of *M.methylotrophus*, could be cultured in Luria broth (L-broth) or on L-agar plates.

1 litre of L-broth contained;

10g Difco Bacto Tryptone

5g Difco Yeast Extract

5g Sodium chloride

1g Glucose

with the pH adjusted to 7.2.

1 litre of L-agar contained;

10g Difco Bacto Tryptone

5g Difco Yeast Extract

5g Sodium chloride

with the pH adjusted to 7.2.

15g Difco agar.

For *in vivo* labelling of DNA with radioactive thymine or when growing cells in the presence of sulphonamide or trimethoprim, a minimal medium was used.

1 litre of minimal medium contained;

2g Ammonium sulphate

14g Dipotassium hydrogen phosphate

6g Potassium dihydrogen phosphate

1g Trisodium citrate

0.2g Magnesium sulphate.$7H_2O$

0.2% Glucose

Minimal medium was supplemented with amino acids as required at 40μg/ml, with thiamine at 4μg/ml and with thymine at 20μg/ml. Minimal agar plates were made by solidifying minimal medium with

15mg/ml of Difco Agar. Isosensitest (Oxoid) broth or agar were prepared simply by addition of the required amount of powder to distilled water and autoclaving.

*M.methylotrophus* was grown in methanol minimal medium (MMM) or on methanol minimal agar (MMA) plates.

1 litre of MMM contained

1.8g Ammonium sulphate

0.2g Magnesium sulphate.$7H_2O$

1.4g Sodium dihydrogen phosphate.$2H_2O$

1.9g Dipotassium hydrogen phosphate

(These constituted a ×1 stock salts solution).

5ml Methanol

1ml Trace element solution *

* 1 litre of trace element solution contained

980mg Ferric chloride

20mg Cuprous sulphate.$5H_2O$

100mg Manganous sulphate.$4H_2O$

100mg Zinc sulphate.$7H_2O$

1.8g Calcium carbonate

36.6ml 1M hydrochloric acid.

MMA was made by solidifying MMM with 15mg/ml Difco Agar.

When required antibiotics were added at the following levels, (where different, the levels for *M.methylotrophus* are shown in brackets).

Ap 30-100 (100); Cm 25; Km 10 (25); Sm 15; Su 500-1000**; Tc 10 (2-10); Tp 10 (100)** μg/ml.

**These antibiotics cannot be used in L-broth or in L-plates.

## 2.2 Preparation of dialysis tubing

All dialysis tubing, used either for dialysis or electroelution, was treated as follows:-

The tubing was cut into convenient lengths (≈30cm) and boiled for 10min in a large volume of 2% sodium bicarbonate and 1mMEDTA. After rinsing in distilled water, the tubing was boiled for a further 10min in fresh distilled water, before being stored at 4°C in 50% ethanol.

Before use the tubing was washed inside and out with distilled water, gloves being worn on all occasions when the tubing was handled.

## 2.3 Preparation and storage of phenol

The "phenol" referred to in this thesis was prepared for use either from redistilled crystalline phenol or from freshly opened liquid phenol (80% w/w in water). The latter was dispensed into smaller lots (100ml) and stored at -20°C until needed.

8-hydroxyquinoline was added to redistilled phenol or thawed liquid phenol to a final concentration of 0.1%. This compound, which coloured the phenol yellow (providing a convenient identification for the phase), is a partial inhibitor of RNase and a chelator of metal ions (Kirby,1956). The phenol was extracted with an equal volume of 0.5M Tris-HCl pH 8.0 then twice with equal volumes of 0.1M Tris-HCl pH 8.0 and stored at 4°C under equilibration buffer.

## 2.4 Plasmids and plasmid DNA isolation

### Plasmids

| Name | Incompatibility | Phenotype (in *E.coli*) | Source |
|---|---|---|---|
| pBR322 | - | ApTc | B.Ely |
| pBR328 | - | ApCmTc | B.Wilkins |
| pLG221 | IncIα | KmNmSm | G.Boulnois |
| pTB92 | IncQ | KmSmSu | P.T.Barth |
| R64*drd*11 | IncIα | SmTc | B.Wilkins |
| R144*drd*3 | IncIα | Km | P.T.Barth |
| R300B | IncQ | SmSu | P.T.Barth |
| R483 | IncIα | SmSpTp | P.T.Barth |
| R751 | IncP | Tp | D.Pioli |
| RP4 | IncP | ApKmTc | P.T.Barth |
| S-a | IncW | CmKmSmSu | N.J.Grinter |

All plasmids with the prefix "pGSS-" are described in this thesis.

## 2.5 Isolation of plasmid DNA

### 2.5a)Preparative scale.

This procedure was routinely undertaken to produce plasmid DNA, in quantities up to several milligrams, for restriction analysis or for the subsequent preparation and isolation of specific plasmid fragments. The scale was dependent on the ultimate use of the DNA, but also on the ultracentrifuge rotor available. Culture sizes were

standardized to 250ml (for Beckmann 50Ti or Sorvall T865.1 rotors) or 500ml (for Sorvall T865 or TV865 rotors). The experimental methods are identical with the larger scale cultures using twice the quantities of the 250ml cultures, described below.

Plasmid-containing cells from a 250ml L-broth stationary phase culture (normally about 16hr growth from a 1:100 inoculum), with antibiotic selection, were harvested by centrifugation in a Sorvall GS-A rotor for 10min at 6000rpm. The supernatant liquid was decanted off and the cell pellet resuspended in 3ml of ice-cold 25% sucrose in 50mM tris(hydroxymethyl)aminomethane [Tris], pH8.0. To this mixture was added 0.5ml of a freshly prepared lysozyme solution (10mg/ml in water) and after 15min incubation on ice, 1ml of 0.25M ethylene-diamine tetra-acetic acid (EDTA) pH8.5 was added. After a further 15min on ice, cell lysis was brought about by the addition of 4ml of 0.1% Triton X-100 in 50mM Tris.HCl pH8.0, 62.5mM EDTA pH8.5. The lysate was then cleared by centrifugation at 18,000rpm for 30min in a Sorvall SS-34 rotor. The supernatant was recovered, and 7.4ml of this "cleared lysate" were mixed with 7.1g of caesium chloride and 270$\mu$l of ethidium bromide solution (10mg/ml in water). This mixture was transferred to a 13.5ml polyallomer centifuge tube, overlayed with liquid paraffin, and capped. The equilibrium density gradient separation was achieved by centrifugation for 40hr at 40,000rpm.

Plasmid DNA was usually visible as a pink/red band just below the centre of the caesium gradient, and was accompanied by a less intense band of chromosomal DNA, a few millimeters above. The plasmid DNA was collected using a syringe and 18-gauge needle which was pushed through one side of the polyallomer tube, just below the plasmid DNA band, and the DNA gently withdrawn.

The ethidium bromide was removed by two extractions with isopropanol equilibrated with a saturated caesium chloride solution,

and the DNA solution was then transferred to a dialysis bag. The caesium chloride was removed by dialysis against 20 vol of TE buffer (10mM Tris.HCl, 1mM EDTA pH7.2), for 2-3hr, and including three buffer changes. On completion of the dialysis, the DNA solution was extracted twice with equal volumes of phenol, then three times with diethyl ether. Excess ether was blown off with compressed air, the solution made 0.3M in sodium acetate, and the DNA precipitated by the addition of 0.54vol of isopropanol and chilling for several hours at -20°C. The DNA was collected by centrifugation at 12,000rpm for 15min, and the pellet dried *in vacuo* before resuspending in 200µl of water and determining the $OD_{260}$ to give a measure of the concentration from

$$[DNA]\mu g/ml \simeq 50 \times OD_{260}$$

Typically 100-300µg of plasmid DNA were obtained from one 250ml culture.

**2.5b) Small scale plasmid DNA preparations.**

It is often convenient and sometimes necessary to prepare a small amount of a plasmid, to assess the size of a plasmid or even to screen plasmids for the presence of a particular restriction site or fragment. Throughout the duration of this thesis, three different methods have been employed, although I now have a preference for the rapid boiling method (Holmes and Quigley,1981). All three methods work equally well for *E.coli* and *M.methylotrophus*.

i) Mini SDS-cleared lysates (Broome-Smith,1980)

Cells from single colonies were streaked across an L-agar plate, and after overnight incubation, the bacterial mass was resuspended in 1.5ml Eppendorf tubes containing 25% (w/v) sucrose in 50mM Tris.HCl (pH8.0), 10mM EDTA. Cells were lysed in 0.5% sodium

dodecyl sulphate (SDS), and cleared lysates obtained by spinning samples for 12min in an Eppendorf centrifuge. Then 40μl samples of SDS-cleared lysates were subjected to 1% agarose gel electrophoresis, run for 2-3hr at 100V, and stained in 5μg/ml ethidium bromide before being photographed.

**ii) Alkaline extraction procedure** (Birnboim and Doly,1979)

Total cellular DNA was denatured with sodium hydroxide, then the plasmid DNA allowed to rapidly renature upon neutralization, while chromosomal DNA formed an insoluble network. Most protein and ribosomal RNA was removed by precipitation with SDS and high salt.

1ml of a 5ml L-broth overnight roller culture was centrifuged for 15sec in a 1.5ml Eppendorf tube. The tightly packed cell pellet was resuspended in 100μl of lysis solution (2mg/ml lysozyme; 25mM Tris.HCl pH8.0; 10mM EDTA pH8.0; 50mM glucose) by vortexing. 200μl of alkaline SDS solution (1% SDS in 0.2M NaOH) were added and the suspension immediately became clear and slightly viscous. After 5min incubation at 0°C, 150μl of high salt solution (3M sodium acetate pH4.8) were added and mixed by inversion. A coarse white precipitate formed, and incubation was continued at 0°C for 60min. The precipitate was recovered by spinning for 5min in an Eppendorf and 400μl of supernatant were removed. (It was important to avoid contamination with any floating precipitate). 1ml of cold ethanol was added to the supernatant which, after 30min at -20°C, was centrifuged for 2min. The supernatant was discarded, the pellet taken up in 100μl of acetate/Tris solution (0.1M sodim acetate in 50mM Tris.HCl pH8.0) and re-precipitated with 200μl of cold ethanol at -20°C for 10min. After centrifuging for 2min , the pellet was dissolved in 40μl of water and analysed by gel electrophoresis.

Plasmid DNA prepared by this method was suitable for

transformation or restriction, though in the latter case it was sometimes necessary to treat the sample with RNase, especially if looking for a small fragment which could be masked by the RNA smear.

iii) Rapid boiling method (Holmes and Quigley,1981)

This method was particularly suitable for analysis of plasmid content from single colonies or small streaks from single colonies. In my hands it gave the best results when scaled up slightly from the original method and when small overnight cultures were used.

Single colonies were toothpicked into 500µl of L-broth in 1.5ml Eppendorf tubes and incubated at 37°C overnight. The cells were pelleted by spinning in an Eppendorf centrifuge for 15sec and were then resuspended in 100ml of STET buffer (8% sucrose; 5% Triton X-100; 50mMEDTA; 50mM Tris.HCl pH8.0). 8µl of lysozyme solution (10mg/ml in water) were added and the mixtures placed in boiling water for 60sec then immediately spun in an Eppendorf centrifuge for 10min. The supernatant was carefully removed (or the gelatinous pellet was removed from the supernatant with a Gilson pipette) and an equal volume of isopropanol added. After chilling at -20°C for 10min, the plasmid was collected by centrifugation for 5min, dried *in vacuo* and resuspended in 20µl of water. 5µl aliquots could be used for analysis by agarose gel electrophoresis or digested with restriction enzymes, in which case the addition of RNase (4µg/ml) may be necessary.

**2.6 Plasmid DNA manipulation and analysis**

All enzymes used were obtained from commercial sources, with the exception of the T4 DNA ligase used in the preparation of all

recombinant plasmids up to and including pGSS15. This enzyme was purified and kindly given to me by Barry Ely and Ron Wilson.

a) Restriction enzymes

Restriction enzymes are endonucleases, usually isolated from prokaryotes, that are able to recognize specific sequences within double-stranded DNA. They fall into three classes, Types I, II and III, but only Type II enzymes are relevant to this thesis.

Type II restriction/modification systems consist of a separate restriction endonuclease and a modification methylase. A large number of Type II enzymes have been isolated (Roberts,1984), many of which are useful in molecular cloning. These enzymes cut DNA within, or near to, their particular recognition sequence, typically four or six nucleotides in length with a twofold rotational symmetry.

Restriction enzymes can generate DNA fragments with protruding 5'-tails, 3'-tails, or cleave at the axis of symmetry to produce blunt ends. These enzymes, because of their individual recognition sequences, allow termini generated with a particular enzyme to be base-paired with any other similarly cut DNA, to form novel recombinant molecules.

At the start of this work all restriction endonucleases were used in buffer cocktails, specific for a particular enzyme, as prescribed by the manufacturers. More recently a multipurpose buffer, "core buffer", has been supplied by Bethesda Research Laboratories (BRL) in which many different restriction enzymes will function. This has been particularly useful for double digests, which can now be performed simultaneously. A further refinement to a totally universal enzyme buffer has been made (O'Farrell *et al.*,1980) and this has now largely superceded the individual cocktails.

O'Farrell Buffer

33mM Tris acetate pH7.9

66mM Potassium acetate

10mM Magnesium acetate

0.5mM Dithiothreitol (DTT)

100μg.ml Bovine serum albumin (BSA) nuclease-free

Stored as a 10X stock at -20°C

**b) T4 DNA ligase**

In duplex DNA, T4 DNA ligase catalyses the repair of single-stranded nicks with adjacent 3'-hydroxyl and 5'-phosphate termini, as well as the end-to-end joining of DNA fragments with staggered (sticky) or blunt (flush-ended) termini. The enzyme, a single polypeptide of 68,000 daltons, requires $Mg^{++}$ and ATP as cofactors (Weiss *et al.*,1968).

The T4 DNA ligase from B.Ely and R.Wilson was used empirically at 1-3μl per microgam of DNA, in the following buffer cocktail:

66mM Tris.HCl pH7.5

1mM EDTA pH8.0

10mM $MgCl_2$

10mM 2-mercaptoethanol

0.1mM ATP

The T4 DNA ligase as supplied by BRL was used as recommended at ≈0.1 unit per ≈10pmol of ends, for sticky end fragments, and 1 unit per ≈10pmol of ends, for blunt-end fragments, in the following buffer:

T4 DNA ligase buffer

66mM Tris.HCl pH7.5

6.6mM $MgCl_2$

10mM DTT

0.4mM ATP

Incubation conditions varied greatly depending on the enzyme source, type of joining, or for experimental convenience, and have been included in the descriptions of the experiments.

c) T4 Polynucleotide Kinase

This enzyme catalyses the transfer of the $\gamma$-phosphate group of ATP to the 5'-hydroxyl terminus of DNA, RNA and synthetic polynucleotides. Activity requirements include $Mg^{++}$ as a cofactor and DTT as a reducing agent. The use of $\gamma$-$^{32}$P ATP allows transfer of the labelled phosphate group to produce an end-labelled nucleic acid molecule.

The conditions for use of this enzyme were as described by Maniatis *et al.*,1982.

## 2.7 Agarose Gel electrophoresis

This is one of the most powerful and extensively used techniques for the analysis of mixtures of DNA fragments.

In general, gels were cast 3-4mm thick on special Perspex plates, 240mm long and 130mm wide, that had 5mm side-walls and specially

designed comb holders. Prior to pouring the gel, adhesive tape was fixed across each end of the plate to contain the molten agarose, and the appropriate comb was placed in the holders. The whole assembly was then laid on a flat surface.

The correct amount of powdered agarose (Litex - Denmark) was added to a measured quantity of Tris acetate gel electrophoresis buffer,

40mM Tris acetate
20mM sodium acetate } pH 7.8
0.2M EDTA

and dissolved by heating in a microwave oven. When the solution had cooled to about 50-60°C, it was poured onto the prepared plate and allowed to set (30-40min at room temperature). The comb and tape were carefully removed and the gel placed in the electrophoresis tank, (a Perspex box with an electrode at each end and divided in half by a low Perspex wall which supported a Perspex platform, upon which the gel was positioned). Electrophoresis buffer was added until it just covered the whole gel by 1-2mm. Samples were mixed with one tenth of a volume of loading buffer (20% Ficoll and 0.025% bromophenol blue tracking dye in gel buffer) and then applied to the gel slots using a Gilson pipette.

The strength of a gel was determined by the size of the fragments to be separated, and the following table was used as a guide:

| Agarose % | Separation of linear molecules in kb |
|---|---|
| 0.3 | 60 - 5 |
| 0.6 | 20 - 1 |
| 0.7 | 10 - 0.8 |
| 0.9 | 7 - 0.5 |
| 1.2 | 6 - 0.4 |
| 1.5 | 4 - 0.2 |
| 2.0 | 3 - 0.1 |

## 2.8 Recovery of DNA from agarose gels

It is important to note that only the highest quality agarose is suitable for any of the methods described below if the DNA is required for further manipulations involving enzymes. Most grades of agarose contain sulphated polysaccharides which are extracted from the gel together with the DNA; these substances are potent inhibitors of most of the enzymes commonly used in subsequent cloning steps.

### 2.8a) Electroelution into dialysis bags. (McDonnel *et al.*,1977)

The ethidium bromide-stained gel was visualised under long-wave U.V. light (to minimize damage to the DNA) and the appropriate band cut out, using a sharp scalpel, and slipped into a dialysis bag which was sealed at one end with a clip. The bag was filled with buffer and the gel slice allowed to sink to the bottom. Most of the excess buffer and any air bubbles were squeezed out of the bag and a second clip applied to seal the other end. The sealed bag was immersed in a shallow layer of half-strength buffer in an electrophoresis tank and the DNA electroeluted at 100v for 1-3hr. During this time the DNA migrated to the inner wall of the bag. The polarity of the current was reversed for 30sec to release the DNA from the bag wall into the buffer. The dialysis bag was opened and all the buffer surrounding the gel slice was carefully removed. The DNA was then purified by phenol-extraction, after the solution had been passed over polyallomer wool to remove any particles of agarose, and then ethanol precipitated.

**2.8b) Electrophoresis onto DEAE-cellulose paper.** (Dretzen *et al.*,1981)

The ethidium bromide-stained gel was visualised under U.V light as above. Using a sharp scalpel, slits were made in the gel parallel

to and just below the required bands. Small strips of Whatman DE-81 DEAE-cellulose paper were inserted into the slits and the gel squeezed firmly against the papers to close the incision. The gel was then returned to the tank and electrophoresis resumed until the DNA had entered the paper strips; this was verified by observation under U.V. illumination. The method allows for recovery of bands that are close together, and by inserting paper strips above bands of interest, contamination by larger fragments is eliminated.

The paper into which the DNA has been electrophoretically transferred was removed from the gel and washed with cold water. The papers could be kept in water for several hours at 4°C unless they contained very small fragments. The papers were then drained and dried by blotting onto filter paper, then transferred to 1.5ml Eppendorf tubes and 300-700$\mu$l of elution buffer per 50mm$^2$ of paper added (elution buffer: 20mMTris.HCl pH7.5; 1mM EDTA; 1.5M NaCl). The papers were shredded by vortexing and incubation at 37°C for 2hr with occasional agitation. The mixture was transferred to a fresh tube containing a plug of polyallomer wool (to trap the paper fibers) and a small hole made in the bottom of the tube with a fine needle. This tube was then placed into the top of a plastic tube, in which it was a snug fit, and after centrifugation for 5min at 10krpm, the eluate was extracted with 3 volumes of *n*-butanol saturated with water, and the DNA precipitated by addition of 2 volumes of ethanol.

Reported recoveries by this method range from 60% for a 20kb fragment to about 80% for fragments of 2kb or less.

Both of the above methods are also suitable for recovery of DNA from polyacrylamide gels.

## 2.9 Polyacrylamide gel electrophoresis

Polyacrylamide gels were used to analyse and prepare fragments of DNA less than 1kb in length (Maniatis *et al.*,1975). In this thesis I have exclusively used gels with a polyacrylamide concentration of 5% for effectively separating fragments in the range 80-500bp.

Stock solutions

1. 30% acrylamide

| | |
|---|---|
| acrylamide | 29g |
| N,N'-methylene bisacrylamide | 1g |
| water | to 100ml |

2. Tris borate electrophoresis buffer (TBE).

0.089M Tris-borate

0.089M boric acid

0.002M EDTA

A concentrated 10× stock contained (per litre)

108g Tris base

55g boric acid

40ml 0.5M EDTA pH 8.0

3. 1.6% Ammonium persulphate

| | |
|---|---|
| ammonium persulphate | 0.16g |
| water | to 10ml |

Short gels (20×20cm) for analyses required 50ml of acrylamide solution, long gels (40×20cm), or short preparative gels, required 100ml.

Quantities for 100ml of acrylamide solution

| | |
|---|---|
| 16.66ml | 30% acrylamide |
| 71.6ml | water |
| 1.66ml | 1.6% ammonium persulphate |
| 10ml | 10× TBE |

The ingredients were mixed in a Buchner flask and suction applied to deaerate the solution.

The gel plates, one plain and one notched, were carefully cleaned with detergent and dried. Spacers were placed down either side and along the bottom between the plates to form a mould for the gel. The plates were sealed all round with electrical tape.

100$\mu$l of TEMED (N,N,N',N'-tetramethylethylene diamine) were added to the deaerated solution, swirled gently, then poured into the space between the two plates. When the space was almost full to the top, the appropriate comb was inserted, taking care not to trap any air bubbles under the teeth. The gel was allowed to polymerize at room temperature for up to 1hr. The comb was removed and the wells flushed out with water.

The tape and spacer were removed from the bottom of the gel and the assembly attached to the electrophoresis tank with large bulldog clips. The reservoirs were each filled with 400ml of 1× TBE and a bent syringe needle used to ensure good buffer contact with the bottom of the gel by removing air bubbles. Loading buffer (20% ficoll in TBE + bromophenol blue or xylene cyanol) was added to the DNA samples at one tenth of a volume, then the samples were loaded into the gel slots with a Gilson pipette or a drawn out glass capillary. Electrophoresis was usually for 30v overnight. (In a 5% polyacrylamide gel, bromophenol blue migrates at the same rate as a 65-bp DNA fragment, xylene cyanol the same as a 260-bp fragment).

At the end of a run the gel assembly was detached from the tank, the tape removed and the plates separated. The spacers were removed and the gel, on one plate, was stained and photographed exactly as described for agarose gels.

## 2.10 Transformation of *E.coli* by plasmid DNA.

Most transformation methods are based on the original observation of Mandel and Higa (1970) that calcium-treated cells were able to take up bacteriophage DNA. The method was later shown to be also applicable to plasmid DNA (Cohen *et al.*,1973).

Once inside the cell, the plasmid DNA replicates and expresses any phenotypic markers that it carries (usually resistance to antibiotics). Plasmid-containing cells can then be selected as resistant colonies growing in the presence of the antibiotic.

Thus, 100ml of L-broth in a 500ml flask were inoculated with 1ml of a fresh overnight culture and incubated at 37°C with vigorous shaking for 100-200min to an $OD_{630}$ of ≈0.5. The culture was chilled on ice for 10min then harvested by centrifugation at 4krpm for 5min at 4°C. The pellets were resuspended in half a volume of an ice-cold, sterile solution of 50mM $CaCl_2$ in 10mM Tris.HCl pH8.0 and incubated on ice for 15min. After centrifugation as before, the pellets were resuspended in one fifteenth of a volume of the calcium buffer.

200μl aliquots were dispensed into pre-chilled tubes and stored at 4°C for 1-24hr, [it has been reported by Dagert and Ehrlich (1979) that a period of storage of 12-24hr increases the efficiency of transformation 4-6×]. Plasmid DNA (up to 40ng in up to 100μl of TE) was then added to the tubes which were then stored on ice for 30min. After a heat shock at 42°C for 2min the cells were dispersed in 5ml of L-broth and incubated at 37°C for 30min (for Tc selection) or 60min (for Ap or Km selection). This period allowed the bacteria to recover and to express antibiotic resistance. 200μl of cells were then spread onto selective agar plates and incubated at 37°C for 12-16hr.

## 2.11 Mobilization of IncQ-derivative plasmids

This is the key step required of a broad host range vector, the ability to be transferred from organism to organism. Generally the plasmids were transferred from an *E.coli* donor strain by plate matings on solid media.

A colony of the recipient strain was suspended in 50$\mu$l of sterile Spizizen salts and spread over half the surface of a selective medium plate. Colonies of the donor strain were then streaked in one direction across the plate into the recipient. The selective medium plates were designed so that only recipient cells containing plasmid from the donor were able to grow.

For many of the mobilizations of the plasmids described in this thesis I have used conjugative plasmids from the incompatibility group I$\alpha$, and this is discussed in detail elsewhere, but plasmids also used have included the IncP plasmids RP4 and R751, and the IncW plasmid S-a.

# CHAPTER 3

## Results

### Vector construction

#### 3.1 Preparation of $Ap^r$ derivatives of R300B.

Introduction of the $Ap^r$ gene from pBR322 into R300B, as a first step to expanding the latter's cloning capability , was approached by ligation of *Hae*II-digested pBR322 to partially *Hae*II-digested R300B. [The entire $Ap^r$ gene is contained within the largest *Hae*II fragment 1 of pBR322 (Fig.4)].

Thus, 1$\mu$g of pBR322 DNA, in 50$\mu$l, was digested to completion with 1 unit of *Hae*II at 37°C for 90min, the reaction stopped by heating to 70°C for 10min, then rapidly chilled on ice. 4$\mu$g of R300B DNA, in a 50$\mu$l reaction, were similarly treated with 1 unit of *Hae*II, but for only 60min. Agarose gel electrophoresis of the digests revealed only very partial cutting of the R300B DNA, and what appeared to be a significant proportion of full-length linear molecules. A 60$\mu$l ligation reaction was prepared to include 25$\mu$l each of the *Hae*II digested plasmids. After incubation at 12°C overnight, 10$\mu$l aliquots were used to transform *E.coli* strain ED8654 to ampicillin- resistance.

Several hundred $Ap^r$ colonies were obtained, and a sample of one hundred were picked off to determine which other antibiotic resistances were being expressed. 100/100 were $Ap^r$ and $Su^r$, 26/100 were $Sm^r$ and 3/100 were $Tc^r$. Single colony lysates of thirty isolates showed evidence of a plasmid band in only thirteen cases, and these were not of uniform size. Cleared lysate DNA preparations of sample plasmids from each size class showed that three plasmids were sufficiently large to suggest that they could be $Ap^r$ derivatives of R300B. These plasmids were designated pGSS1, pGSS2, and pGSS3.

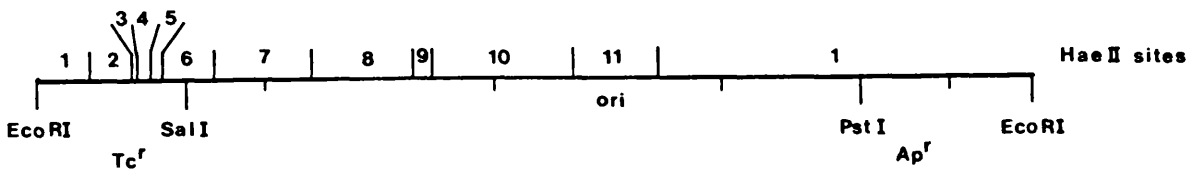



**Figure 4.**

A *Hae*II restriction map of pBR322. The plasmid is shown linearised at its *Eco*RI site, which lies within the large *Hae*II fragment 1. The fragments are numbered consecutively for position rather than for size.

All three carried the genes for resistance to ampicillin and sulphonamide, but pGSS3 gave only resistance to reduced levels of streptomycin, (2-5μg/ml), compared to pGSS1 and pGSS2 which determined resistance to 25μg/ml. The plasmid pGSS3, unlike pGSS1 and pGSS2, was unable to transform *E.coli* strain NM36 (*pol*A), which suggested that the replication region of R300B, or a region responsible for DNA polymerase I-independence, may have been disrupted. Subsequent gel electrophoretic analysis of pGSS3 DNA after *Hae*II digestion revealed the presence of a band corresponding to the 370-bp *Hae*II fragment 11 from pBR322. This fragment, which is contiguous with the $Ap^r$ fragment 1, carries the pBR322 origin of replication, and may not have been cleaved in the initial *Hae*II digestion. This would support the theory that the R300B replication region had been disrupted, and suggested that pGSS3 is using a pBR322 replication origin, [NM36 lacks the DNA polymerase I activity that is essential for replication from a ColE1 origin (Kingsbury and Helinski,1970)].

Restriction analysis of pGSS2 was consistent with the $Ap^r$ *Hae*II fragment 1 from pBR322 being inserted into R300B in the orientaion shown in Fig.5, which made it of limited use as a vector since removal of duplicated *Pst*I and *Eco*RI sites, for example, would not be easy.

Attempts to mobilize these plasmids into *M.methylotrophus* resulted in only one or two $Ap^r$ colonies being isolated in each case, but no plasmids could be detected suggesting that these colonies were spontaneous $Ap^r$ mutants. The fact that the plasmids were not found in *M.methylotrophus* may indicate that mobilization or maintainance functions on R300B had been disrupted, but, in the case of pGSS3, it was more likely to be due to a disruption of the R300B replication region. At this stage I decided to postpone further work with these

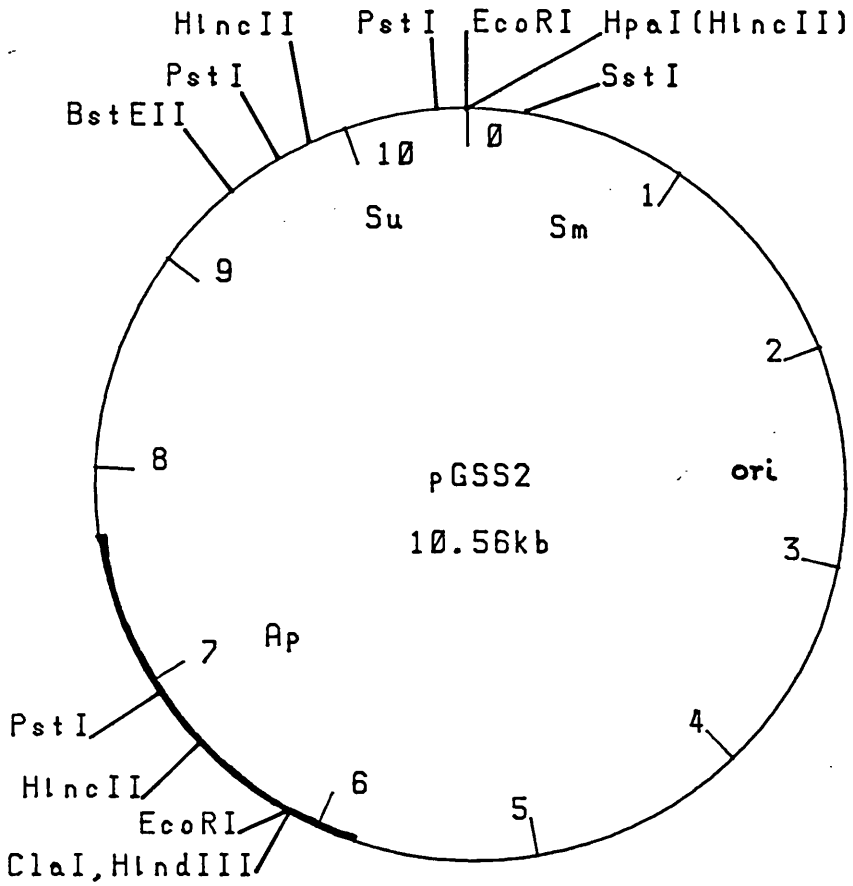



Figure 5.

A restriction map of pGSS2. This plasmid was derived from R300B by insertion of the $Ap^r$, *Hae*II fragment 1 from pBR322. The latter fragment is shown emboldened.

plasmids.

This preliminary experiment had however highlighted an important aspect of using partial digests, the potential to produce full-length linear plasmid molecules. A variety of such molecules, generated from different single cut sites in pBR322 and R300B, could be ligated together and transformed into *E.coli*. Selected colonies should contain relatively stable arrangements of cointegrate plasmids which could form a useful starting point from which to construct further vectors.

**3.2 Construction of co-integrates of pBR322 and R300B.**

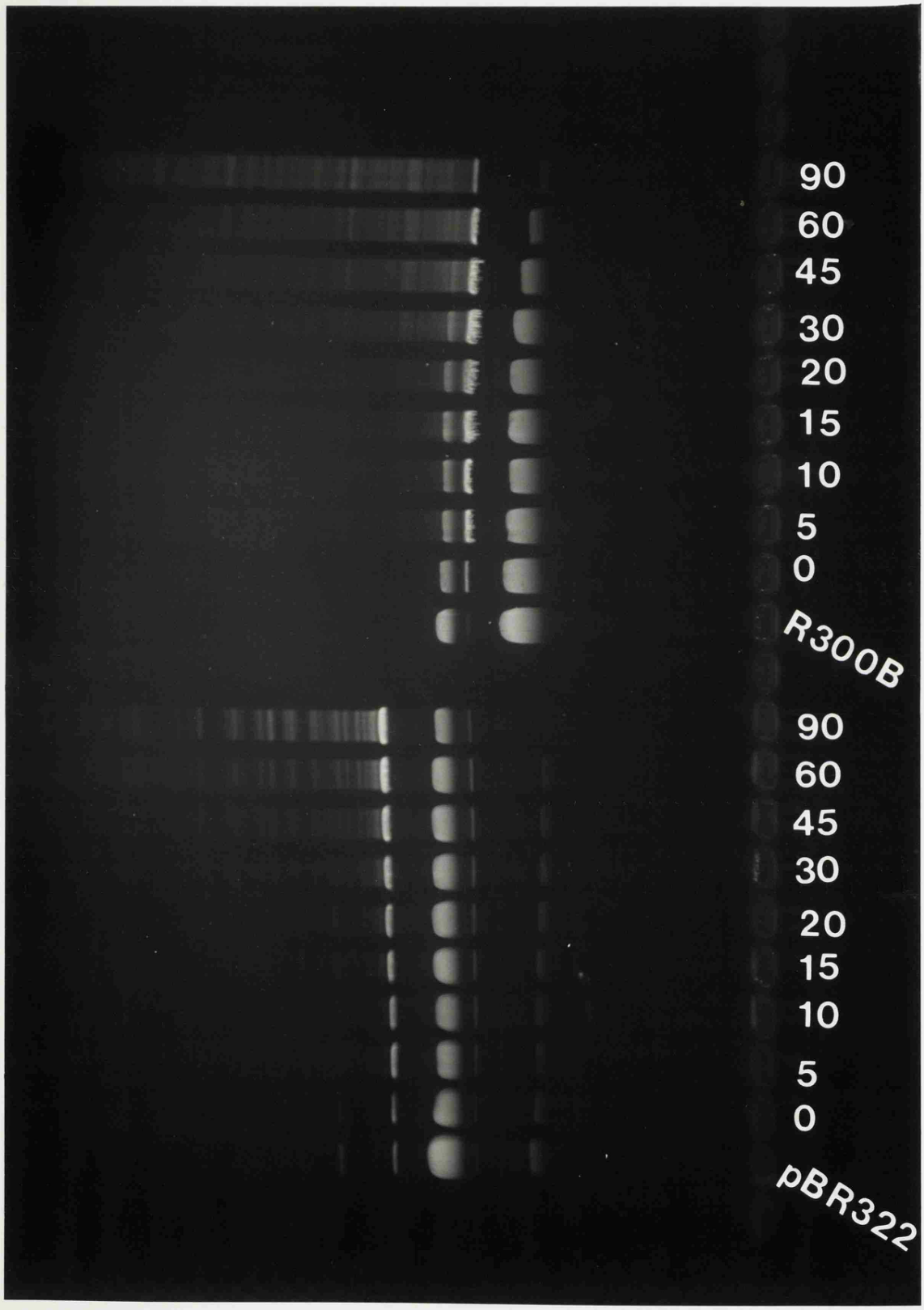
As it was not known how much of R300B was required to retain its broad host-range characteristics, digestion conditions were chosen which produced the highest proportion of full-length linear molecules as observed by agarose gel elecrophoresis of samples taken during the restriction enzyme reaction [Fig.6]. [It has been reported (Bagdasarian *et al.*, 1981), that only about 25% of the RSF1010 genome, principally the segment of antibiotic resistance genes, could be removed without deleterious effects.]

4μg of pBR322 DNA in 120μl of buffer were incubated at 37°C with 2 units of *Hae*II and a 20μl sample, taken immediately after the enzyme was added, was heated at 70°C for 10min then chilled on ice. Subsequent samples were removed during the incubation at 5, 10, 20 and 60min, and treated as before. 10μl of each sample were run out on an agarose gel to determine the level of digestion. In a similar way, 4μg of R300B DNA were treated with 0.5 unit of *Hae*II at 37°C for 30min and samples analysed as for pBR322.

Two 50μl ligation reactions were prepared, the first using 10μl of the "5min" pBR322 digest with 10μl of the "5min" R300B digest, and the second using 10μl of the "10min" pBR322 digest with 10μl of the

**Figure 6.**

Agarose gel electrophoretic analysis of partial *Hae*II digests of R300B and pBR322 plasmid DNAs. Samples were removed from limit digests of the plasmid DNAs at the time intervals indicated. After heating to stop the enzyme activity, a portion of each sample was subjected to agarose gel electrophoresis. The gel clearly shows the progress of the digestions.


90
60
45
30
20
15
10
5
0
R300B
90
60
45
30
20
15
10
5
0
pBR322

"0min" R300B digest. After incubation at 4°C overnight, the entire ligation mixtures were used for transformation of *E.coli* ED8654, with selection for resistance to tetracycline.

Approximately two thousand $Tc^r$ colonies were obtained from the first mix, and six hundred from the second. Replica plating to look for resistance to ampicillin and streptomycin showed all colonies to be $Ap^r$ but only twenty eight colonies were also $Sm^r$. After re-streaking, only six colonies were still $Ap^rTc^rSm^r$. Single colony lysates revealed only three possible co-integrate plasmids; the others appeared to contain products of a multimeric ligation or were double transformants. These three plasmids, designated pGSS4, pGSS5 and pGSS6, were all larger than R300B, with pGSS5 and pGSS6 appearing to be of equal size and larger than pGSS4. *Eco*RI digestion of pGSS4 gave two fragments of approximately 5.4kb and 4.8kb, indicating, with a plasmid size of only 10.2kb, that a nett deletion had occurred. Similar digestions of pGSS5 and pGSS6 produced identical patterns of two fragments, approximately 10kb and 3kb, consistent with the size predicted for a total co-integrate.

**3.3 Analysis of pGSS4**

The structure of the cointegrate plasmid pGSS4 was characterised by further restriction mapping. Following the identification of two *Eco*RI sites, unique *Sal*I and *Hind*III sites were identified. A double digest with *Eco*RI+*Sal*I showed that the *Sal*I site was located on the largest *Eco*RI fragment. The size reduction of this fragment caused it to co-migrate with the smaller *Eco*RI fragment on a 1% agarose gel, thus confirming their original size difference at about 650bp. *Pst*I produced three fragments, one of about 800bp equivalent to the small *Pst*I fragment from R300B that carries all, or most of, the $Su^r$ gene. Double digests with *Pst*I+*Sal*I

and *Pst*I+*Eco*RI showed that the largest *Pst*I fragment contained the *Sal*I site and that both *Eco*RI fragments contain *Pst*I sites. Further mapping with *Bst*EII+*Sst*I, *Hinc*II, *Pvu*II, *Ava*I, *Eco*RI+*Pvu*II, *Eco*RI+*Bst*EII, *Pvu*I and *Hae*II has allowed a map of pGSS4 to be drawn as shown [Fig.7].

Immediately apparent from the map is the extent and position of the deletion of R300B sequence that has occurred in this construction, from about 2.5 through to about 4.7 on the R300B map [Fig.3].

The plasmid pGSS4 was able to transform and replicate in *E.coli* NM36, a characteristic of the parental R300B, but was unable to be mobilized into *M.methylotrophus*. When colonies containing pGSS4 were subjected to analysis for plasmid by an alkaline SDS procedure (Birnboim and Doly, 1979), two properties were observed which distinguished it from R300B. Firstly its copy number appeared to be raised to a level more comparable with pBR322 and secondly, there was no evidence of a diffuse band that always appeared on gels, associated with R300B, when subjected to this type of analysis. This band is probably a result of R300B being able to form a relaxation complex, such that on treatment with alkaline SDS, any complexed molecules are disrupted leaving one strand of the DNA with a nick. Under the denaturing/renaturing conditions, the nicked strand will not be able to reassociate as quickly as any circular molecules and will thus leave some single-stranded circles or only partly re-associated plasmid molecules. These will migrate faster than duplex, super-coiled plasmid when subjected to electrophoresis, and produce the diffuse band as observed. This is such a distinct property [Fig.8], that it could be used diagnostically to look for plasmids capable of forming relaxation complexes.

Thus pGSS4 was behaving like pBR322, but with the ability to

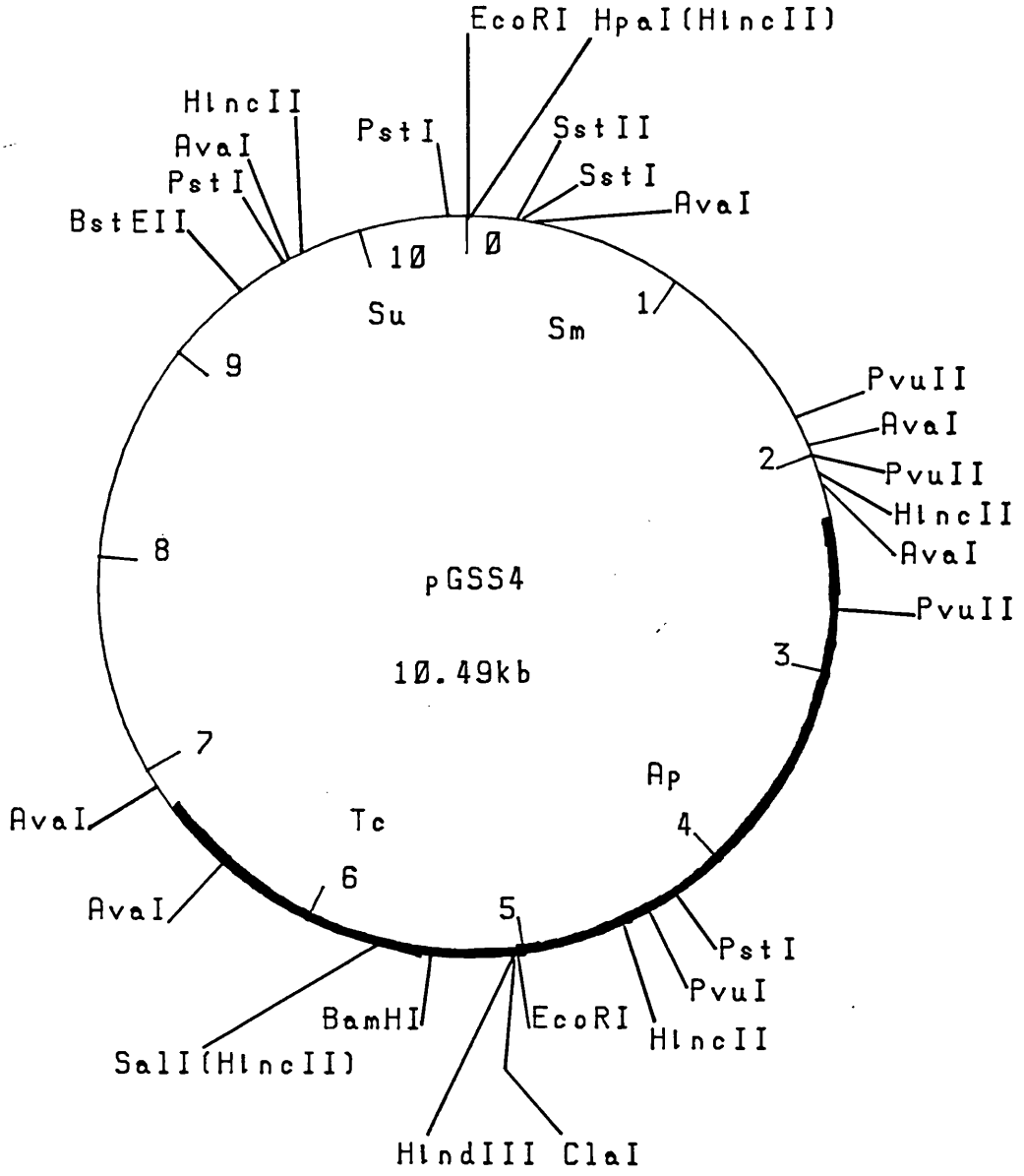



**Figure 7.**

A restriction map of pGSS4. This plasmid was constructed by ligation of *Hae*II-linearised pBR322 DNA with partially *Hae*II-cut R300B DNA. The sequences derived from pBR322 are shown emboldened and, by comparison with Fig.3, the extent of the R300B-sequence deletion can be seen.

**Figure 8.**

Detection of plasmids by the alkaline SDS method (Birnboim and Doly,1979). Track 1. pBR322, 2. R300B, 3. pGSS15, 4. pGSS6, 5. pGSS4. All plasmids were in *E.coli* strain NM36, except for pBR322 which was in ED8654. The most intense bands in tracks 1 and 5 are the CCC bands from plasmids pBR322 and pGSS4, respectively. The arrowed bands in the centre tracks, below the CCC bands, only occur when the plasmids are mobilizable and may be used to form the basis of a screen for this property.

replicate independently of DNA polymerase I. This was further evidenced by the fact that a derivative of R300B, pTB92, encoding resistance to kanamycin (Barth *et al.*,1981), was compatible with pGSS4. Since pGSS4 seemed to be replicating from a pBR322 origin, it would be interesting to see if a gene product was being made, encoded by the R300B-derived portion of the plasmid, that was similar to DNA polymerase I, giving rise to the observed independence.

The *pol*A genotype is associated with an increased sensitivity to U.V., and I decided to check if a *pol*A strain of *E.coli* carrying pGSS4 showed any increased resistance to U.V. over its plasmid-free counterpart. As controls, a *rec*A strain (KM4196) and a *pol*A$^+$ strain (ED8654) were included. Sections of fresh cross-streaks were subjected to different lengths of exposure to U.V. and then allowed to grow up overnight. ED8654 was unaffected, the growth of KM4196 was severely cut back and there was no detectable difference between NM36 with or without pGSS4. The suggestion was therefore, that whatever was allowing this plasmid to replicate in a *pol*A strain was not capable of complementing this particular mutation.

In summary, pGSS4 displayed properties of both parents in being high copy, non-mobilizable and compatible with pTB92, like pBR322, but DNA polymerase I-independent, like R300B.

**3.4 Analysis of pGSS6**

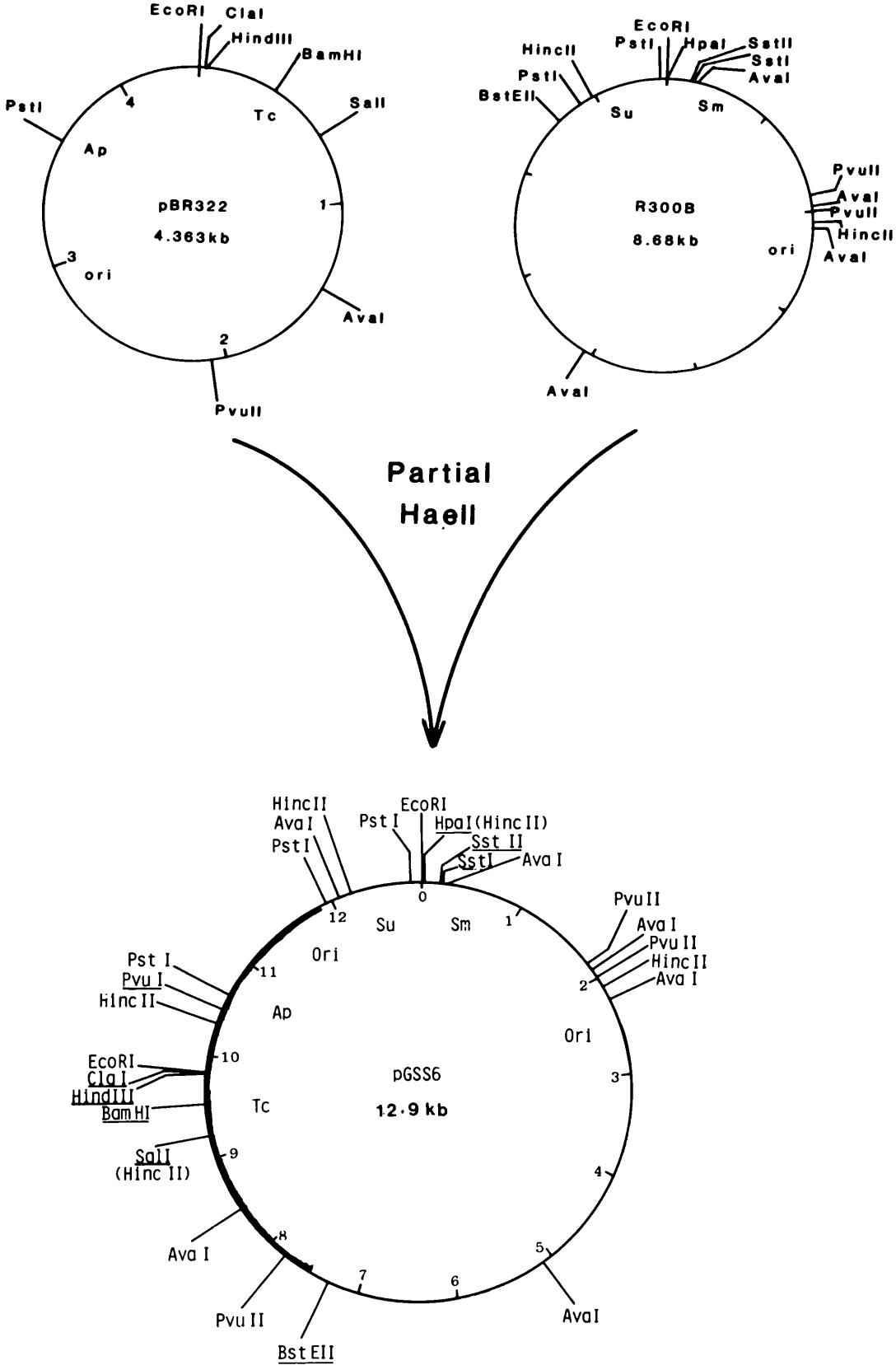
Since pGSS5 and pGSS6 appeared identical, from now on I have concentrated on the analysis of pGSS6 alone. This plasmid could be linearised by *Sal*I or *Hind*III and a double digest with *Eco*RI+*Sal*I showed that the *Sal*I site, and hence the Tc$^r$ gene, was contained on the larger *Eco*RI fragment. An *Eco*RI+*Pst*I digest indicated that there were three *Pst*I sites which all occurred within the smaller *Eco*RI fragment. From a *Bst*EII+*Eco*RI digest it was found that the *Bst*EII

site was on the larger *Eco*RI fragment. From this preliminary data it could be deduced that the pBR322 molecule had been inserted into the R300B molecule at a *Hae*II site located between the *Bst*EII site and the adjacent *Pst*I site. Also, the smaller *Eco*RI fragment carried $Ap^r$ and $Su^r$ genes, while $Tc^r$ and $Sm^r$ genes were on the larger *Eco*RI fragment. Further analysis with *Bam*HI, *Bst*EII+*Sst*I, *Hinc*II, *Eco*RI+ *Sst*I, *Ava*I, *Pvu*II and *Hae*II has allowed a restriction map to be constructed as shown [Fig.9]. The data were also consistent with pGSS6 being a total co-integrate of pBR322 and R300B.

The plasmid pGSS6 could transform *E.coli* strain NM36 and could be mobilized into *M.methylotrophus* by a number of different conjugative plasmids including RP4, S-a, and R64*drd*11. This indicated that the properties of R300B which make it a good candidate for a broad host-range cloning vector have been inherited intact. Insertion into the region between the *Bst*EII site and the adjacent *Pst*I site, did not appear to affect antibiotic resistance, replication, mobilization or broad host-range. Thus pGSS6, while having some applications as a cloning vector in its own right, appears to have more potential as a starting point for the development of new vectors.

Figure 9.

Construction of a total co-integrate plasmid, pGSS6, made from pBR322 and R300B following ligation of full length line ar molecules obtained by partial *Hae*II digestion. In pGSS6 the pBR322-derived sequence is shown emboldened and unique restriction sites are underlined.


EcoRI
ClaI
HindIII
BamHI
SalI
AvaI
PvuII
PstI
Ap
Tc
ori
pBR322
4.363kb
HincII
PstI
BstEII
EcoRI
HpaI
SstII
SstI
AvaI
Su
Sm
PvuII
AvaI
PvuII
HincII
AvaI
ori
R300B
8.68kb
Partial
HaeII
HincII
AvaI
PstI
EcoRI
HpaI(HincII)
Sst II
SstI
AvaI
PvuII
AvaI
PvuII
HincII
AvaI
Su
Sm
Ori
Ap
Ori
Tc
pGSS6
12.9 kb
Pst I
Pvu I
HincII
EcoRI
ClaI
HindIII
BamHI
SalI
(HincII)
AvaI
PvuII
BstEII
AvaI

# CHAPTER 4

## Resolution of cointegrate plasmids

### 4.1 Derivation of pGSS7

The restriction analysis of pGSS4 assigned the pBR322 origin of replication to the smaller *Eco*RI fragment, but because of the ability of pGSS4 to replicate in an *E.coli polA* strain, there was the possibility that a functional R300B origin of replication was still present. However, the large deletion of R300B sequence that had occurred during the construction of pGSS4, suggested that this was unlikely. To obtain more information, I decided to re-circularize the individual *Eco*RI fragments to see if they were capable of autonomous replication.

Thus 2$\mu$g of pGSS4 DNA, in a 100$\mu$l reaction, were digested with 5 units of *Eco*RI at 37°C for 90min, then the enzyme action terminated by heating at 70°C for 10min. After treating with T4 DNA ligase at 4°C overnight, at a low DNA concentration (3$\mu$g/ml) to favour re-circularization, *E.coli* strain ED8654 was transformed with selection for resistance to ampicillin and tetracycline. Only eighty five $Tc^r$ colonies were obtained, but all of these were also $Ap^r$ and probably arose from uncut parental molecules. This began to suggest that the $Tc^r$ gene could not be separated from the $Ap^r$ gene, on a viable plasmid, and implied that essential replication functions were contained on the smaller *Eco*RI fragment of pGSS4. This hypothesis was further endorsed by the isolation of approximately eight hundred $Ap^r$ colonies of which only about 10% were also $Tc^r$. Replica plating to look for $Sm^r$ revealed only weak and variable growth consistent with this gene not having its own promoter, but rather that it was relying on transcription which had initiated within the $Ap^r$ gene,

derived from pBR322, and to which it would now be juxtaposed. One of the stronger $Sm^r$ colonies was grown up and the plasmid isolated from it was designated pGSS7. Restriction analysis of pGSS7 showed it to have the structure predicted from the smaller *Eco*RI fragment of pGSS4 having been circularised, [Fig.10]. One striking property that became apparent in the preparation of pGSS7 DNA was the extremely high yield obtained, being about ten times a normal yield [Fig.11]. Thus the deleted section of pGSS4 sequence, made up from both R300B and pBR322 sequences, appeared to have been involved in plasmid copy number control. The copy number may actually need to be very high to produce the resistance to streptomycin if, for example, the transcription was not very efficient. This could be a contributing factor to the observed variation in resistance to streptomycin, although it is clearly not the whole explanation.

The plasmid pGSS7 was incapable of transforming a *pol*A strain of *E.coli*, was incompatible with pBR322 but compatible with R300B, and could not be mobilized into *M.methylotrophus*. The loss of DNA polymerase I-independence, occurring as a result of the construction of pGSS7 from pGSS4, would support the theory that there was a region of R300B responsible for that property and that it must have been located between 5.0 and 7.5 on the R300B map. Although pGSS7 was not a broad host-range plasmid it was interesting because of its properties, and potentially useful as a further high-yielding source of the $Ap^r$ gene fragment.

**4.2 Derivation of pGSS8 and pGSS9**

In a similar experiment to that described for pGSS4, pGSS6 DNA was digested with *Eco*RI, ligated under conditions of low DNA concentration, and the ligation products transformed into *E.coli*

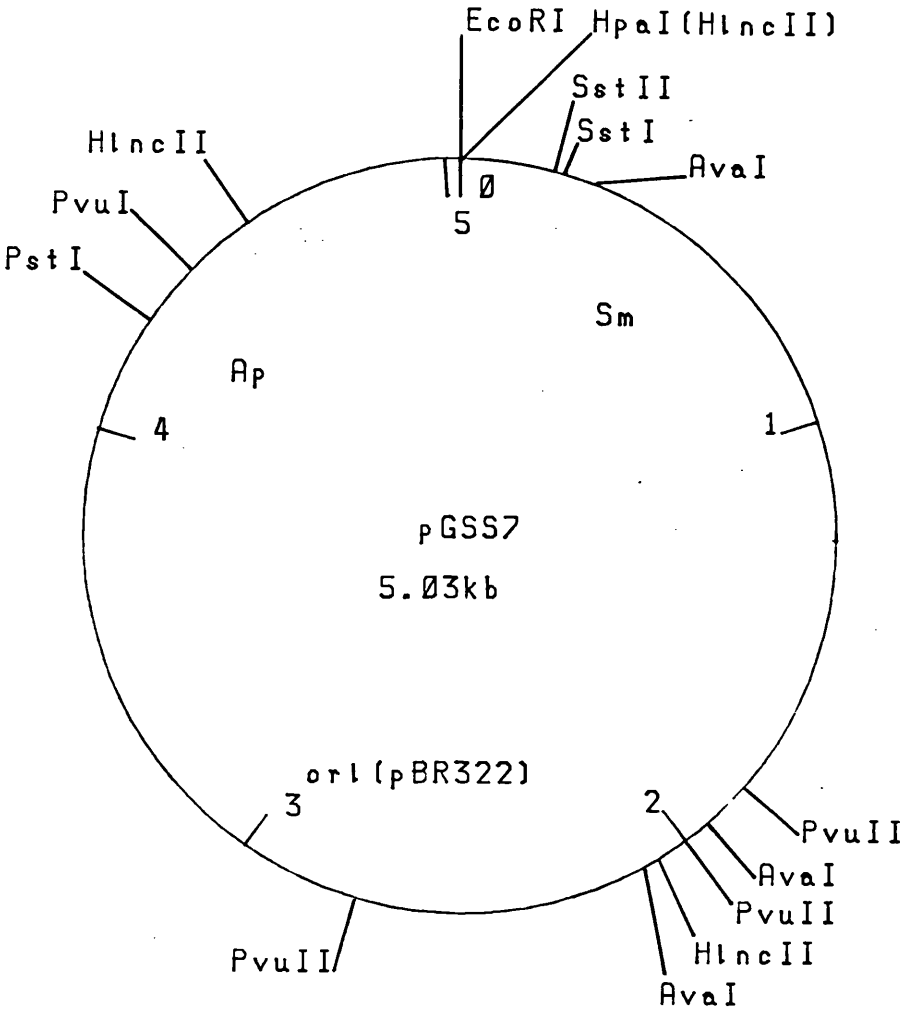



**Figure 10.**

A restriction map of pGSS7. This plasmid was the only viable product obtained after *Eco*RI digestion and re-ligation of pGSS4 [Fig.6].

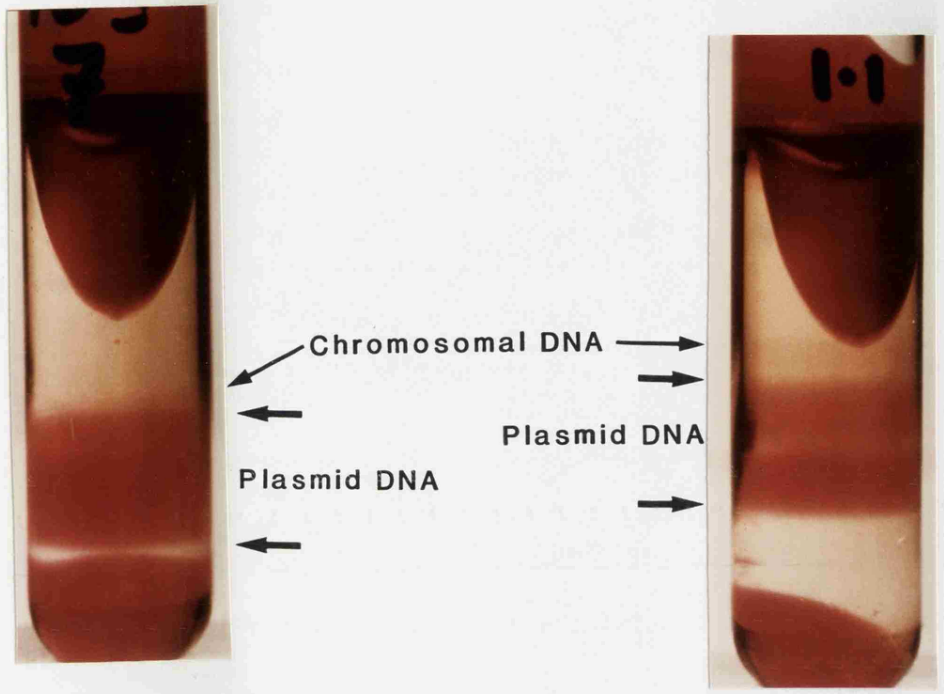



Figure 11.

Photographs of ethidium bromide-caesium chloride equilibrium density gradient DNA preparations of pGSS7 and pWS1.1. This demonstrates the extremely high yields of plasmid DNA routinely obtained using these plasmids.

strain ED8654 with selection for resistance to tetracycline or ampicillin. Out of approximately nine hundred $Tc^r$ colonies, 30% were $Ap^s$, $Sm^r$ and $Su^s$. A plasmid isolated from one of these colonies was designated pGSS8. Of the four hundred and thirty three $Ap^r$ colonies that were obtained two were found to be $Su^r$, $Sm^s$ and $Tc^s$. The plasmid isolated from one of these colonies was designated pGSS9.

The fact that two distinct plasmids have been derived from pGSS6, each with two antibiotic resistance genes in the predicted arrangements, clearly demonstrated that the replication origins of both pBR322 and R300B must have been present, and functional, on the co-integrate. Further, from extrapolation of data presented on RSF1010, (de Graaff *et al.*,1978), the R300B replication origin will be on pGSS8, with the pBR322 origin of replication on pGSS9.

Restriction analysis confirmed that pGSS8 and pGSS9 were the structures that could be predicted from *Eco*RI cleavage of pGSS6, and further confirmed the point of insertion of pBR322 into R300B as being between the *Bst*EII site and the *Pst*I site. The experiment also demonstrated that cleavage at the *Eco*RI site of R300B would separate but not inactivate the $Su^r$ and $Sm^r$ genes, and indicated a potential use of this site, and the adjacent *Hpa*I site, as cloning sites within a transcription unit that would not give rise to fused proteins.The plasmid pGSS9 was found to require DNA polymerase I for replication and could not be mobilized into *M.methylotrophus*. In contrast, pGSS8 was DNA polymerase I-independent and retained the mobilization functions of the parental R300B.

The fragment from pBR322 containing the $Tc^r$ gene has introduced unique *Cla*I, *Hind*III, *Bam*HI, and *Sal*I sites, together with additional *Ava*I and *Pvu*II sites. Cloning into the *Bam*HI or *Sal*I sites inactivated the $Tc^r$ gene and produced recombinants with a $Tc^s$ $Sm^r$ phenotype. The *Sst*I and *Sst*II sites, retained from R300B, provided

insertional inactivation into the $Sm^r$ gene, and both these enzymes produce 3' single-stranded ends which are ideal substrates for homopolymer tailing. Thus by having the $Su^r$ gene replaced by a $Tc^r$ gene in R300B, the cloning properties are improved and pGSS8 becomes a useful broad host-range vector.

In pGSS8 the unique *Eco*RI site now forms one pBR322/R300B sequence junction and lies between the $Tc^r$ and $Sm^r$ genes. Since the $Tc^r$ fragment contained an *Ava*I site and a *Pvu*II site, an *Eco*RI/*Hae*II fragment from pBR322 would need to be 2354bp long and would contain all the *Hae*II fragments except number 11 (containing the pBR322 *ori*) and most of number 1 (containing the $Ap^r$ gene), [Fig.4]. As pGSS8 possessed a *Bst*EII but no *Pst*I sites, the second pBR322/R300B sequence junction must be located within the 270bp region falling between the *Bst*EII site and the adjacent *Pst*I site of R300B, [Fig.12].

Confirmation of this structure for pGSS8 came from analysis of heteroduplexed molecules formed by annealing *Sal*I-linearized pBR322 and pGSS8, [Fig.13a] or *Sst*I-linearized R300B and pGSS8 [Fig.13b]. From the average measurement of twenty molecules in the former category it was calculated that 53% of pBR322 was homologous to pGSS8, corresponding to approximately 2325bp which is in good agreement with the required 2354bp. Plasmid R300B formed a very tight heteroduplex with pGSS8 and only a small loop of non-homology was apparent which essentially corresponded to the $Su^r$ gene of R300B and the $Tc^r$ gene of pGSS8. This confirmed the uninterrupted nature of the first 7.6kb of the R300B replicon present in pGSS8, and may be very important in view of the evidence observed in the formation of pGSS4 and pGSS7, suggesting that essential genes are spread over a large proportion of the genome.

As observed previously, the $Sm^r$ gene does not have its own

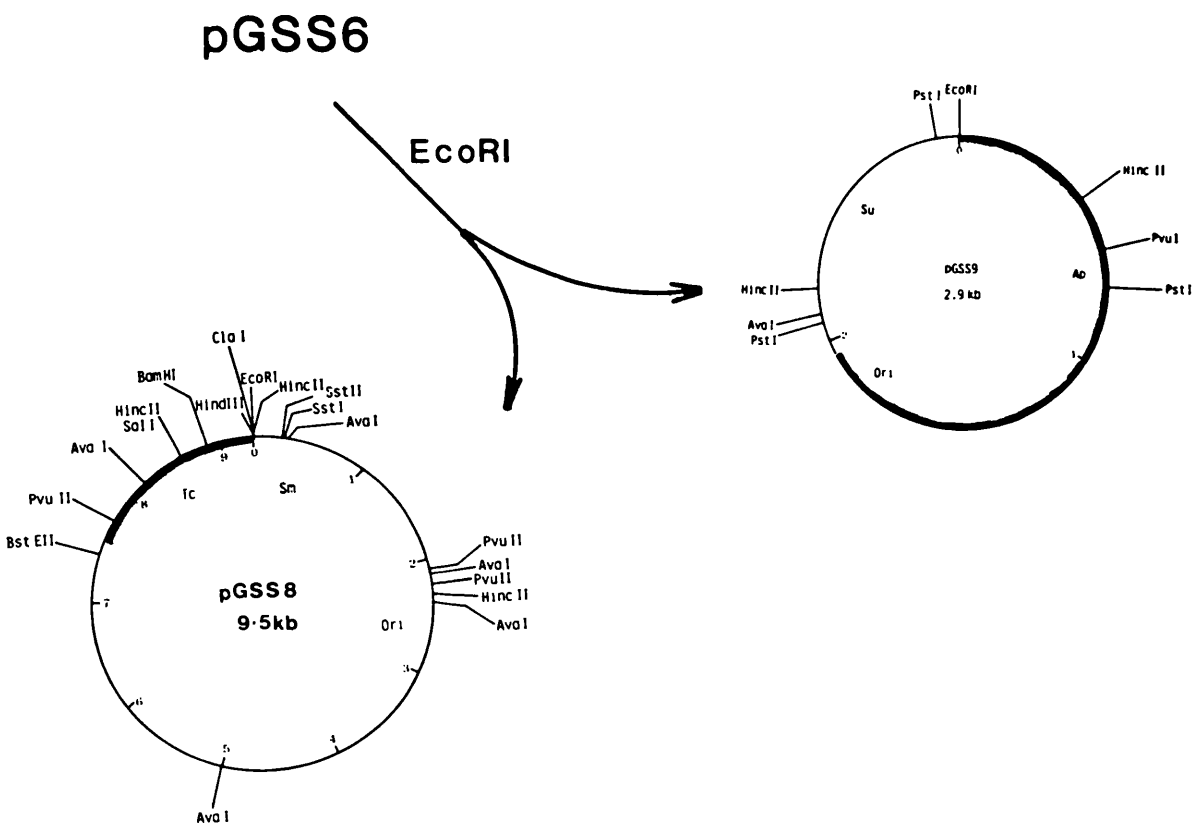



**Figure 12.**

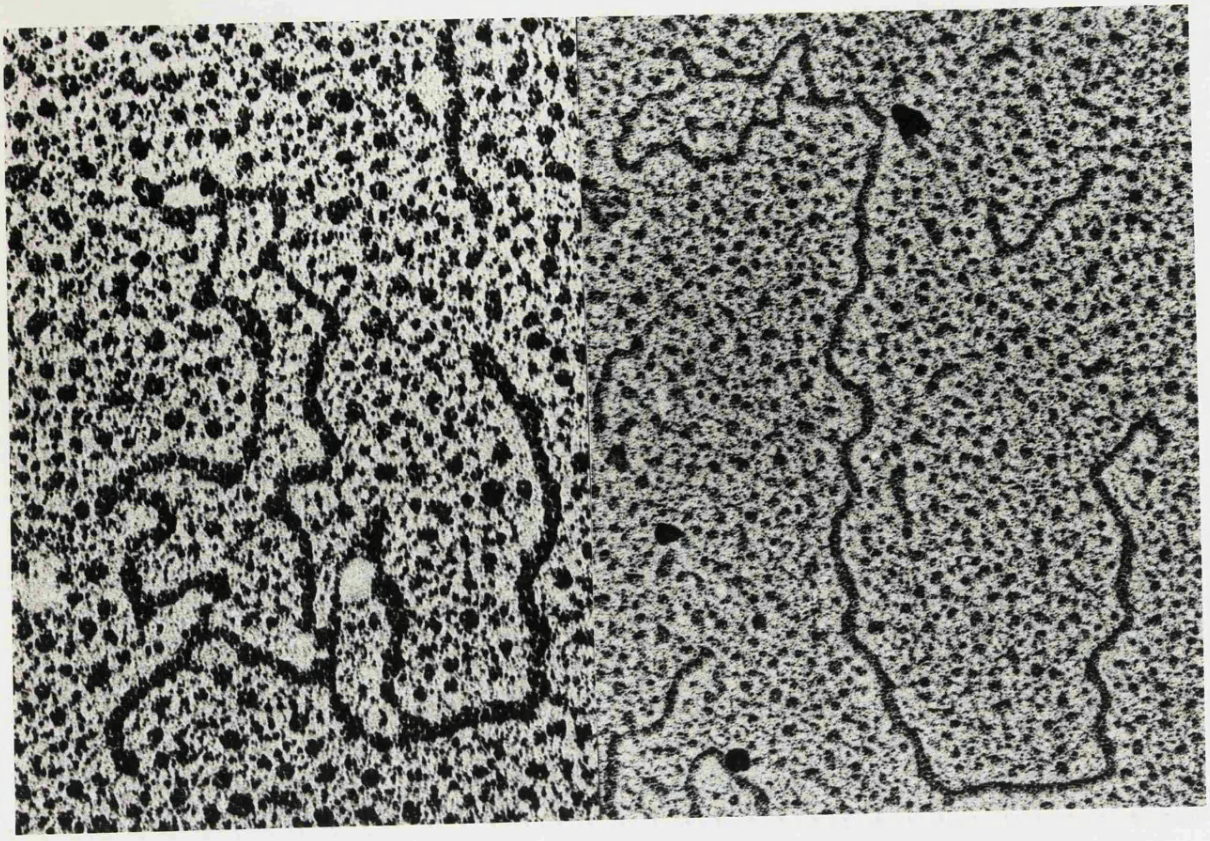
Derivation of plasmids pGSS8 and pGSS9 from pGSS6 by *Eco*RI digestion and religation of the fragments. The pBR322-derived sequences are shown emboldened.

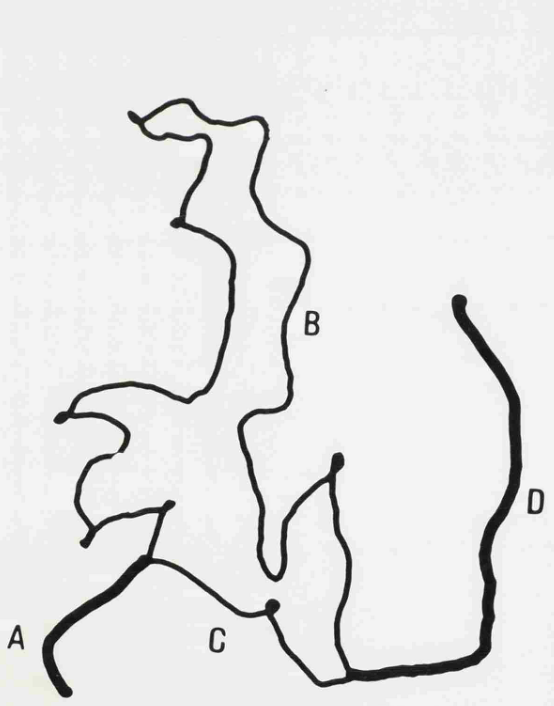
Figure 13.

Electron micrographs of heteroduplexed DNA formed by using the formamide technique designed by Westmoreland *et al.* (1969).

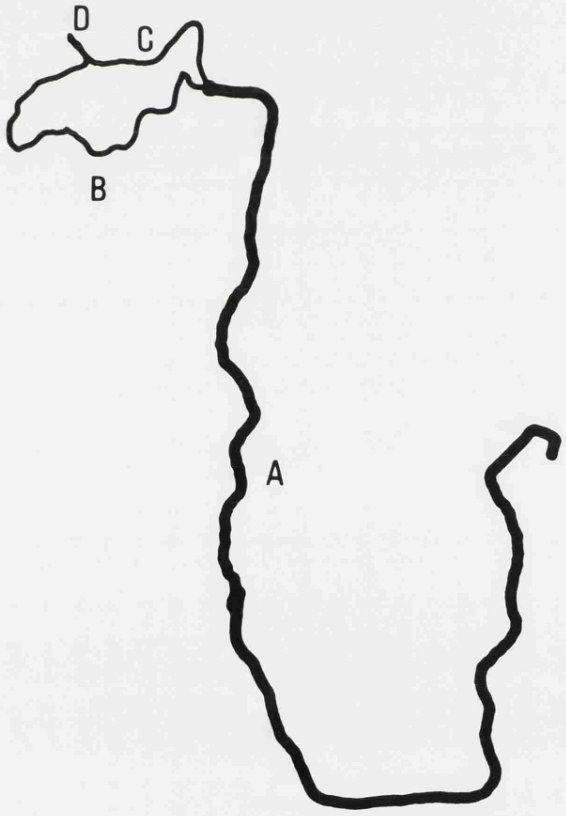
(a) Heteroduplex formed between *Sal*I-linearized pGSS8 and pBR322. Average measurements of twenty such molecules revealed that 53% of pBR322 was homologous to pGSS8 [(A+D/A+D+C) × 100].

(b) Heteroduplex formed between pGSS8 and R300B after linearizing them with *Sst*I. Regions A and D show the extent of uninterrupted homology, with only a small loop of nonhomology corresponding to the R300B sequence containing the $Su^r$ gene (region C), and the pBR322 sequence containing the $Ap^r$ gene (region B).






a




b

promoter and in R300B is transcribed from the promoter for the $Su^r$ gene. In pGSS7 it appeared to be transcribed from a promoter within the $Ap^r$ gene and in pGSS8 transcription must be initiated from a promoter near the start of the $Tc^r$ gene, (Stueber and Bujard,1981). However, this may cause a problem if a cloning procedure disrupts the region between the *Hind*III site and the *Eco*RI site, as both selective markers could be lost. To provide an alternative secondary selection, and extra cloning sites, I decided to clone the $Ap^r$ gene from pBR322 into pGSS8.

# CHAPTER 5

## Introduction of an $Ap^r$ gene into pGSS8

### 5.1. Construction of pGSS12, pGSS13, pGSS14 and pGSS15.

In view of the potential loss of selective markers in certain applications of pGSS8, I decided to introduce an alternative antibiotic-resistance gene. The advantages of this would be twofold: first, to provide a selection in the absence of $Tc^r$ and $Sm^r$; and second, to introduce extra unique cloning sites. Since *Hae*II partial digests had proved successful in producing linear molecules of R300B and pBR322 to produce the co-integrate pGSS6, this was the strategy adopted for introduction of the $Ap^r$ gene of pBR322, (contained on the single *Hae*II fragment 11), into pGSS8.

Thus 2μg of pBR322 DNA, in a 50μl reaction, were digested with two units of *Hae*II at 37°C for 90min, and 1.5μg of pGSS8 DNA, in a 50μl reaction, were digested with one unit of *Hae*II for 5min at 37°C. A 20μl aliquot of the *Hae*II cut pBR322 and 40μl of the partially cut pGSS8 were mixed and incubated at 30°C for 15min. The reaction volume was increased to 100μl, to include ligase buffer and five units of T4 DNA ligase, then incubated overnight at 10°C. Transformation of *E.coli* strain NM36 was effected with 10μl aliquots of the ligation mix and selection was for resistance to ampicillin. (The NM36 strain, while not the most efficient transformation host, was designed to select against any plasmids only able to replicate from a pBR322 origin of replication). Only five $Ap^r$ colonies were obtained and these were scored for other resistance determinants. One was $Ap^r$ $Tc^s$ $Sm^s$, another was $Ap^r$ $Tc^r$ $Sm^s$ and three were $Ap^r$ $Tc^s$ $Sm^r$. The colony that was only resistant to ampicillin was not of interest as a cloning vector, but the others were analysed for

plasmid content and all contained a single plasmid species, of similar sizes but all larger than the parental pGSS8. The $Ap^r$ $Sm^r$ plasmids were designated pGSS12, pGSS13 and pGSS14, while the $Ap^r$ $Tc^r$ plasmid was designated pGSS15.

Preliminary restriction analysis of these four plasmids with *Pst*I, *Eco*RI and *Bam*HI+*Pst*I was sufficient to indicate position and orientation of the $Ap^r$ insert, at least in the $Ap^r$ $Sm^r$ trio. In pGSS12, the insertion had occurred at a *Hae*II site within the $Tc^r$ gene 1209bp from the *Eco*RI site of pGSS8. The orientation was such that it produced *Eco*RI fragments of ≈8kb and ≈3kb, and *Bam*HI-*Pst*I fragments of ≈9.2kb and ≈1.8kb, which indicated that a small deletion had also occurred (predicted size 11.44kb). Plasmids pGSS13 and pGSS14 showed restriction patterns which were very similar and further analysis of just pGSS14 with *Eco*RI+*Sal*I, *Ava*I, *Hind*III, *Pvu*II+*Sal*I, and *Pvu*II+*Hind*III confirmed a structure as shown in Fig.14. However, pGSS13 produced two small bands at ≈1100bp when cleaved with *Bam*HI+*Pst*I and probably contained an extra *Bam*HI site introduced on a partial *Hae*II segment of pBR322 comprising fragments 1 and 2. This plasmid was not analysed further.

**5.2. Analysis of pGSS15**

The $Ap^r$,$Tc^r$ plasmid, pGSS15, was much more interesting since it carried the same resistance genes as pBR322, and pGSS15 only had a single *Eco*RI site. Digestion with *Bam*HI+*Pst*I produced fragments of ≈10.5kb and ≈1.1kb: there was only a single site for *Hind*III, no site for *Sst*I but two sites for *Sal*I approximately equidistant from a unique *Pst*I site. Rather than adding extra *Eco*RI and *Hind*III sites, the inserted *Hae*II fragment appeared to have substituted these sites. This could be possible if the incoming fragment had replaced one or

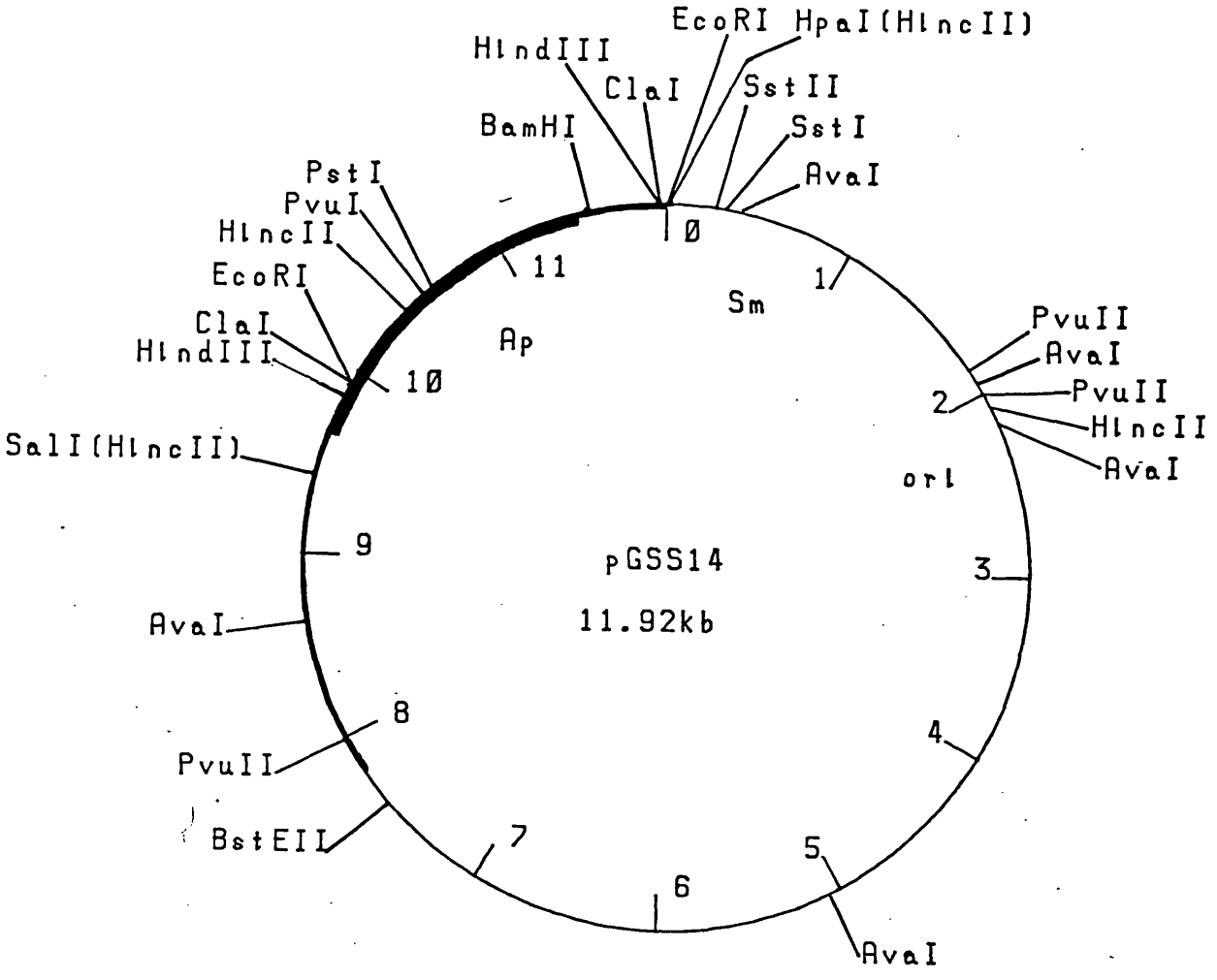



**Figure 14.**

A restriction map of pGSS14. The pBR322-derived sequence is shown by the heavy line and the inserted *Hae*II fragment by the very heavy line. The point of insertion was the *Hae*II site 418bp from the *Eco*RI site, into the $Tc^r$ gene.

more *Hae*II fragments extending from the *Hae*II site, 237bp into the $Tc^r$ gene from the *Eco*RI site, through the *Eco*RI site-junction of pBR322 and R300B sequence in pGSS8, and on beyond the *Sst*I site. Such an event would re-constitute the $Tc^r$ gene and place the $Ap^r$ gene in exactly the same position, relative to the $Tc^r$ gene, as it would be in pBR322. Digestion of pBR322 and pGSS15 DNAs with *Bam*HI+*Pst*I produced single small fragments from each plasmid which were identical.

Analysis of a *Hae*II digest of pGSS15 by PAGE through a 5% gel revealed the possible presence of the pBR322 *Hae*II-fragment 11 and a double-strength band at the position of the pBR322 *Hae*II-fragment 6 [Fig.15]. This latter band was shown to disappear and be replaced by two other double-strength bands of 122bp and 105bp, suggesting that the pBR322 *Hae*II fragment 6 was duplicated in pGSS15, accounting for the presence of the extra *Sal*I site. Digestion of the gel-purified small *Sal*I fragment from pGSS15 with *Hae*II clearly showed that the pBR322 *Hae*II fragment 11 was present [Fig.16]. This fragment carries the origin of replication from pBR322 and, being contiguous with the $Ap^r$ fragment 1, may not have been completely cleaved by *Hae*II in the original digest. However, Fig.15 also shows a double-strength band at 105bp which would be consistent with the *Sal*I fragment of pGSS15 being bounded by inverted repeats of the smaller *Sal*I-*Hae*II portions of the pBR322 fragment 6 [Fig.17].

It is clear, therefore, that extra *Hae*II fragments from pBR322 were ligated into pGSS8 along with the $Ap^r$ fragment, and that their introduction was an unfortunate consequence of not isolating a purified fragment prior to ligation. Nevertheless, the fortuitous arrangement of the antibiotic resistance genes was a bonus, since it would facilitate sub-cloning from pBR322 or pAT153 (Twigg and Sherratt,1980), and warranted further structural analysis of pGSS15.

**Figure 15.**

Restriction analysis of pGSS15 using *Hae*II and *Hae*II + *Sal*I. Track: 1. pGSS15 cut with *Hae*II; 2. pBR322 cut with *Hae*II; 3. R300B cut with *Hae*II; 4. pGSS15 cut with *Hae*II+*Sal*I; 5. pBR322 cut with *Hae*II+*Sal*I; 6. DNA size markers from *Alu*I-cut pBR322.

The *Hae*II fragment of 370bp from pGSS15 (track 1), which was suspected of containing the pBR322 *ori*, could have arisen from pBR322 or R300B (see also Fig.16). The 227-bp *Hae*II fragment from pGSS15 was clearly derived from pBR322 but the intensity of its band on the gel was consistent with it being a doublet. *Sal*I-cleavage of this fragment was observed to produce the expected 122-bp and 105-bp sub-fragments (track 4), but again each of these were doublets suggesting that the 227-bp fragment must be duplicated.

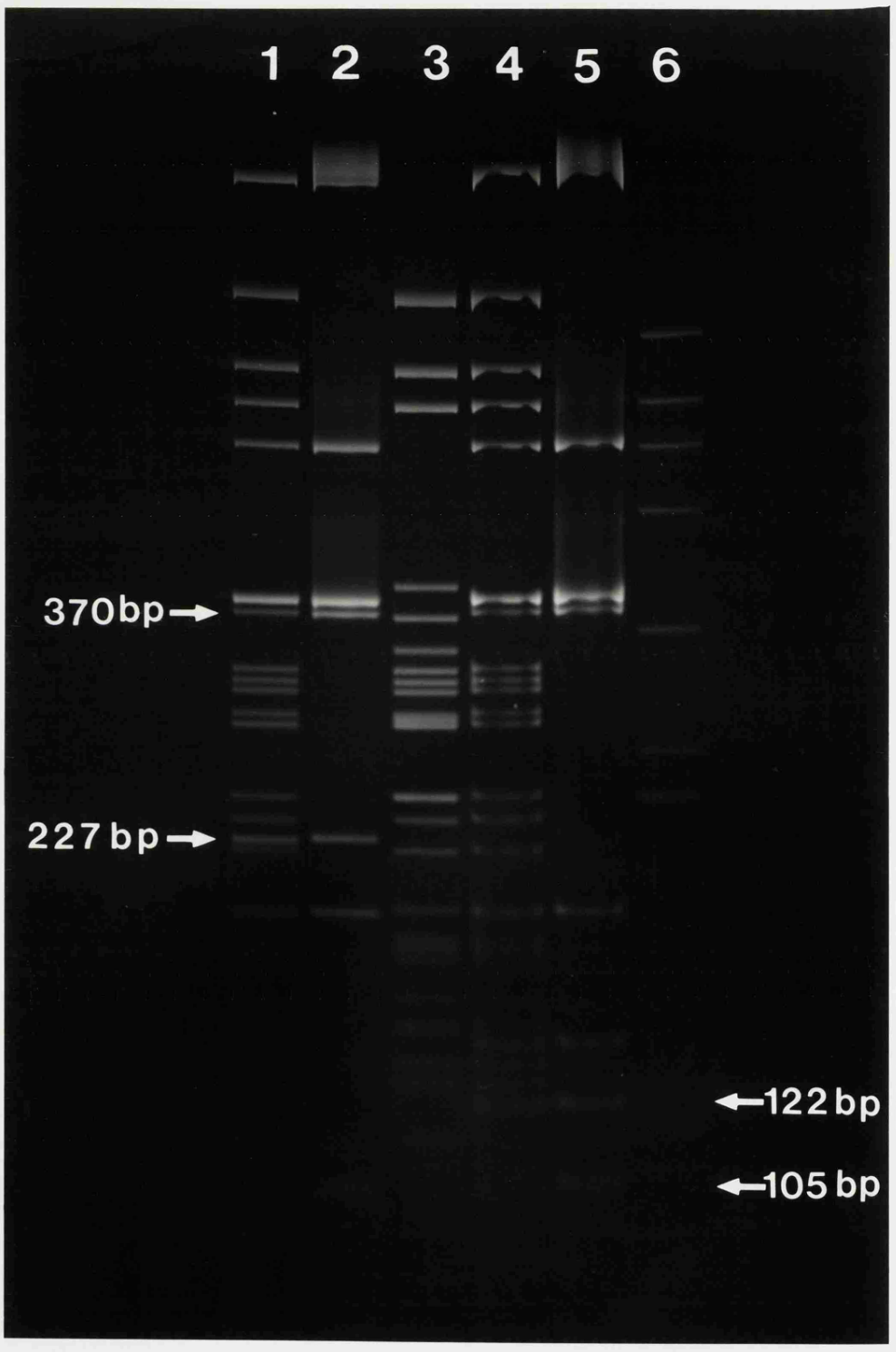

1 2 3 4 5 6
370bp
227 bp
122 bp
105 bp

**Figure 16.**

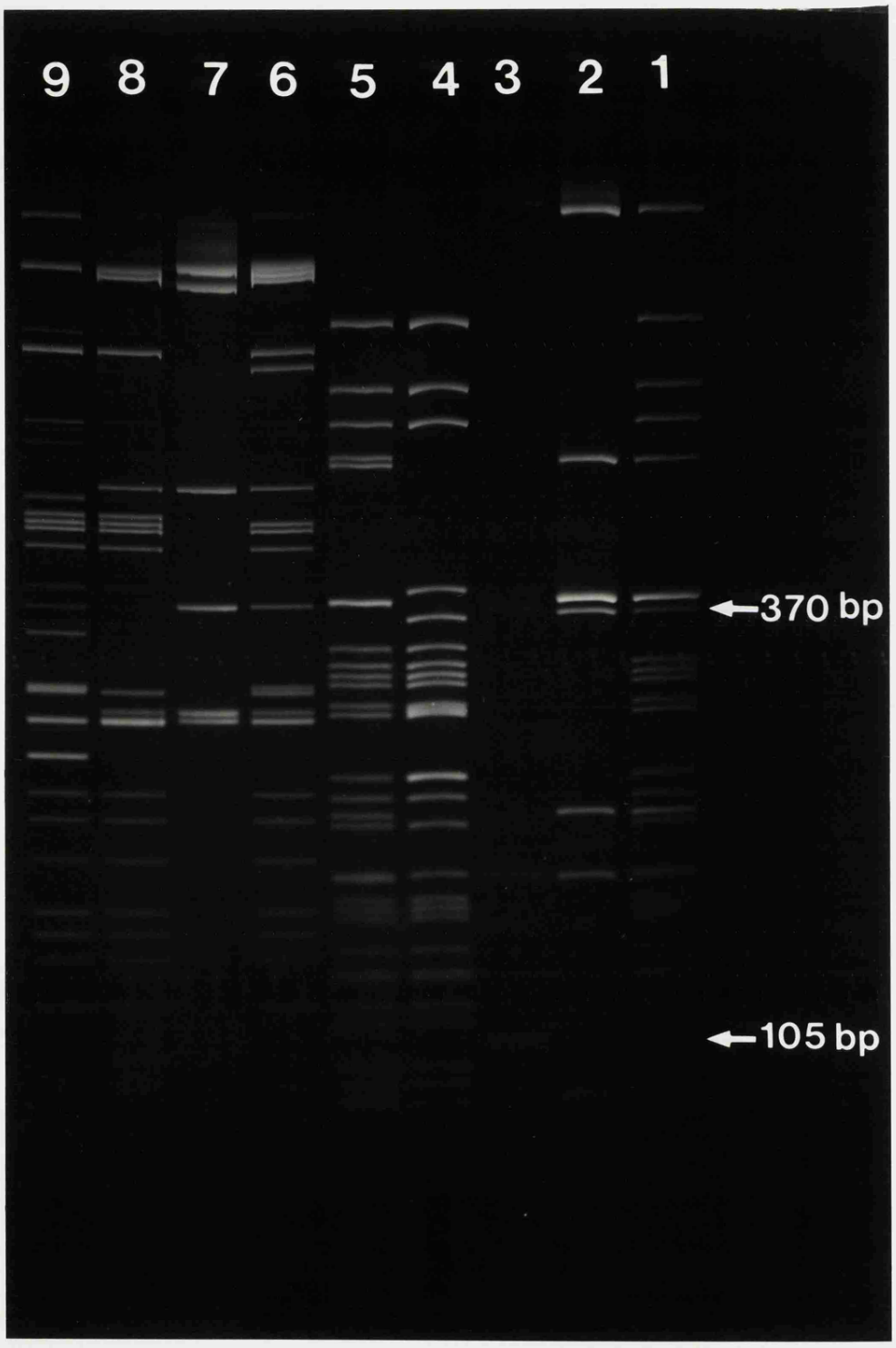
Restriction analysis of pGSS15 using *Hae*II and *Taq*I.

Track: 1. pGSS15; 2. pBR322; 3. small *Sal*I-fragment from pGSS15; 4. R300B; 5. pGSS8 - all cut with *Hae*II.

Track: 6. pGSS15; 7. pBR322; 8. pGSS8; 9. R300B - all cut with *Taq*I.

The important track on this gel was 3, in which the small *Sal*I-fragment from pGSS15 was gel-purified and re-cut with *Hae*II. This clearly showed the presence of the pBR322 *Hae*II-fragment 11 (370bp) which contained *ori* and showed a double-strength band at 105bp indicating that the extra 227-bp *Hae*II-fragment 6 must have been in the inverted orientation (see Fig.17).

The *Taq*I digests were less informative, but pGSS15 did have the largest *Taq*I fragment from pBR322 which would require the $Ap^r$ gene and *ori* to be contiguous. Also if the *Hae*II-fragment 6 from pBR322 was joined to the *Hae*II-fragment 11 from pBR322 as proposed, then a new *Taq*I fragment of 320bp would be predicted. The plasmid pGSS15 has such a band present.


9 8 7 6 5 4 3 2 1
370 bp
105 bp

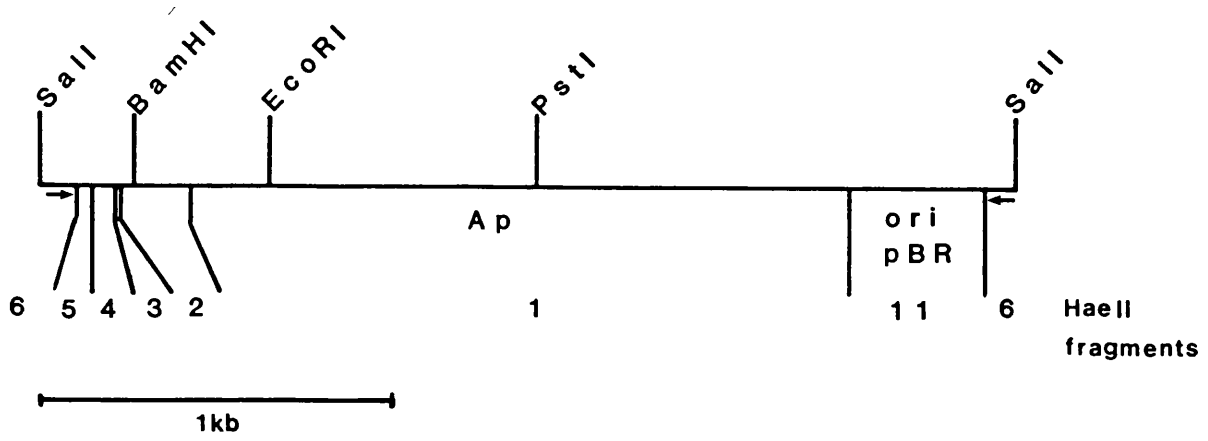



**Figure 17.**

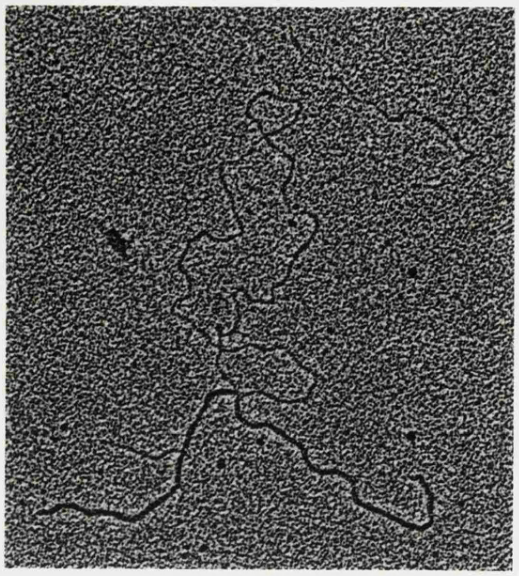
Diagram of the small *Sal*I-fragment from pGSS15. The *Hae*II fragment numbers are those of pBR322 (Fig.4) and the orientations of the 105-bp sub-fragments of 6 are indicated. The PAGE analysis of these fragments was shown in Fig.16 track 3.

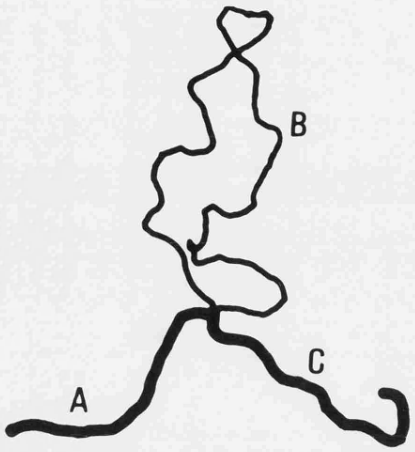
Analysis of heteroduplex molecules, formed between *Eco*RI-linearized pBR322 and pGSS15 DNAs, was used to assess the extent of homology between the two plasmids. A typical example is shown in Fig.18a, where the two double-stranded "tails", A and C, represent sequences from pBR322 which are also present in pGSS15, and the single-stranded loop arises from R300B-derived sequences together with inverted repetition sequence from pBR322. Average measurements of the regions A and C from twenty molecules produced values for their lengths of 2.35kb and 1.98kb, respectively. These were in very good agreement with the predicted lengths of 2.35kb and 2.01kb. However, amongst these molecules there was an occasional occurrence of a predominantly single-stranded molecule with a relatively short, internal, double-stranded region. This could have resulted from the "snap-back" of the inverted repetition in pGSS15 brought about by the duplication of the *Hae*II fragment 6 from pBR322 as described earlier.

**Figure 18.**

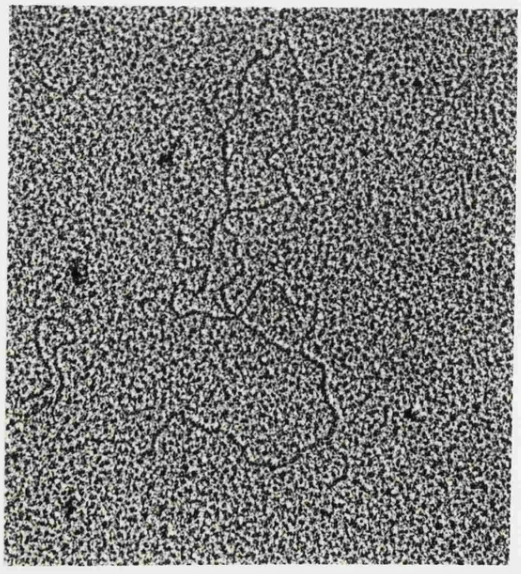
(a) Heteroduplex formed between pGSS15 and pBR322 DNAs linearized with *Eco*RI. Regions A and C represent sequence homology corresponding to the entire pBR322 genome. Region B is the loop of nonhomologous R300B-derived sequence which also includes the duplicated *Hae*II fragments 6 and 7 from pBR322.

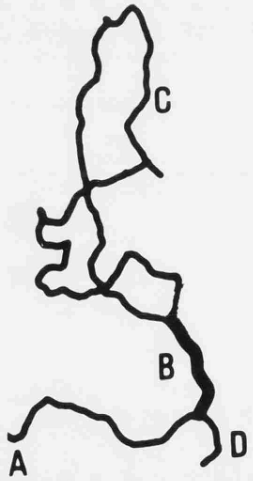
(b) Snap-back analysis of single stranded pGSS15 DNA linearized at its *Eco*RI site. Regions A and D are derived from the pBR322 sequence which lies between the inverted repeats and which has been cleaved asymmetrically by the *Eco*RI. Region B is the duplexed repeated sequence, and region C is all of the R300B-derived sequence together with *Hae*II fragments 8, 9 and 10 of pBR322.






a






b

To test this more rigorously, an experiment designed to look specifically for the ability of molecules with internal inverted repeats to rapidly form "stem-loop" structures was undertaken. Denatured, *EcoR*I-cut, pGSS15 DNA was allowed to re-anneal for only 10-40min at 37°C compared with the 5hr at 37°C required to form the heteroduplex with pBR322 DNA. The experiment did produce many more of these snap-back molecules [Fig.18b] but the average length of the double-stranded region obtained from measurements of twenty molecules was 655bp, much longer than than expected from a repeat of just the pBR322 *Hae*II fragment 6 (227bp). However, it would be consistent with a repeat of the *Hae*II fragments 6 and 7, which together would be 657bp.

This was confirmed by a *Hinf*I digest of pGSS15 [Fig.19], which showed a double-strength band at 154bp corresponding to the pBR322 *Hinf*I fragment C [Fig.20], which is totally contained within the pBR322 *Hae*II fragment 7. The *Hinf*I restriction analysis was also extremely informative in confirming the total structure of pGSS15, especially when compared to a published *Hinf*I map of the identical plasmid, R1162 (Meyer *et al.*,1982) [Fig.20]

The plasmid pGSS15 contained every *Hinf*I fragment from pBR322 with the exception of the 344-bp fragment G. This fragment contained the *Hae*II site at co-ordinate 2353 that formed one of the pBR322/ R300B sequence junctions, and consequently formed hybrid *Hinf*I fragments in pGSS15. Thus 322bp of pBR322 fragment G joined with ≈2.9kb of R300B *Hinf*I fragment K to produce the largest *Hinf*I fragment observed in both pGSS8 and pGSS15. The remaining 22bp of the fragment G would then be joined to 81bp of pBR322 *Hinf*I fragment A, which was brought in on the duplicated *Hae*II fragment 6, to form a new fragment of 103bp unique to pGSS15.

A detail worth noting here was the presence of two, or possibly

Figure 19.

Restriction analysis of pGSS15 using *Hinf*I (right-hand tracks) and *Hae*III (left-hand tracks).

Unique bands of 103bp and 430bp can be seen in the pGSS15 *Hinf*I track. Also the intensity of the band at 154bp is indicative of a doublet.

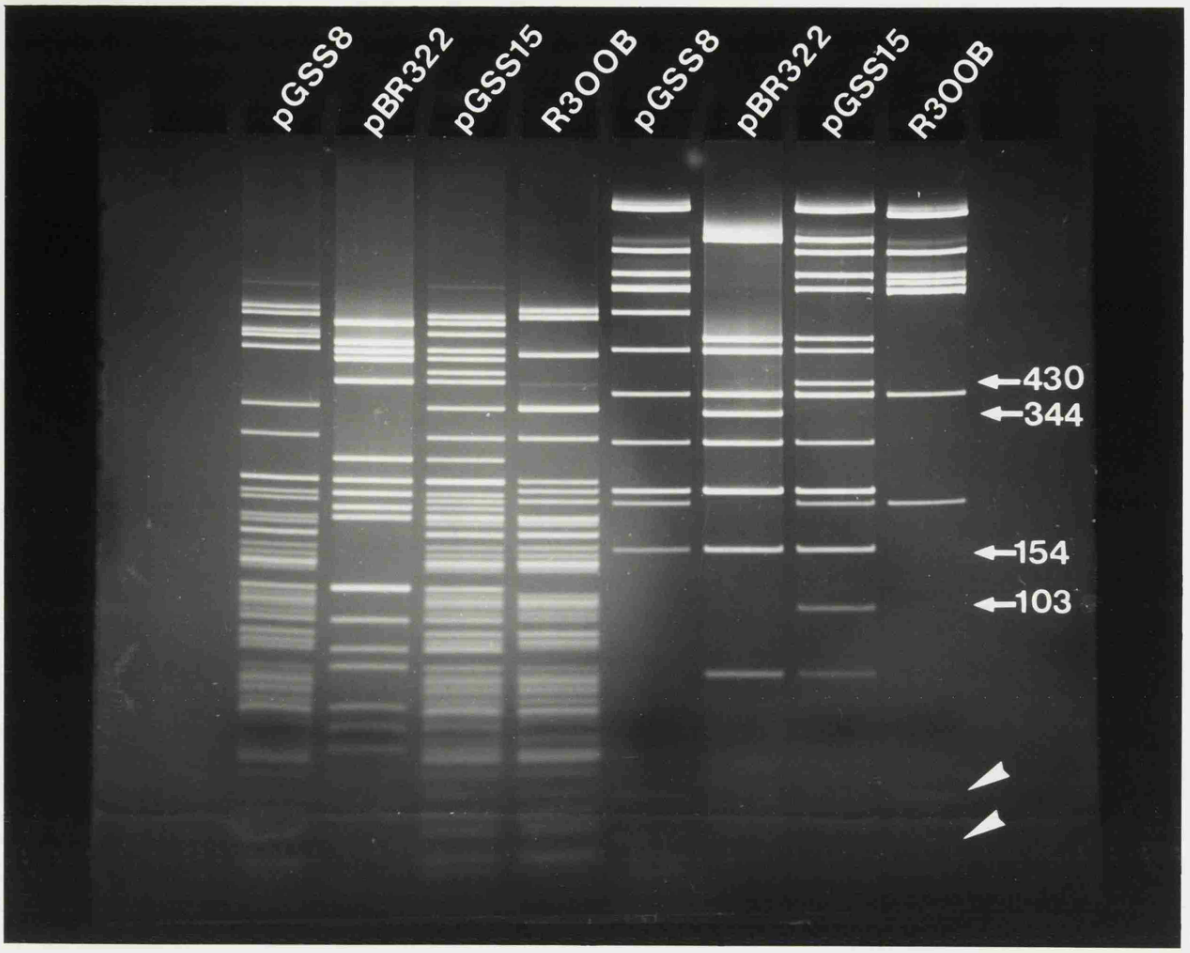
The arrowed bands in the R300B track are the 36-bp and 25-bp fragments C and B respectively [Fig.20]. The three bands close to the top of this track correspond to R300B *Hinf*I fragments G, A and D/E (in order of decreasing size). In both pGSS15 and pGSS8 the A-fragment band is missing (associated with the $Su^r$ gene) and in pGSS15 the bottom band of the trio is less intense due to the deletion of part of fragment D, the remainder of which contributes to the 430-bp band.

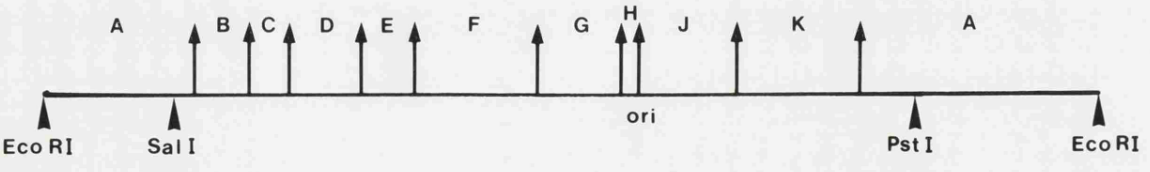
Without total sequence data, the *Hae*III digests of R300B, and hence of

pGSS15 and pGSS8, were too complex to be useful.

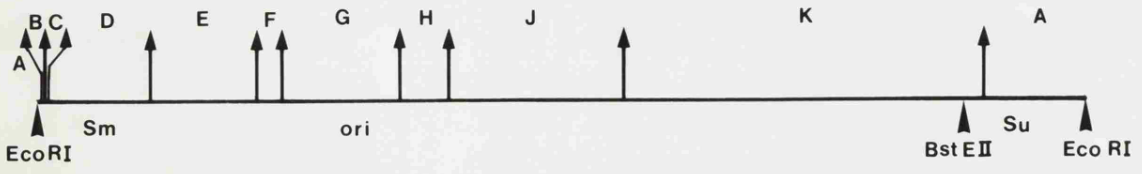
Figure 20.

*Hinf*I restriction maps of pBR322 and R300B. The plasmids are shown linearised at their unique *Eco*RI sites.




pBR322


R300B

three, bands between 20 and 50bp in size, which could be seen on my *Hinf*I digest of R300B. These bands were also present in the *Hinf*I digest of pGSS8 but not of pGSS15. This suggested that that they may be clustered between the *Eco*RI site and the *Sst*I site in R300B and could then have been lost in the deletion occurring during the construction of pGSS15. Evidence in support of this came from some sequence data on R300B of a region around the *Eco*RI site [Fig.21] (M.Guyer, personal communication). This revealed three *Hinf*I sites within the first 120bp from the *Eco*RI site giving rise to two extra fragments of, 25bp and 36bp, which had not been reported by Meyer *et al.* (1982).

Another *Hinf*I band unique to pGSS15 was observed at ≈430bp and must arise from a hybrid fragment at the other pBR322/R300B sequence junction made up from 202bp of pBR322 *Hinf*I fragment D and ≈230bp of R300B *Hinf*I fragment D. Thus, since the evidence for a deletion between the *Eco*RI and *Sst*I sites of pGSS8 during the construction of pGSS15 is very strong, and since R300B *Hinf*I fragments A, B, C and D are missing in pGSS15, then the *Hinf*I site at R300B co-ordinate 950 would form the other end of this ≈430bp hybrid fragment. The data then allow the size of the deletion to be fairly accurately estimated at 720bp. After further restriction mapping to confirm the presence of other sites, the map of pGSS15 [Fig.22] could be drawn with some confidence.

### 5.3. Properties and mobilization of pGSS15.

The plasmid pGSS15 could be propagated in a *pol*A strain of *E.coli*, thus confirming replication from a non-ColE1 origin, and could be readily mobilized into *M.methylotrophus* by *e.g.* RP4, S-a or R64*drd*11. The latter plasmid, belonging to the incompatibility group Iα, is reported to have only a fairly narrow host range (Bukhari *et*

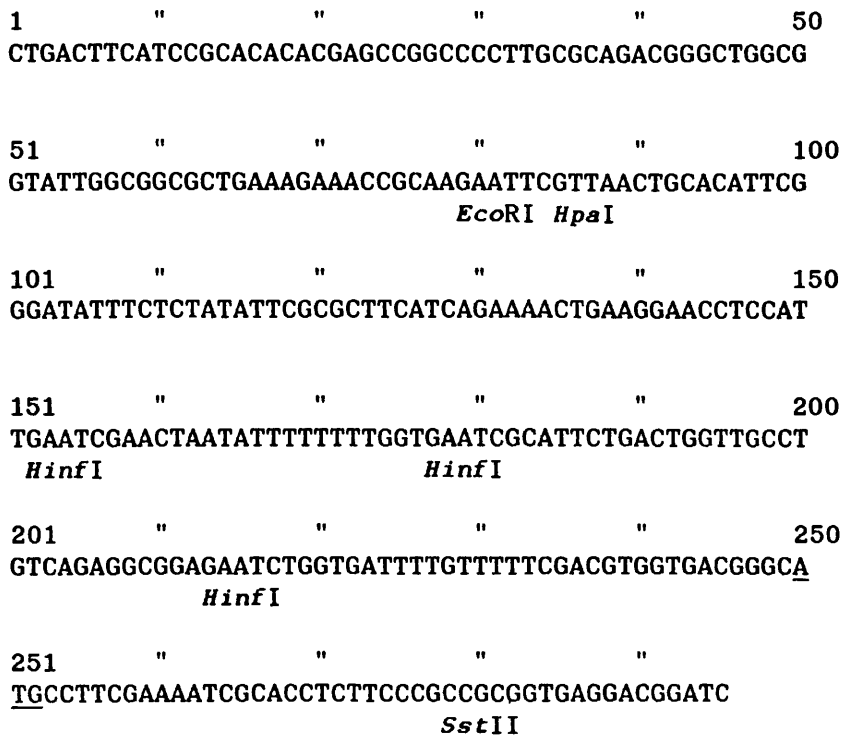



Figure 21.

DNA sequence around the *Eco*RI site of R300B (M.Guyer, personal communication). The sequence confirmed the position of the *Hpa*I site (GTTAAC) immediately adjacent to the *Eco*RI site (GAATTC), confirmed the presence of three *Hinf*I sites close to the *Eco*RI site and showed a potential start (ATG) to the $Sm^r$ gene.

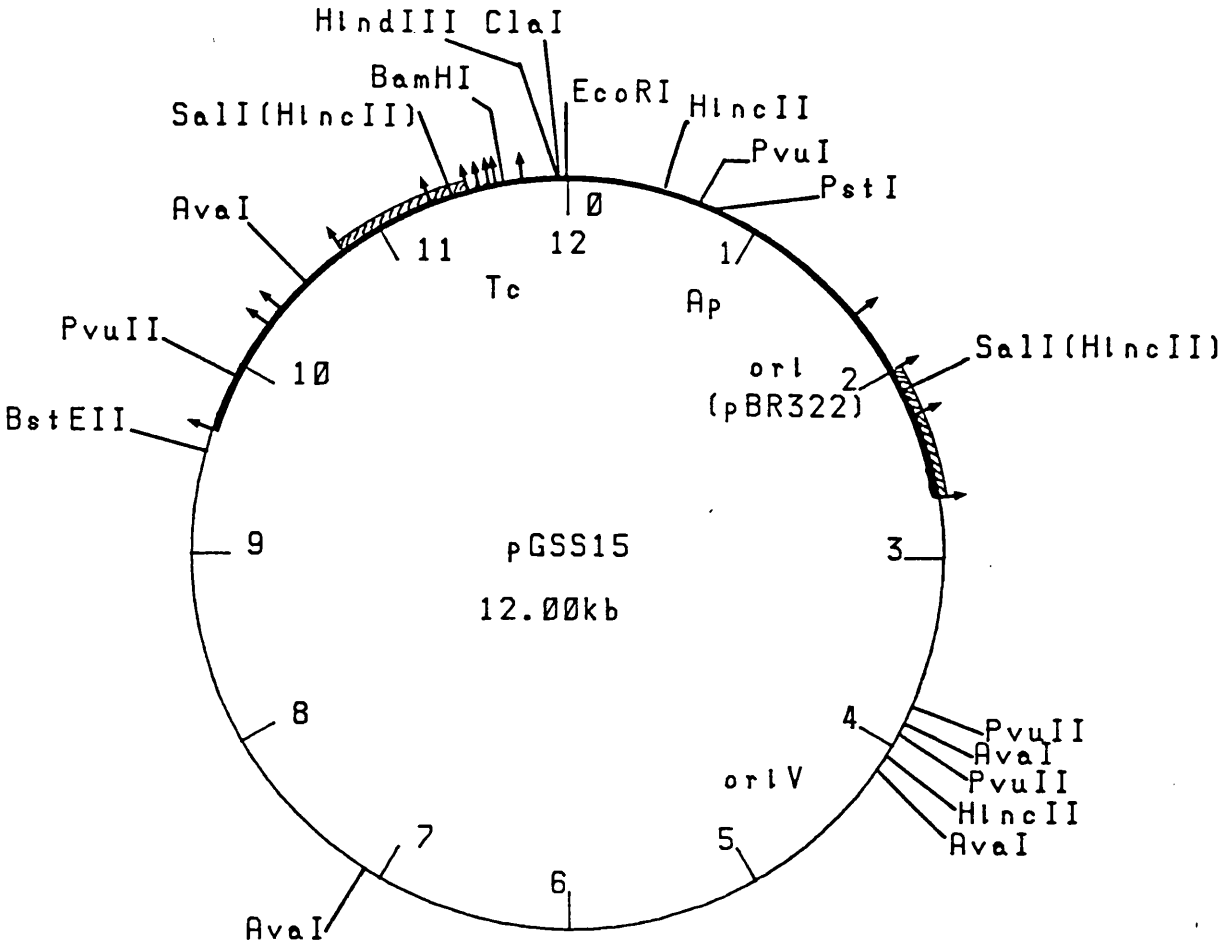



**Figure 22.**

A restriction map of pGSS15. The extent of the region derived from pBR322 is shown emboldened and the inverted repeats are shown as the hatched areas. The small arrows indicate the positions of the *Hae*II sites, but only for the pBR322-derived sequence.

*al.*,1977) and in fact I could never detect the presence of this plasmid in *M.methylotrophus*. This proved to be an added bonus in providing delivery of just pGSS15 without the complications of having the mobilizing plasmid present, giving extra or duplicated antibiotic resistances and ability of further mobilization . In a mating on solid media of R64*drd*11-containing *E.coli* and *M.methylotrophus*, no resistant colonies were detected, but when an Iα plasmid carrying Tn*5*, pLG221 (Boulnois,1981) was used in a similar cross, $Km^r$ *M.methylotrophus* colonies were observed. Single colony lysates to detect the presence of plasmids showed no bands other than chromosomal DNA and attempts to retrieve the $Km^r$ by mobilization from *M.methylotrophus* into *E.coli* proved negative. The probable conclusion to be drawn from this observation was that Iα plasmids were capable of promoting conjugation with *M.methylotrophus* but incapable of stable maintenence in that host. However, the Iα::Tn*5* plasmid provided a convenient transposon donor for the methylotroph. The implications from this observation may be far-reaching in terms of using transposon-carrying Iα plasmids as *in vivo* mutagens, since they may be able to conjugate with many other bacterial species beyond their prescribed host range. More recently, following these observations, pLG221 has been shown to behave in an analogous manner in *Pseudomomas putida* (Boulnois *et al.*,1984)

The plasmid pGSS15 is therefore a cloning vector with similar cloning site facilities as pBR322, but the with the ability to be mobilized into a wide range of Gram negative bacteria. The use of IncIα plasmids to promote the mobilization provides a clean delivery of the vector plasmid with a degree of safety, since subsequent mobilization would require introduction of another conjugative plasmid.

## 5.4 Attempts to remove the extra *Sal*I site from pGSS15

a) By partial *Sal*I digestion and "filling-in" the ends.

By using a partial *Sal*I digest there was a 50% chance of leaving the *Sal*I site in the $Tc^r$ gene intact. Thus the possibility arose that the extra *Sal*I site could be eliminated by filling in the cut ends with large fragment of DNA polymerase I, ligating the blunt ends and using this DNA for the transformation of *E.coli* selecting for $Tc^r$. In two such experiments I was unable to obtain any transformants. This may have been due to problems with blunt-end ligation, although I would have expected transformants from uncut molecules.

b) By treatment of the *Sal*I-cut ends at the extra site with *Bal*31.

So as to protect and retain the *Sal*I site in the $Tc^r$ gene the small *Bst*EII-*Hind*III fragment of pGSS15 was purified by electroelution from agarose gel slices. The two *Sal*I fragments were isolated in a similar manner.

0.4μg of the large *Sal*I fragment and 0.15μg of the small *Sal*I fragment were treated, in separate reactions, with 0.2unit of *Bal*31 at 30°C for 90sec. The *Bal*31 activity was stopped by heating the reaction mixtures to 70°C for 10min, followed by phenol extraction and ethanol precipitation. The dried precipitates were included in 25μl reactions in which the large fragment was digested with *Bst*EII at 60°C for 90min and the small fragment digested with *Hind*III at 37°C for 90min. After phenol extraction and ethanol precipitation, the dried pellets were each dissolved in 3μl of water and mixed with 0.2μg of the small *Bst*EII-*Hind*III fragment of pGSS15. This mixture was treated with T4 DNA ligase at 10°C overnight then 2μl aliquots were used for the transformation of *E.coli* strain ED8654. No

transformants were obtained.

c) By attempting to join a *Hae*II-generated sticky end to a *Sal*I-generated sticky end using a synthetic single-stranded octanucleotide linker, (a gift from M.Edge).

This experiment was designed to remove both the extra *Sal*I site and the pBR322 *ori* from pGSS15. The following fragments were purified by electroelution from an agarose gel: the largest BstEII-*Sal*I fragment and the smaller *Bst*EII-*Hind*III fragment of pGSS15 and the largest *Hind*III-*Hae*II fragment from pBR322. The arrangement required for the ligation of the fragments is shown in Fig.23, where the starred junction represents the juxtaposition of the *Hae*II single- stranded 3' tail and the *Sal*I single-stranded 5' tail. This leaves an 8-base space to be filled in by the synthetic octanucleotide linker.

```
            A
5'------  GCGCTCGAC------3'
            G

  ------T        G------
         CGCGAGCT
```

This sequence would no longer be recognised by *Sal*I and the *Hind*III-*Hae*II fragment would not contain the pBR322 *ori*.

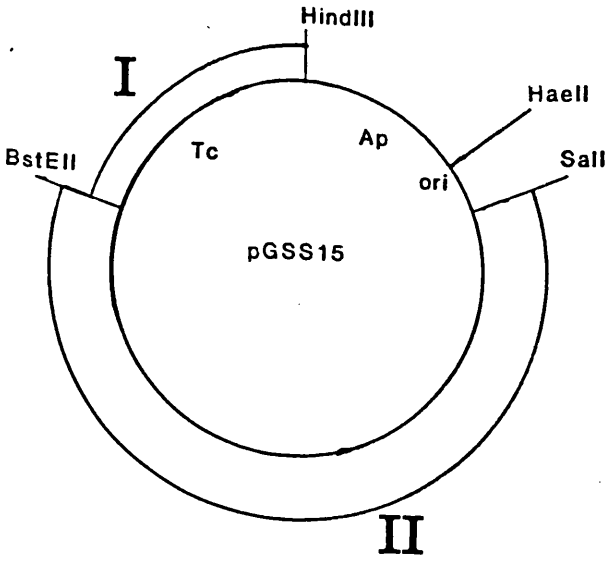
The synthetic octanucleotide was treated with T4 kinase to provide a 5'-phosphoryl group, making a suitable substrate for T4 DNA ligase. It was convenient at this stage to introduce a radioactive tag from $^{32}P\gamma ATP$ to be able to check reaction efficiencies and the fate of the fragment.

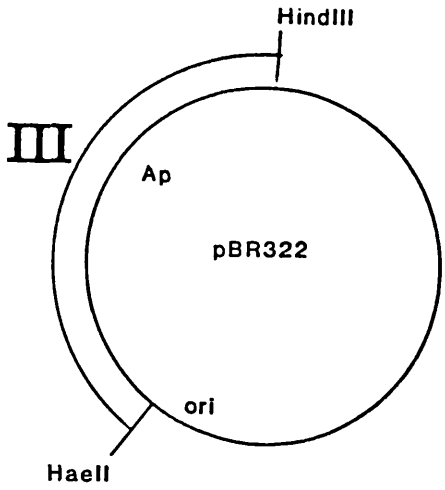
Thus 0.5μg of fragments A and C, 1μg of fragment B and 0.15μg of the kinased octanucleotide were treated with T4 DNA ligase in a 100μl reaction at 22°C for 90min. 10μl aliquots were used for the transformation of *E.coli* strain ED8654, with selection for resistance to Ap or Tc.

**Figure 23.**

Diagram of the approach used in an attempt to remove the pBR322 *ori* and the extra *Sal*I site from pGSS15.

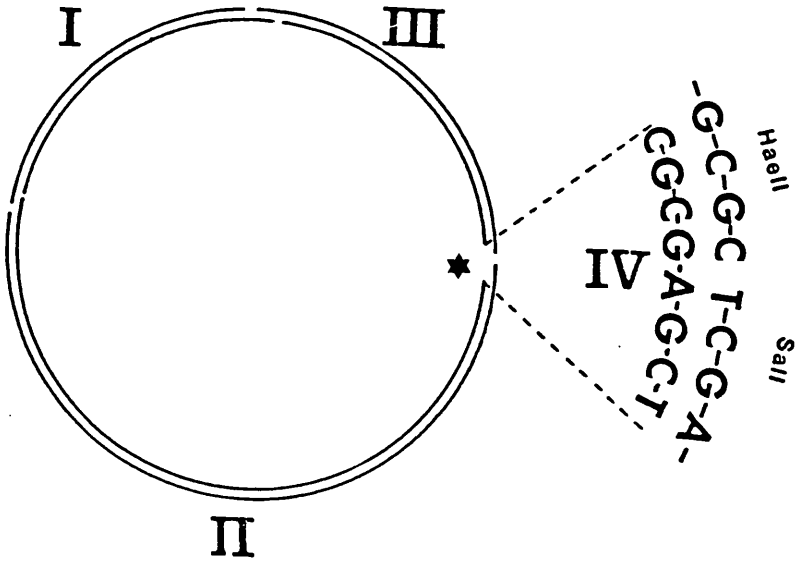
Fragments I and II were isolated from digests of pGSS15 DNA and fragment III from a *Hind*III+*Hae*II digest of pBR322 DNA. Fragment IV was the synthetic octanucleotide which was to be used to link a *Hae*II sticky end to a *Sal*I sticky end. (✱)


HindIII
I
HaeII
Tc
Ap
BstEII
SalI
ori
pGSS15
II


HindIII
III
Ap
pBR322
ori
HaeII

CGCGAGCT

IV


I
III
HaeII
IV
SalI
II

Two $Ap^r$ colonies were detected and found to be also $Tc^r$. Small scale plasmid analysis revealed that these two isolates each contained a plasmid indistinguishable from pGSS15 and which still had two *Sal*I sites.

A small sample of the ligation mix was run out on a 1% agarose gel, with pGSS15 DNA as a marker, and the gel autoradiographed. On both the photograph of the ethidium bromide-stained gel and the autoradiograph a faint band in the ligation mix was detected at a size equal to pGSS15. This suggested that at least some of the labelled octanucleotide had been ligated to fragments which produced a product equivalent in size to pGSS15, and there appeared to be more than two colonies-worth of DNA!

To recover parental plasmid from a ligation reaction of this type indicates that at least one fragment preparation was contaminated with uncut or linear pGSS15 DNA. The most likely candidate for this would be the large *Bst*EII-*Sal*I fragment of pGSS15. The experiment was repeated with this fragment gel-purified, re-cut and re-purified. However the end result was similar to before with only one $Ap^r$ colony being isolated which again had two *Sal*I sites.

# CHAPTER 6

## Uses of pGSS15 as a broad host range cloning vector

### 6.1 Expression of mouse dihydrofolate reductase coding sequence in *M.methylotrophus*.

To assess the potential of pGSS15 as a broad host range cloning vector, and to provide a test system for the expression of a eukaryotic coding sequence in *M.methylotrophus*, a cDNA sequence encoding mouse dihydrofolate reductase (DHFR) was cloned into pGSS15. Subsequent transfer of the recombinant plasmid to *M.methylotrophus* allowed expression levels of the enzyme to be measured. The system provided a preliminary demonstration of the methylotroph's potential for commercial production of mammalian peptides.

Cloning of eukaryotic cDNA into the *Pst*I site of the $Ap^r$ gene of pBR322 by dG.dC-tailing has been demonstrated to lead to the expression of either fused (Itakura *et al.*,1977) or native (Pasek *et al.*,1979) eukaryotic polypeptides, at least in *E.coli*. In most cases the cDNA fragment, together with its short GC tails, can be excised by *Pst*I digestion.

Copy DNA derived from the mRNA for mouse DHFR has been cloned and expressed in *E.coli* (Chang *et al.*,1978); bacteria which synthesize mouse DHFR can be selected because they are resistant to levels of trimethoprim (Tp) that would normally inhibit growth.

The plasmid pDHFR7 (Chang *et al.*,1978) was constructed by annealing a deoxy-C-tailed mouse cDNA to pBR322 DNA which had been cleaved in the β-lactamase gene by *Pst*I and tailed with deoxy-G-residues. The *Pst*I sites were regenerated on both ends of the insert providing a convenient source of the cDNA fragment. 1μg of this DNA was mixed with 0.5μg of *Pst*I-digested pGSS15 DNA and

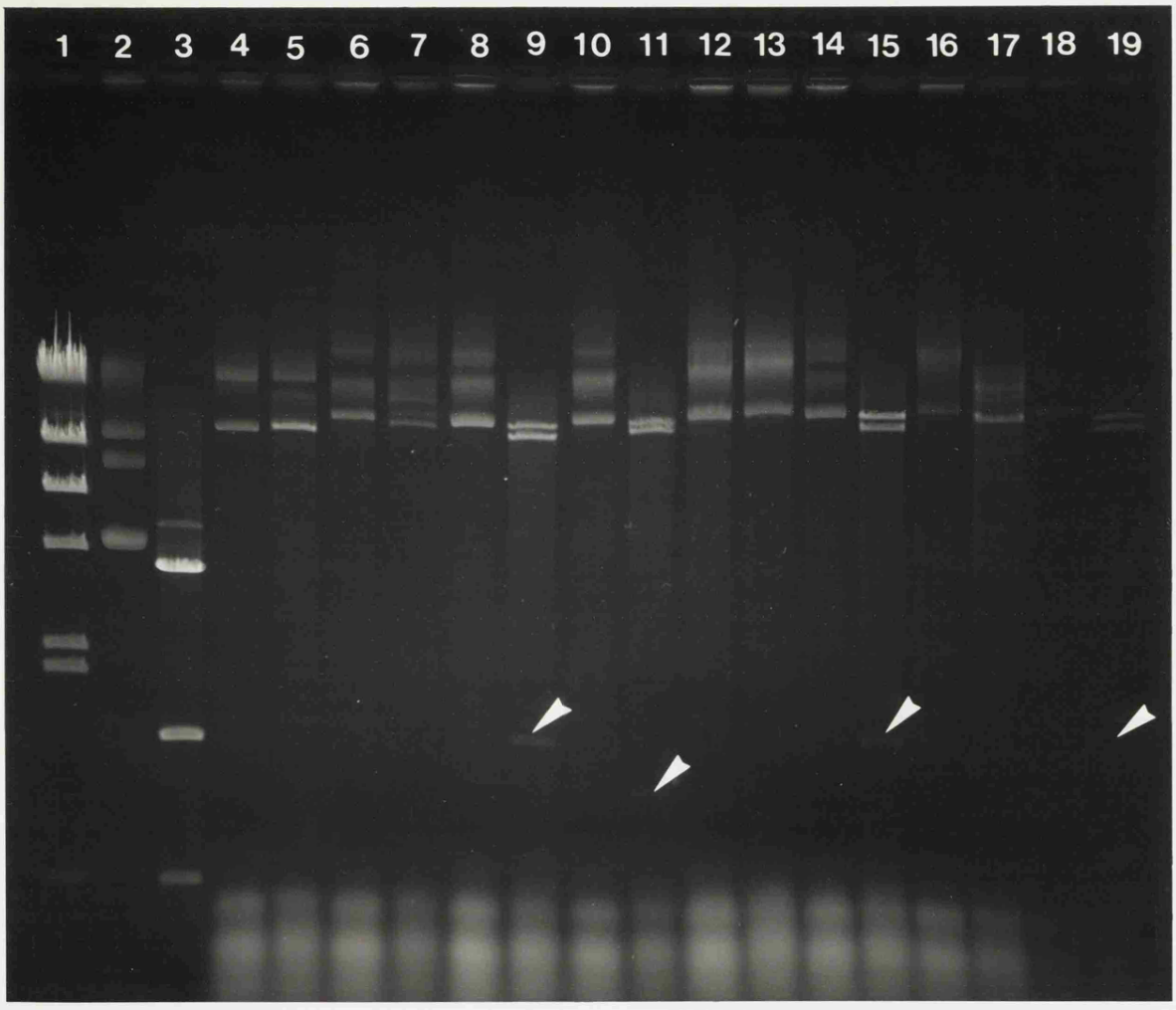
treated with T4 DNA ligase at 12°C overnight. Transformation of *E.coli* strain ED8654 with this ligated mixture yielded 221 $Tp^r$ clones, of which 216 were $Ap^s$, showing that the DHFR fragment had been inserted into the β-lactamase gene. Selection for $Tc^r$ with another aliquot of the ligation mixture yielded several thousand resistant colonies, mostly containing parental pGSS15, but which included some $Ap^s$, $Tp^r$ colonies. Six randomly sel ected $Ap^s$, $Tp^r$ colonies and one $Ap^s$, $Tp^s$ colony were examined for plasmid content, and all contained a plasmid larger than pGSS15. The DHFR cDNA has a single *Bgl*II site which allows orientation of the fragment. Restriction digests with *Eco*RI+*Bgl*II on seven of these colonies [Fig.24] established that only one orientation, that identical to pDHFR7, led to functional expression of the DHFR sequence. This confirmed that productive transcription was from the β-lactamase promoter. The plasmid isolated from one of these $Tc^r$, $Tp^r$, $Ap^s$ colonies was designated pDHFR2.43.

The IncIα plasmid R64*drd*11 was introduced into the *E.coli* strain carrying pDHFR2.43, and used to mobilize this hybrid plasmid into *M.methylotrophus*. Table II shows the levels of DHFR activity in extracts of *E.coli*(pDHFR2.43) and *M.methylotrophus*(pDHFR2.43). When trimethoprim (100μg/ml) was added to the *M.methylotrophus* extracts, complete inhibition of the bacterial enzyme was observed, but the activity in the pDHFR2.43 extract was reduced to only ≈25%. This was consistent with observations that the mammalian enzyme is less prone to inhibition by trimethoprim (Burchall,1971). Furthermore, changing the substrate to folate instead of dihydrofolate abolished the activity of the bacterial enzyme while the mammalian enzyme retained a low level of activity (Baccanari *et al.*,1975). These two observations confirm that the trimethoprim resistance of *M.methylotrophus* carrying pDHFR2.43 is a consequence of the

**Figure 24.**

1% agarose gel electrophoresis of plasmid DNAs prepared from recombinants of pGSS15 and pDHFR7 after digestion with *Eco*RI+*Bgl*II.

Track 3 shows an *Eco*RI+*Bgl*II digest of pDHFR7 and the plasmids in tracks 9, 15 and 19 all have an identical fragment (arrowed) indicating that the insert orientations were the same as pDHFR7. Track 11 shows the fragment size from a clone with the insert in the opposite orientation and this plasmid did not give trimethoprim resistance. Other tracks represent plasmids without insertions.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Cultures for the DHFR assays were grown with shaking at 37°C overnight, collected by centrifugation and the pellets resuspended in 3 vol 50mM potassium phosphate buffer (pH7). The cells were disrupted by sonication on ice in the presence of 10μM phenylmethyl-sulphonyl fluoride and 10μM benzamidine. Cell debris was removed by centrifugation at 10,000 × g for 15min, then 100,000 × g for 60min. Enzyme assays were performed at 37°C. The standard assay mixture contained 50μM dihydrofolate, 60μM NADPH, 12mM 2-mercaptoethanol, 50mM poassium phosphate buffer pH7 and enzyme (normally 100μl of bacterial extract) in a final volume of 1.0ml. The oxidation of NADPH was followed at 340nm for ≈10min. Enzyme activity is expressed as nmol of NADPH oxidised per min per mg protein (mean ± s.e.m.). Numbers in parentheses are the number of separate experiments. Protein concentration was determined by the method of Lowry *et al.* (1951).

**Table II** Dihydrofolate reductase activities of pDHFR 2.43 in extracts of *E. coli* ED 8654 and *M. methylotrophus*

| Strain | Substrate | Addition | Activity |
|---|---|---|---|
| *E. coli* ED 8654 | Dihydrofolate | — | 0.70 ± 0.14 (4) |
| *E. coli* ED 8654 (pDHFR 2.43) | Dihydrofolate | — | 1.58 ± 0.26 (5) |
| *M. methylotrophus* | Dihydrofolate | — | 2.34 ± 0.68 (5) |
| *M. methylotrophus* | Dihydrofolate | Trimethoprim* | 0 (2) |
| *M. methylotrophus* | Folate† | — | 0 (2) |
| *M. methylotrophus* (pDHFR 2.43) | Dihydrofolate | — | 12.6 ± 1.7 (4) |
| *M. methylotrophus* (pDHFR 2.43) | Dihydrofolate | Trimethoprim* | 2.57 ± 0.53 (4) |
| *M. methylotrophus* (pDHFR 2.43) | Folate† | — | 0.50 (1) |

* Trimethoprim (100 μg $ml^{-1}$) was added to the enzyme assays.
† Folate (70 μM) was included in the assay mixture instead of dihydrofolate.

expression of the mouse dihydrofolate reductase coding sequence.

As pGSS15 retains the broad host range characteristics of R300B, I have been able to transfer pDHFR2.43 into another methylotroph, *Methylomonas methylovora*, and into an autotrophic bacterium, *Alcaligenes eutrophus*. Both these strains were rendered trimethoprim resistant and the resistance was shown to correspond to the acquisition of pDHFR2.43, as observed by electrophoretic analysis of small scale alkaline-SDS lysates.

**6.2 High expression vectors from pGSS15.**

The presence of the extra *Sal*I site in the repeated sequence of pGSS15 was, in most cases, inconvenient rather than a serious problem, and had not prevented the plasmid from being used successfully. Thus a logical progression from this vector would be to provide for increased levels of expression of cloned coding sequences, from both prokaryotes and eukaryotes. This may be achieved by the introduction of a suitable promoter, recognised by the host RNA polymerase, and a translation initiation signal, to allow the ribosomes to bind to the resultant mRNA. In the case of an *E.coli* mRNA, a ribosome binding site includes the translational start codon (AUG or GUG) and another sequence that is complementary to bases on the 3' end of 16S ribosomal RNA – the so-called Shine-Dalgarno (SD) sequence (Shine and Dalgarno,1975). Translation initiation signals can be provided either by joining a eukaryotic coding sequence to the beginning of a bacterial structural gene, resulting in a fused polypeptide product (Itakura *et al.*,1977; Mercerau-Puijalon *et al.*, 1978; Villa-Komaroff *et al.*,1978; Martial *et al.*,1979; Emtage *et al.*, 1980), or by constructing a hybrid

ribosome-binding site with a bacterial SD sequence followed by a translation initiation codon at the beginning of the eukaryotic gene (Goeddel *et al.*,1979; Roberts *et al.*,1979; Edman *et al.*, 1981). This latter case is usually more desirable since either a normal eukaryotic protein, or one carrying an additional N-terminal methionine residue, is produced (Wetzel *et al.*,1981; Olsen *et al.*, 1981). As described earlier however, the mouse DHFR coding sequence was translated from its own initiator codon when cloned by dG:dC tailing into the *Pst*I site of pBR322 (and also when sub-cloned into pGSS15) even though this site was within the β-lactamase gene. The oligo-dG tract preceding the initiation codon may provide some SD sequence homology in this case (Chang *et al.*, 1978).

The most extensively used bacterial promoters for expression systems have been the *E.coli lac* (Goedel *et al.*,1979; Roberts *et al.*, 1979; Edge *et al.*,1981; DeMaeyer *et al.*,1982) and *trp* promoters (Emtage *et al.*,1980; Edman *et al.*,1981). They have been particularly suitable because they have been so well characterized, their control is well understood and they can promote highly efficient transcription.

**6.2(i) Introduction of the *lac* and *lac*UV5 promoters into pGSS15 and expression of a synthetic human α-1 interferon gene in *M.methylotrophus*.**

The *lac* promoter/operator region from the *E.coli* lactose operon can be isolated on a 95-bp *Alu*I-generated fragment. This has been cloned into pAT153, between filled-in *Eco*RI and *Bam*HI sites, to produce the derivative plasmid pPM50 (Edge *et al.*,1981). The construction allowed the *Eco*RI and *Bam*HI site to be regenerated thus producing a convenient "portable" *lac* promoter fragment. Expression of a synthetic human α-1 interferon gene from this *lac* promoter, on

plasmid pIFS1, has been demonstrated (Edge *et al.*,1981) and the *lac* promoter-synthetic interferon fragment was transferred from pIFS1 to pGSS15 and used to demonstrate expression in *M.methylotrophus* (DeMaeyer *et al.*, 1982).

These initial constructions were made using the wild type *E.coli* promoter fragment and were found to produce only low level expression. This was because the fragment lacked a sequence which is recognised by the positive activator of the *lac* promoter, catabolite activating protein (CAP). Maximal expression from the wild type *lac* promoter in *E.coli* requires stimulation by CAP. A plasmid carrying a similar fragment to that in pPM50 has been constructed, but in this case the fragment was derived from a mutant lactose operon promoter, UV5 (Miller and Reznikoff,1978) and is designated pPM60 (P.Meacock, personal communication). When the synthetic interferon gene was transcribed from the *lac*UV5 promoter, as in plasmid pIFS101 (DeMaeyer *et al.*,1982), antiviral activity was increased almost ten fold over that from pIFS1.

To compare the effect of changing from the wild type *lac* promoter to the *lac*UV5 mutant promoter in *M.methylotrophus*, I subcloned the latter promoter, together with the synthetic interferon gene, from pIFS101 into pGSS15. The fragment was readily obtained from an *Eco*RI+*Sal*I digest of pIFS101 (5$\mu$g) and electroelution of the DNA from the band corresponding to ≈600bp after electrophoresis through a 1%agarose gel. The two *Sal*I sites in pGSS15 made the sub-cloning more complicated than it would otherwise have been, but the approach described below was relatively straight forward and provided unambiguous fragments for the ligation step.

Thus 20$\mu$g of pGSS15 DNA were digested with 20units of *Bst*EII, in 100$\mu$l total volume, for 90min at 60°C. The reaction was terminated by phenol extraction and ethanol precipitation. The DNA was

resuspended in 22μl of water and digested with *Eco*RI in a 25μl reaction. After electrophoresis through a 0.8% agarose gel, the two *Bst*EII-*Eco*RI fragments were cut from the gel and the DNA electroeluted. The larger *Bst*EII-*Eco*RI fragment was then cut with *Sal*I and the larger *Bst*EII-*Sal*I fragment gel-purified as above. Three fragments were thus prepared for ligation: the smaller *Eco*RI-*Sal*I pIFS101 fragment; the small *Bst*EII-*Eco*RI fragment from pGSS15; and the large*Bst*EII-*Sal*I fragment from pGSS15, [Fig.25]. Each of these purified fragments was resuspended in 20μl of water of which 10μl of each were included in a 150μl ligation reaction with 1 unit of T4 DNA ligase. After incubation at 10°C overnight, 100μl of the mixture were used for the transformation of *E.coli* strain C600. [The transformation was undertaken by David Pioli within the GMAG CatII laboratory at ICI, Runcorn]. A small test portion of the transformation mix (0.5%) was plated out and gave 14 $Tc^r$ colonies. Five out of six of these colonies, chosen at random, were found to be $Tc^r$,$Ap^s$ and their plasmid content, as analysed using a small-scale SDS lysis method (Birnboim and Doly,1979), showed a *Hinf*I restriction pattern consistent with the predicted structure (D.Pioli, personal communication). Thus pGSS15 contained the synthetic interferon gene transcribed from an *E.coli lac*UV5 promoter and no longer carried the pBR322 *ori*.

Assays for the antiviral activity for pGSS15:IFS1 and pGSS15:IFS101 in *M.methylotrophus* confirmed a similar situation to that found in *E.coli*, the *lac*UV5 promoter was more efficient than the wild type *lac* promoter, Table III.

**6.2(ii) Introduction of *trp* promoters into pGSS15 and expression of the *E.coli lacZ* gene in *M.methylotrophus*.**

To make available the *E.coli trp* promoter on a broad host range

**Figure 25.**

Diagram of the approach used to sub-clone the *lac*UV5-synthetic interferon fragment from pIFS101 into pGSS15.

Since pGSS15 contained two *Sal*I sites, two separate fragments (I and II) had to be isolated to allow insertion of the *Eco*RI-*Sal*I fragment.

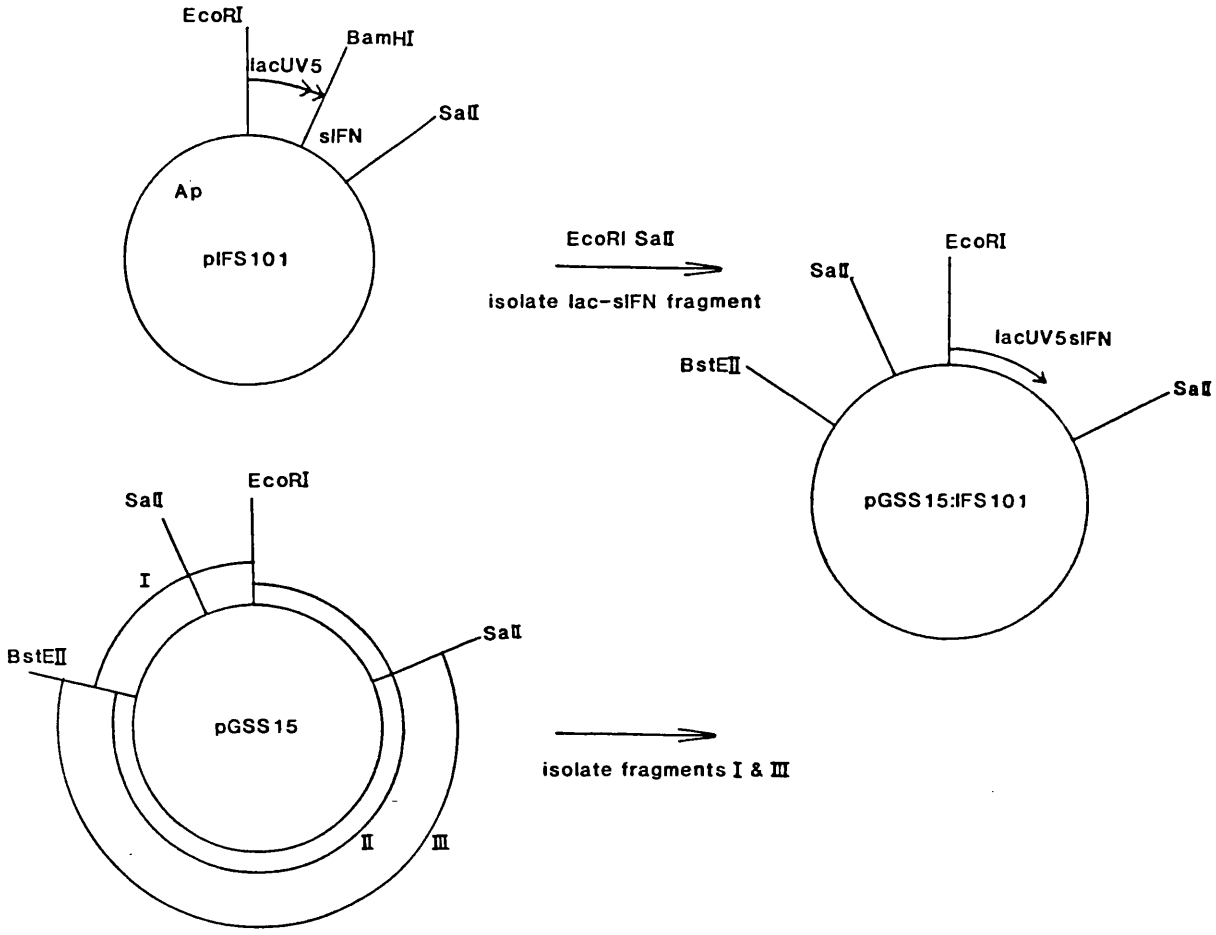

EcoRI
BamHI
lacUV5
SalI
sIFN
Ap
pIFS101
EcoRI SalI
isolate lac-sIFN fragment
SalI
EcoRI
lacUV5sIFN
BstEII
SalI
pGSS15:IFS101
SalI
EcoRI
I
SalI
BstEII
pGSS15
II
III
isolate fragments I & III

Table III

Comparison of α-1 interferon expression levels in *E.coli* and *M.Methylotrophus*.

| Plasmid† | Promoter | Interferon activity* | |
|---|---|---|---|
| | | *E.coli* | *M.meth.* |
| pPM60 | - | 300 | - |
| pGSS15 | - | 300 | 300 |
| pIFS1 | wt *lac* | $1.9.10^4$ | - |
| pGSS15:sIFN1 | wt *lac* | $1.9.10^4$ | $3.8.10^4$ |
| pIFS101 | *lac*UV5 | $1.3.10^6$ | - |
| pGSS15:sIFN101 | *lac*UV5 | $3.0.10^5$ | $1.2.10^6$ |

* assayed on MDBK cells; expressed as antiviral units/0.15ml of cell extract.

† pGSS15:sIFN1 and pGSS15:sIFN101 contain the *Eco*RI-*Sal*I "expression units" from pIFS1 and pIFS101 respectively.

The interferon assays were carried out at ICI Pharmaceuticals Division by Val Moore and were as described in De Maeyer *et al.* (1982).

MDBK cells are a bovine kidney cell line as described by Madin and Darby 1958.

plasmid I decided to replace the promoter from the $Tc^r$ gene of pGSS15 with a *trp* promoter fragment. There were two convenient sources of *trp* promoters obtainable on fragments which could be readily inserted into pGSS15 without affecting the phenotype of the latter. A wild type *E.coli trp*E gene, together with its promoter and operator was available from the plasmid pDBE102 (Burt,1981). However, a second and more elegant source was an 82-bp synthetic DNA fragment encoding the *E.coli trp* promoter and operator sequences together with the first Shine-Delgarno sequence of the *trp* operon. This fragment had been cloned into pAT153 to give the plasmid pSTP2 (Windass *et al.*,1982). In the construction of pSTP2, the promoter of the $Tc^r$ gene had been replaced by the synthetic *trp* promoter, in a similar way to that envisaged for pGSS15. This type of sub-cloning then becomes very easy and demonstrates the convenience of having the resistance genes arranged as in pBR322 and pAT153, since an *Eco*RI-*Bam*HI fragment could now be used.

Thus 10μg each of pGSS15 and pSTP2 DNAs were digested with *Eco*RI+*Bam*HI and the products separated by electrophoresis through a 0.8% agarose gel. Bands corresponding to the 11.2-kb of pGSS15 and the 435-bp fragment of pSTP2 were cut out of the gel and the DNA electroeluted. The purified and lyophilized DNA pellets were each dissolved in 50μl of water and 12.5μl of each solution were then included in a 50μl ligation reaction. After incubation at 22°C for 1hr the mixture was used for the transformation of *E.coli* strain ED8654 with selection for resistance to tetracycline. Twenty seven resistant colonies were obtained, four of which, selected at random, were all shown to contain a plasmid with a newly-acquired unique *Hpa*I site. The plasmid was designated pGSS18.

Plasmid pDBE102, containing the wild type *E.coli trp*E promoter, operator and leader sequences, was digested with *Eco*RI and *Hind*III

(10μg of DNA in a 50μl reaction). Without further purification, 5μl of this digestion mix were ligated with 5μl of similarly digested pGSS15 DNA, in 50μl at 16°C overnight. Transformation of *E.coli* strain BE6 to ampicillin-resistance and *trp+*, produced three colonies which were each found to contain a plasmid larger than pGSS15. Analysis of the plasmid DNA with *Eco*RI+*Hind*III revealed a fragment with a size identical to that in pDBE102. This plasmid was designated pGSS19.

The plasmids pGSS18 and pGSS19 were used as *Hind*III-cloning vectors in an experiment to quickly establish if the *E.coli lac*Z gene could be expressed in *M.methylotrophus*. Such an experiment would test the efficacy of including the chromogenic substrate, 5-bromo-4-chloro-3-indolyl-B-D-galoactoside (X-gal) in methanol minimal plates.

[X-gal is not an inducer of B-galactosidase but is cleaved by B-galactosidase releasing a blue indolyl derivative. Since X-gal is not an inducer, only mutants constitutive for B-galactosidase produce blue colonies on medium containing X-gal].

The bacteriophage λDB219 (Burt and Brammar,1982) carries an intact *E.coli lac*Z gene which could be excised on an 8.45-kb *Hind*III fragment and inserted into the unique *Hind*III sites of pGSS18 or pGSS19. Thus 3μg of pGSS18 DNA and 3μg of λDB219 DNA were mixed and digested with 6 units of *Hind*III for 90min at 37°C. In a separate re action, 3μg of pGSS19 DNA and 3μg of λDB219 DNA were similarly digested. After phenol extraction and ethanol precipitation, to kill the enzyme activity, the DNA pellets were resuspended and treated with T4 DNA ligase at 12°C overnight. [This experiment was undertaken in collaboration with David Burt and the transformation and selections described below, were done by him].

Transformation of *E.coli* strain DB202 (Burt,1981), with selection

for $Ap^r$ on Lac-MacConkey (Oxoid) agar, produced *lac+* clones which were identified by a red pigmentation and *lac*$^-$ colonies which were white. Several *lac+* colonies were picked and their plasmid DNAs analysed by restriction with *Eco*RI. Both orientations of the *lac*Z fragment were obtained in pGSS18 and are shown diagrammatically in Fig.26.

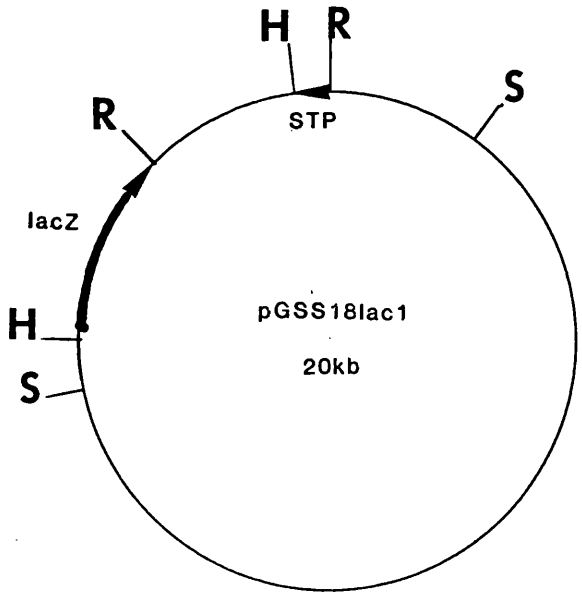
The *lac+* plasmids were subsequently mobilized into *M.methylotrophus* using the Inc Iα plasmid R64*drd*11. When cultures of *M.methylotrophus* containing pGSS18lac1, pGSS18lac13 or pGSS19lac7 were plated out onto methanol minimal plates containing X-gal, all plasmid-containing colonies were blue.

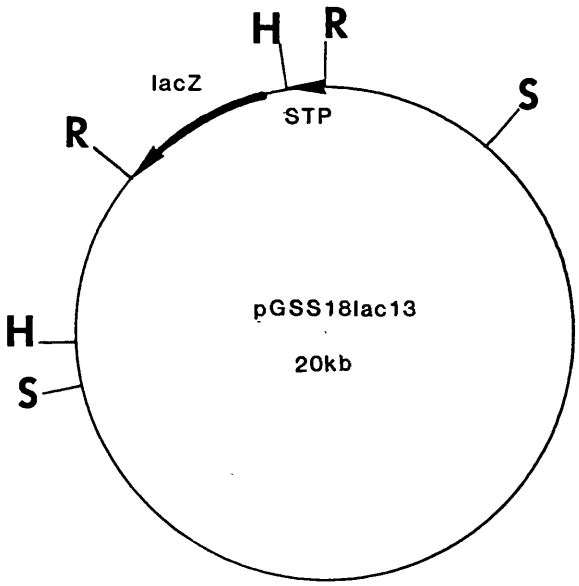
This experiment, although fairly crude, demonstrated for the first time that the *E.coli lac*Z gene could be expressed efficiently in *M.methylotrophus* and that inclusion of X-gal into methanol minimal plates would allow chromogenic testing for *lac+* or *lac*$^-$ colonies.

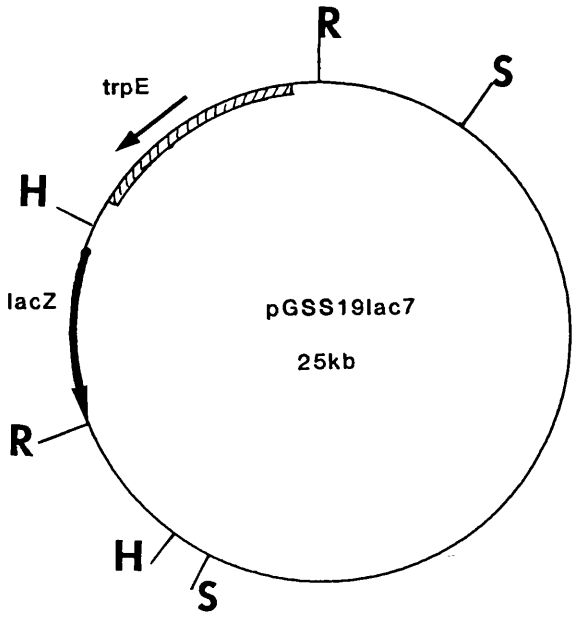
These demonstrations of expression of heterologous coding sequences in *M.methylotrophus*, using the specially constructed broad host range plasmids, are encouraging and suggest that there is now the basis of a viable alternative host/vector system.

Figure 26.

Diagrams (not to scale) to show relative orientations of the *E.coli lacZ* gene inserted into either pGSS18 (having the synthetic *trp* promoter, STP) or pGSS19 (having the *trpE* promoter/operator region).


H
R
S
R
STP
lacZ
pGSS18lac1
20kb
H
S


H
R
lacZ
S
R
STP
pGSS18lac13
20kb
H
S


R
S
trpE
H
lacZ
pGSS19lac7
25kb
R
H
S

# CHAPTER 7

## 7.1 Insertion of the $Cm^r$ gene from pBR328 into some broad host range vectors

The plasmid pBR325 is a $Cm^r$ derivative of pBR322 constructed by the insertion of an S1 nuclease-treated *Hae*II fragment from phage P1$Cm^r$ into *Eco*RI-digested pBR322 after treatment with S1 nuclease (Bolivar,1979). A deletion derivative of pBR325, pBR328, lacks 1089bp of DNA between the vegetative origin of replication and the $Tc^r$ gene (Soberon *et al.*,1980). These plasmids now contain a unique *Eco*RI site within the $Cm^r$ gene which, when inactivated by insertion of a DNA fragment, allow easy selection of recombinant molecules. The $Cm^r$ gene could be excised from pBR328 on a *Bam*HI-*Pst*I fragment and readily inserted into several of the pGSS plasmids described earlier. This would provide a useful alternative selection marker and, in some cases, a unique *Eco*RI cloning site.

Digestion of pGSS6 with *Bam*HI+*Pst*I would remove the whole of the $Ap^r$ gene, the pBR322 *ori*, most of the $Su^r$ gene and some of the $Tc^r$ gene. Insertion of the *Bam*HI-*Pst*I $Cm^r$ fragment from pBR328 would then reconstitute the $Tc^r$ gene, introduce a $Cm^r$ gene and provide the β-lactamase promoter in the correct orientation to maintain transcription of the $Sm^r$ gene. The same *Bam*HI-*Pst*I $Cm^r$ fragment inserted into *Bam*HI+*Pst*I digested pGSS15 would reconstitute both $Tc^r$ and $Ap^r$ genes while introducing the $Cm^r$ gene in an analogous position, relative to Ap and Tc, as it was in pBR328.

A third vector which could benefit from having a $Cm^r$ gene was pWS1.1, a derivative of R300B that has a 2.1kb *Eco*RI-*Pvu*II fragment containing $Sm^r$ replaced by a similar sized *Eco*RI-*Pvu*II fragment from pBR322 containing the $Tc^r$ gene (W. Schuch, personal communication).

(I observed that this plasmid must have a very high copy number as yields of plasmid DNA were consistently high, e.g. approximately 3mg of purified plasmid DNA were obtained from a 400ml shake culture grown to stationary phase!). The $Su^r$ gene of pWS1.1 did not allow a very clean selection when the $Tc^r$ gene was inactivated and $Cm^r$ would represent a significant improvement, as well as bringing in a unique *Eco*RI cloning site.

Thus 5$\mu$g each of pGSS6, pGSS15 and pWS1.1 DNAs were digested in 25$\mu$l reactions with 5units each of *Bam*HI and *Pst*I at 37°C for 2hr. 20$\mu$g of pBR328 DNA in a 100$\mu$l reaction were similarly digested. The enzymes were inactivated by heating at 70°C for 10min, then chilling on ice.

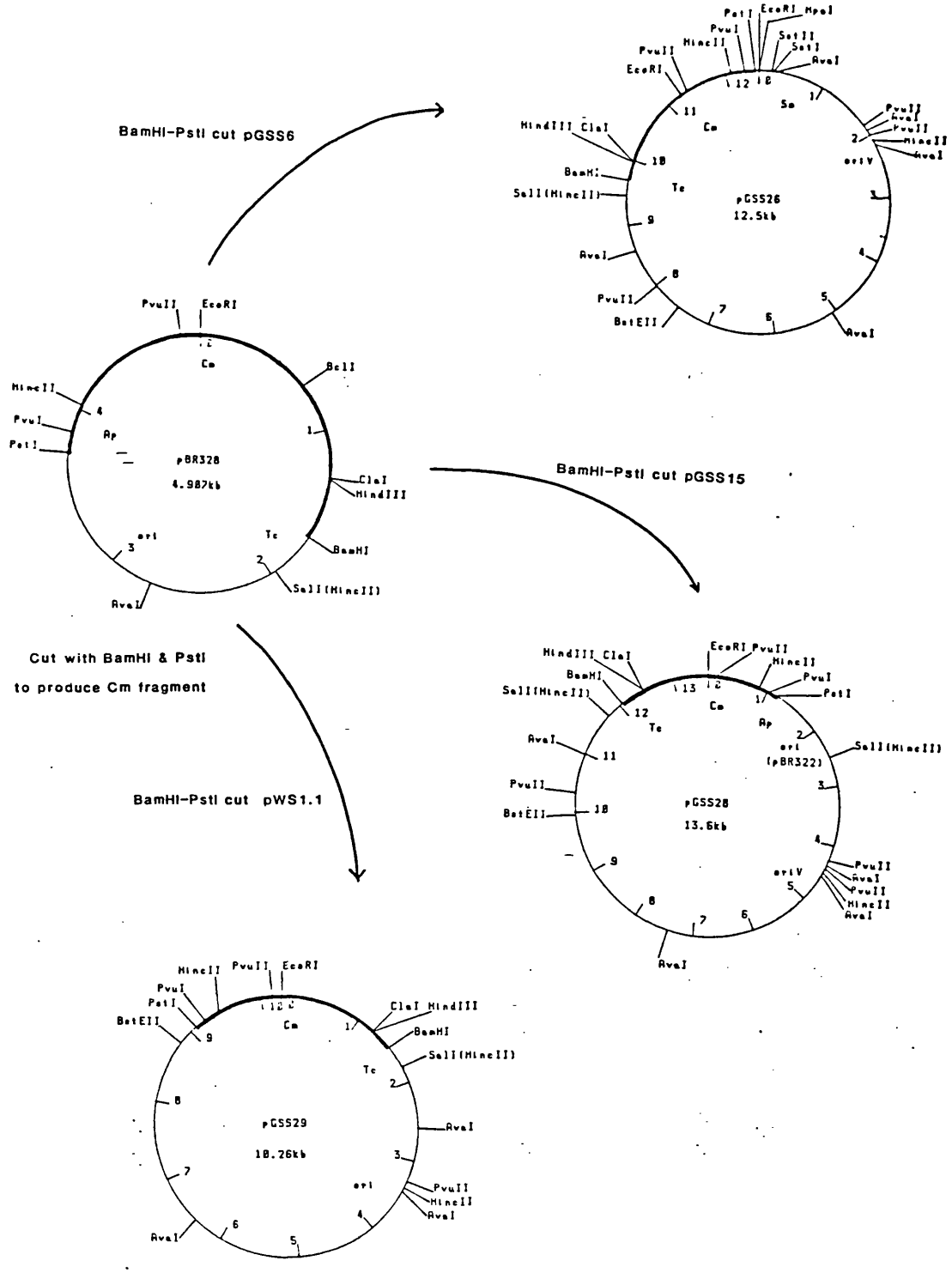
75$\mu$l reactions were made, each containing 5$\mu$l of the cut pBR328, 5$\mu$l of cut pGSS6, pGSS15 or pWS1.1, and T4 DNA ligase. After incubation at 22°C for 90min 10$\mu$l of each ligation mix was used for transformation of *E.coli* strain C600 and plated out with selection for resistance to Cm. 17 out of 67 $Cm^r$ colonies from the ligation with pGSS6 were also $Tc^r$, $Sm^r$, $Ap^s$. 8 out of 21 $Cm^r$ colonies from the ligation with pWS1.1 were found to be $Tc^r$ and $Ap^s$, and 3 out of 25 $Cm^r$ colonies from the ligation with pGSS15, picked at random, were all found to be the expected $Cm^r$ derivative. The $Cm^r$ versions of pGSS6, pGSS15 and pWS1.1 were designated pGSS26, pGSS28 and pGSS29, respectively [Fig.27]. Plasmid pGSS29 still appeared to have a high copy number.

**7.2 Construction of pGSS33**

In the construction of pGSS26 from pGSS6, the pBR322 *ori* had been removed and pGSS26 still contained over half of the β-lactamase gene. If this gene could be completed and the extra *Eco*RI site, derived

**Figure 27.**

Insertion of the $Cm^r$ *Bam*HI-*Pst*I fragment from pBR328 into pGSS6, pGSS15 and pWS1.1. The resulting plasmids were designated pGSS26, pGSS28 and pGSS29 respectively and show the position of the insert as a thicker line.


BamHI-PstI cut pGSS6
pGSS26
12.5kb
pBR328
4.987kb
BamHI-PstI cut pGSS15
Cut with BamHI & PstI
to produce Cm fragment
BamHI-PstI cut pWS1.1
pGSS28
13.6kb
pGSS29
10.26kb

from R300B, could be removed, then a very useful cloning vector would result giving a choice of four antibiotic resistance genes all carrying at least one restriction site capable of producing insertional inactivation. Inspection of the small amount of DNA sequence data from around the *Eco*RI site of R300B [Fig.21] revealed that the *Hpa*I site was immediately adjacent to the *Eco*RI site , and about 160-bp away from what appeared to be the start of the $Sm^r$ gene. Therefore, a deletion of the *Pst*I-*Hpa*I fragment from pGSS26 would remove the extra *Eco*RI site without interfer ing with the $Sm^r$ gene. This would allow the $Ap^r$ gene to be reconstituted by insertion of a *Pst*I-blunt-ended fragment containing the appropriate sequence.

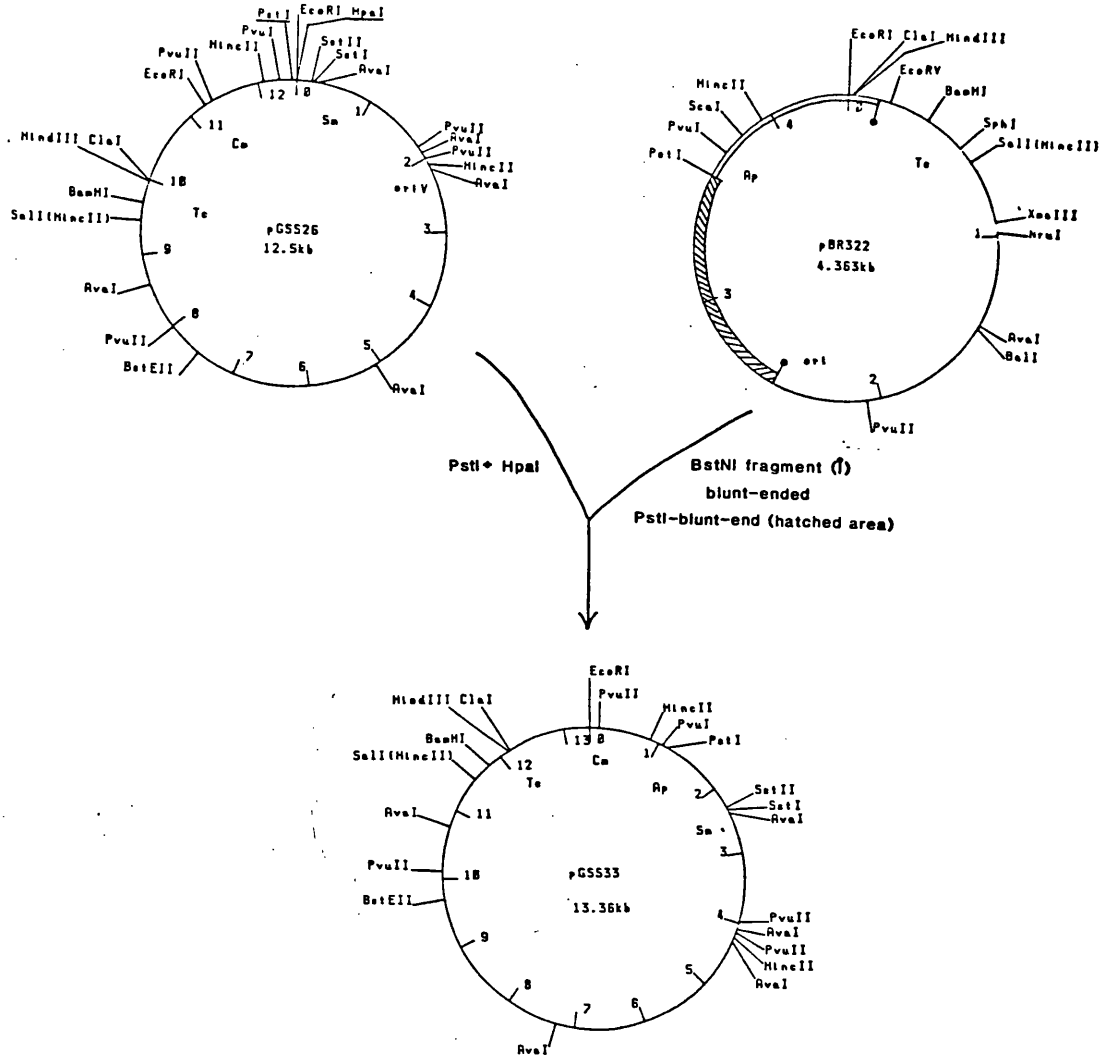
From the sequence of the $Ap^r$ gene in pBR322 no unique target sites for restriction enzymes producing blunt-ends, between the *Pst*I site and the pBR322 *ori*, could be found. However, a 1857-bp *Bst*NI fragment could be isolated from pBR322 which contained the whole $Ap^r$ gene [Fig.28]. *Bst*NI recognises the sequence $^{5'}CC(A/T)GG^{3'}$ and at either end of the fragment of interest, the sequence was $^{5'}CCTGG^{3'}$, leaving a single 5'-T extension after digestion with the enzyme. This was made flush-ended by the addition of $\alpha^{32}P$ dATP in the presence of large fragment polymerase. Subsequent digestion of the filled-in fragment with *Pst*I+*Hinc*II allowed a 978-bp *Pst*I-blunt-ended fragment to be isolated, on a 1% agarose gel, away from other fragments of 286bp and 586bp.

2$\mu$g (10$\mu$l) of pGSS26 DNA that had been digested with *Pst*I+*Hpa*I were mixed with half of the quantity of purified fragment (10$\mu$l) in a 30$\mu$l reaction and treated with T4 DNA ligase for 16hr at 12°C. A 10$\mu$l aliquot of the ligation mix was used for transformation of *E.coli* strain C600 with selection for resistance to Ap. 31 $Ap^r$ colonies were obtained of which 20 were shown to be also resistant Tc, Cm and Sm. The plasmid from one of these colonies was designated

**Figure 28.**

Construction of pGSS33 from pGSS26 and pBR322. (Sharpe,1984)

The *Bst*NI $Ap^r$ fragment of pBR322 (bounded by the black dots) was isolated, made blunt-ended, then cut with *Pst*I+*Hinc*II to allow a *Pst*I-blunt-ended fragment (shaded) to be purified. This fragment was then inserted into pGSS26, between the *Pst*I and HpaI sites (underlined), to reconstitute the $Ap^r$ gene and remove the extra *Eco*RI site. The arrangement of the $Ap^r$, $Cm^r$ and $Tc^r$ genes in pGSS33 then becomes identical to that in pBR328.


PstI
EcoRI HpaI
PvuI
HincII
PvuII
EcoRI
SstII
SstI
AvaI
Sm
Cm
pGS526
12.5kb
ori V
HindIII ClaI
BamHI
SalI(HincII)
Te
AvaI
PvuII
BstEII
PvuII
AvaI
PvuII
HincII
AvaI
EcoRI ClaI HindIII
EcoRV
BamHI
SphI
SalI(HincII)
XmaIII
NruI
AvaI
BclI
PvuII
HincII
ScaI
PvuI
PstI
Ap
Te
pBR322
4.363kb
ori
PstI + HpaI
BstNI fragment (1)
blunt-ended
PstI-blunt-end (hatched area)
EcoRI
PvuII
HincII
PvuI
PstI
HindIII ClaI
BamHI
SalI(HincII)
Te
Cm
Ap
SstII
SstI
AvaI
Sm
AvaI
PvuII
BstEII
pGS533
13.36kb
PvuII
AvaI
PvuII
HincII
AvaI

pGSS33 and subjected to restriction analysis.

There was now only one site for *Eco*RI, as expected, and the size of the *Pst*I-*Sst*I fragment in pGSS33 was ≈1200bp compared to only ≈300bp in pGSS26, and corresponded to a net insertion of ≈900bp. This was consistent with removal of 80-90bp of DNA between the *Pst*I and *Hpa*I sites of pGSS26 and replacement with a 987-bp fragment [Fig.29].

Digestion of pGSS26 with *Ava*I produced four fragments, the largest of which included the *Eco*RI and *Hpa*I sites, and this was the only fragment observed to be larger when pGSS33 was digested with *Ava*I. The recognition sequence for *Hpa*I (GTTAAC) is also recognised by *Hinc*II (GTPyPuAC), however when a filled-in *Bst*NI sequence was joined to the *Hpa*I-cut sequence of pGSS26, both the *Hpa*I and *Hinc*II sites were lost. *Hinc*II also cut pGSS26 at three other sites producing fragments of 7.2, 2.7, 2.2 and 0.4kb, with the *Hpa*I site forming the junction between the 0.4kb and 2.2kb fragments. When pGSS33 was digested with *Hinc*II, both of these fragments disappeared to be replaced by a new fragment of 3.5kb, which include the newly inserted *Pst*I-blunt-end fragment. The other two fragments remained unchanged.

The smallest (2.7kb) of the three major *Pvu*II fragments of pGSS26, which included the *Pst*I and *Hpa*I sites, did not appear in a *Pvu*II digest of pGSS33, but there was now a new band of 3.6kb, again confiming the position and size of the insert. *Bam*HI, *Cla*I, *Eco*RI, *Hind*III, *Pst*I, *Pvu*I, *Sal*I, *Sst*I and *Sst*II all cut pGSS33 only once and hence these enzymes could be used to generate DNA fragments for cloning into appropriately linearized pGSS33. There is however redundancy in the recognition sequences of some restriction enzymes e.g. *Bam*HI recognises and cuts the sequence $^{5'}G\downarrow GATCC^{3'}$ leaving a four base single-stranded 5' extension of GATC. This is the same as

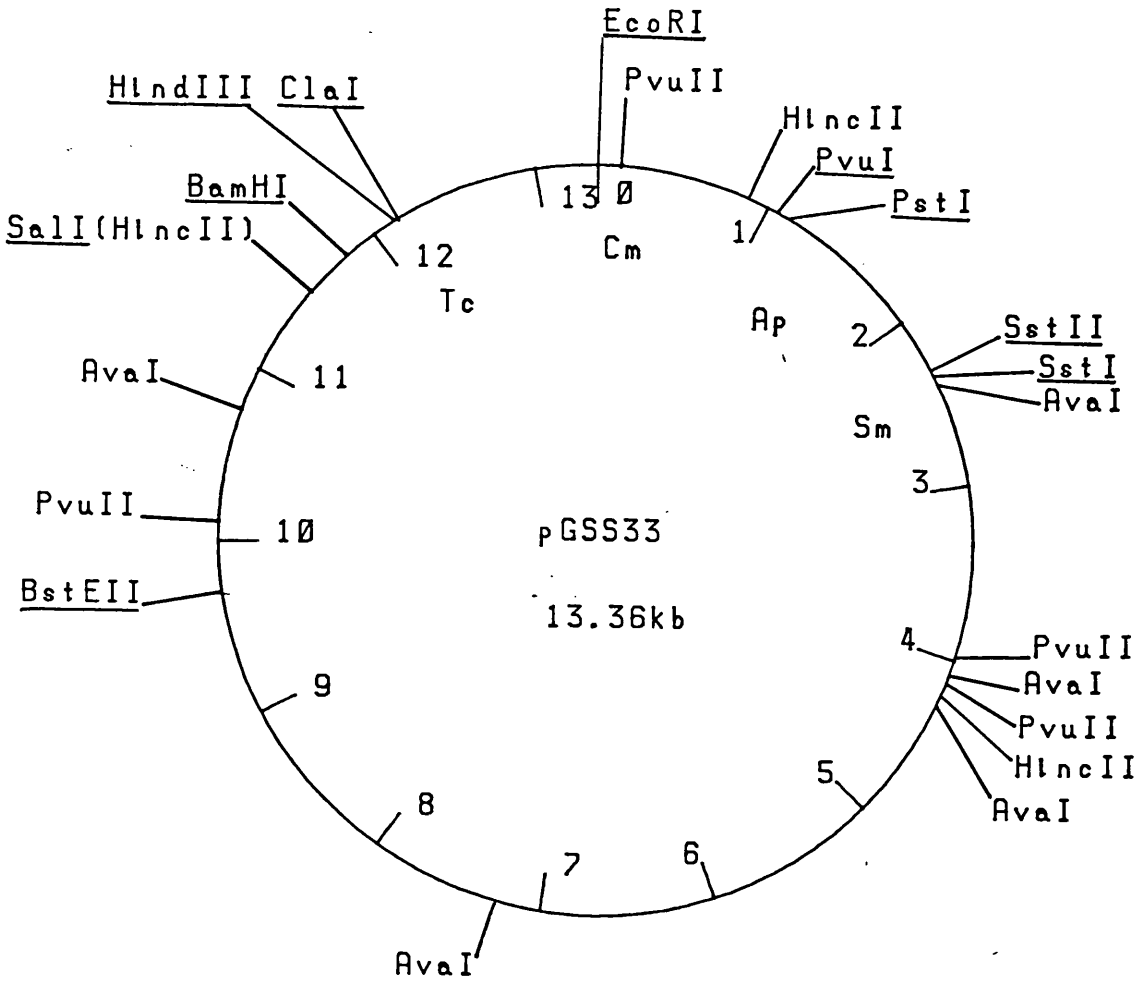



**Figure 29.**

A restriction map of pGSS33.

Each antibiotic resistance gene has at least one restriction enzyme site that will give insertional inactivation. All the unique sites are shown underlined.

that produced by *Bgl*II, which recognises $^{5'}$A↓GATCT$^{3'}$, or *Bcl*I which recognises $^{5'}$T↓GATCA$^{3'}$. Also two enzymes, *Sau*3A and *Mbo*I, recognise and cut the sequence $^{5'}$N↓GATCN$^{3'}$ and thus fragments generated with any of these enzymes could be cloned into pGSS33. Some further examples are listed in TableIV. With the majority of the cloning sites occurring within antibiotic resistance genes, pGSS33 becomes a very powerful and versatile cloning vector able to accept fragments generated by some twenty five different restriction enzymes. In most cases there is a ready selection for recombinants by easily distinguished phenotypic changes.

Table IV

| PLASMID CUT WITH RESTRICTION ENDONUCLEASE | WILL ACCEPT FRAGMENTS GENERATED BY RESTRICTION ENDONUCLEASE | INACTIVATING |
|---|---|---|
| BamHI | BamHI, BclI, BglII, MboI, Sau3A, XhoII | Tc |
| ClaI | ClaI, AcyI, AsuII, HinP$_1$I, HpaII, MspI, NarI, TaqI | (Tc)* |
| EcoRI | EcoRI, EcoRI* | Cm |
| HindIII | HindIII | (Tc)* |
| PstI | PstI | Ap |
| PvuI | PvuI, (RshI, XorII) | Ap |
| SalI | SalI, XhoI, AvaI | Tc |
| SstI | SstI, (SacI), HgiAI, HgiJII | Sm |
| SstII | SstII, (SacII) | Sm |
| BstEII | BstEII | - |
| | ( ) Isoschizomers | |

(Tc)* Does not always cause inactivation (Stüber and Bujard, 1981)

# CHAPTER 8

## 8.1 *In vitro* packaging and cosmid vectors

Originally developed for improving the efficiency with which λ vectors could transfect *E.coli*, *in vitro* packaging (Becker and Gold, 1975; Sternberg *et al.*,1977) has also become a powerful means of introduction into *E.coli* of "cosmids" containing large fragments of cloned DNA (Collins and Hohn,1978). Cosmids are bacterial plasmids which carry a small region of λ DNA in the proximity of the cleavage site (the *cos* site) which is all that is required for recognition by the packaging system (Syvanen,1974; Hohn,1975). [The extent of λ DNA sequence, flanking the cohered cohesive end site, which is required for packaging has recently been accurately defined (Hohn,1983). The region is asymmetrical and includes 36bp of the λ right end and 80bp of the λ left end.] Cosmids thus combine the properties of both phage and plasmid vector systems, with advantages over both. The capacity of the cosmid vectors to clone large DNA fragments, up to 40kb, is an advantage over the phage vectors. Because of the small size of some cosmids [e.g. pHC79, 6.4kb (Hohn and Collins,1980)] and the size-dependency of the packaging step – the only hybrids that are obtained are those able to fill a phage head with approximately 37-50kb of DNA – a selection for large hybrids is achieved automatically. This is in marked contrast to plasmid vectors where the transformation step introduces a bias against the cloning of large fragments (Sherratt,1979). However the major advantage of cosmids is the high efficiency with which input passenger DNA is converted into clones ($5\times10^4$ - $5\times10^5$ clones per μg of DNA), and makes possible the construction of gene-libraries from the entire genome of almost any organism.

Nevertheless, there are problems associated with using cosmids

and unless steps are taken to counteract them, the advantages referred to above will be drastically diminished. The size of the cosmid may be important, since without special treatment, dimers or trimers of just the cosmid vector could be packaged. This is prevented to a large extent by treatment with bacterial alkaline phosphatase which removes the 5' phosphate groups and prevents the cosmid molecules from joining to themselves. The recovery of plasmid molecules then requires a non-phosphatased insert to provide the 5' phosphates and allow ligation, at high DNA concentrations, of vector and insert to produce long, concatemeric DNA molecules which are the ideal substrates for *in vitro* packaging. At this stage though, small inserts could cause problems since an arrangement of

**cosmid - small insert - cosmid - small insert - cosmid**

could produce a packagable dimer. To minimize this problem, partial digests of passenger DNA, with an enzyme known to cut the DNA frequently, could be size-fractionated on a sucrose gradient or an agarose gel, thereby reducing the number of small fragments.

Broad host-range cosmids would allow gene banks to be produced from organisms other than *E.coli*, in a form which would make possible the re-introduction of cloned inserts into mutant hosts without the necessity for sub-cloning. Since there was some interest from collegues wishing to obtain a genetic map of *M. methylotrophus* and whose work would benefit from having a collection of cosmid clones carrying defined regions of the genome, attempts were made to produce broad host-range cosmids from some of the vectors already described.

### 8.2 Preparation of cosmid vectors from pGSS8

The plasmid pGSS8 was chosen for conversion to a cosmid since it

represented an *Inc*Q replicon containing a unique *Bam*HI site within the $Tc^r$ gene. This enabled *Sau*3A to be used to digest λ DNA with the potential of producing a relatively small λ fragment containing the *cos* site.

5μg of λ*cI857* DNA (the kind gift of Chris Boyd) were treated with T4 DNA ligase in a 200μl reaction, then phenol extracted and ethanol precipitated. This covalently closed circular DNA was then digested with 5 units of *Sau*3A in a 50μl reaction at 37°C for 60min, and at the same time 3μg of pGSS8 in a 50μl reaction were digested with 3 units of *Bam*HI at 37°C. 40μl of each digest were included in a 150μl ligation reaction and allowed to anneal at 30°C for 15min before ligase buffer and enzyme were added. After incubation at 10°C overnight, the mixture was used to transform *E.coli* strain NM36 to $Sm^r$. 346 colonies were isolated as being $Sm^rTc^s$, representing 37% of the total transformants.

Colonies containing plasmids with a cloned *cos* region were identified using an *in situ* colony hybridization method (Grunstein and Hogness,1975) with a probe prepared by end-labelling λ DNA with $\alpha^{32}$P-dGTP in the presence of *E.coli* DNA polymerase I. To check that the ends had indeed been labelled, 0.3μg samples of the DNA were mixed with 0.7μg of unlabelled λ DNA and digested with *Eco*RI, *Bam*HI or *Eco*RI/*Bam*HI. After agarose gel electrophoresis, the ethidium bromide stained bands were visualized under UV light and then the gel was autoradiographed to compare the band positions of the labelled DNA. [The right-hand end of λ is contained in the smallest *Eco*RI fragment, and the left-hand end in the largest *Eco*RI fragment, while from a *Bam*HI digest the right-hand end is in a 6.76kb fragment and the left-hand end in a 5.54kb fragment which co-migrates with an internal fragment of equal size. A double digest of λ DNA with *Eco*RI and *Bam*HI unambiguously separated the

ends onto a 5.54kb *Bam*HI fragment for the left end and a 3.38kb *Eco*RI fragment for the right end].

From the autoradiograph it was apparent that both ends had been successfully labelled, the right-hand end preferentially, due to the higher proportion of C-residues.

Thus 3μg of end-labelled λ DNA were digested with 3 units of *Sau*3A in a 100μl reaction at 37°C for 60min and the enzyme activity stopped by heating at 70°C for 10min prior to two phenol extractions. After being passed through a small G50 fine Sephadex column, equilibrated with TE buffer, in a siliconized pasteur pipette, labelled material was obtained in two 250μl fractions which were pooled, boiled for 2-3min and added to the hybridization mix.

The $Tc^S$ colonies were tooth-picked onto nitrocellulose filters, grown up overnight at 37°C, then denatured and baked prior to inclusion in the hybridization reaction. Dried, hybridized filters were autoradiographed for 36hr at -70°C with flash-sensitized film and an intensifying screen. Eight of the 346 colonies hybridized strongly to the probe and single colony lysates revealed that the plasmids in all these isolates were larger than pGSS8.

To confirm that these plasmids carried a *cos* region they were subjected to a further, more stringent hybridisation with two overlapping probes. 10μg of λ DNA were cut as before in a double digest with *Bam*HI and *Eco*RI, and after agarose gel electrophoresis, the bands carrying the left and right-hand ends were cut out of the gel and the DNA electroeluted. The λ ends, now physically separated, were each labelled by nick-translation using 10μCi of $\alpha^{32}$P-dCTP. Thus, to achieve hybridisation to both of these probes required that a *cos* site be present, since this would be the only region common to both probes.

All eight isolates hybridized to both probes suggesting that the

plasmids now carried λ *cos* sites. Plasmid DNA was prepared from cultures containing these new cosmids (designated Cos 1-8) and digested with *Eco*RI/*Sal*I to determine the size of the inserts by estimating the increase in size of the *Eco*RI/*Sal*I fragment from pGSS8 (650bp). A range of insert sizes, from approximately 3-14kb, was observed but the *Eco*RI/*Sal*I patterns were complicated by the introduction of extra *Eco*RI sites. *Bam*HI digests revealed insert sizes of approximately 3kb for Cos1 and Cos2, and approximately 6.5kb for Cos3. At this stage *in vitro* packaging extracts were prepared so that the cosmids could be tested.

**8.3 The packaging principle**

Cosmid and λ DNA, as well as *in vitro* recombinants thereof, are conveniently packaged using a combination of two lysates, each of which is defective in a different step of λ morphogenesis. Empty precursor particles accumulate after induction in bacteria containing a prophage mutant in gene *A*. This mutation can be complemented by addition of the missing *A* protein by supply from an induced λ*E*⁻ lysogen. The *E* protein is the major component of the bacteriophage head and is required for assembly of the earliest identifiable precursor. Mutants in protein *E* accumulate all other components of the the viral capsid. The *A* protein is involved in the insertion of λ DNA into the bacteriophage head, and cleavage of the concatemeric, precursor DNA at the *cos* sites. Mutants in the *A* gene accumulate empty pre-heads. When lysates prepared from induced cultures of these two mutant lysogens are mixed in the presence of ATP, spermidine and putrescine, DNA is packaged, the head matures and tails are attached. The result is a DNase-resistant, infectious particle which can be stored like any *in vivo* produced phage.

The complete genotypes of the strains are :

NS428: *recA* ($\lambda A_{am}11$ *b2 red3* $cI_{ts}857$ $S_{am}7$)

NS433: *recA* ($\lambda E_{am}4$ *b2 red3* $cI_{ts}857$ $S_{am}7$)

The $cI_{ts}857$ repressor allows efficient prophage induction by temperature shift, and the mutation in gene *S* prevents lysis of the induced bacteria. The combination of the bacterial *recA* and phage *red3* mutations and the *b2* deletion, drastically decreases the amount of endogenous packagable DNA.

**8.4 Preparation of extracts for *in vitro* packaging.**

Overnight cultures of the two packaging strains, NS428 and NS433, grown at 32°C were diluted 1:100, NS428 into 150ml of L-broth containing 10mM magnesium chloride in a 500ml flask, and NS433 into 750mlof the same medium in a 2L flask. The cells were shaken with good aeration at 32°C until they reached 1-2.$10^8$ cells/ml (0.2 absorbance at 630nm). This took only 2-2.5hr. At this point, the lysogens were induced by swirling them in a hot (80-90°C) water bath until the temperature of each culture was 42°C. (This was measured by placing an alcohol-sterilized thermometer directly into the culture medium). The flasks were then transferred to a 42°C shaking incubator and aerated vigorously for 20min.

The cultures were then shifted to 38°C for an additional 70min, and again shaken as hard as possible. After this time the cells were quick-chilled in iced water. The effectiveness of induction was tested by transferring a few millilitres of culture to a glass tube and shaking with a few drops of chloroform. The cultures cleared completely in a few minutes, (chloroform did not lyse a non-induced culture). The induced cultures were then ready for processing and were used as rapidly as possible.

**8.5 Preparation of extract A**

150ml of each of the two chilled, induced cultures were mixed and

cenrifuged at 8000rpm for 10min at 4°C in a Sorval GS-A rotor. The pellet was resuspended in a total of 0.6ml of buffer A (20mM Tris.HCl pH8.0; 1mM EDTA; 3mM $MgCl_2$; 5mM 2-mercaptoethanol) transferred to a sterile plastic universal, clamped securely in ice, and sonicated for 15 X 2sec bursts, allowing 30sec between bursts (Micro tip of MSE sonicator, amplitude 6$\mu$, mark 4). The tip of the sonicator probe was sterilized with alcohol prior to use. The culture cleared as the sonication proceeded. The correct level of sonication was checked by spinning 50$\mu$l aliquots of sonicated and non-sonicated cells for 5min in an Eppendorf centrifuge. There was hardly any pellet in the sonicated sample, compared to the non-sonicated sample.

150$\mu$l of cold glycerol was then added to the sonicate and mixed gently with a pipette tip. The mixture was kept on ice. 50$\mu$l aliquots of this mixture were then placed in cold Nunc vials and quickly submerged into liquid nitrogen. The extracts could be stored indefinitely under liquid nitrogen for use in future experiments. Two standard reactions could be performed with one tube of Extract A.

**8.6 Preparation of Extract B**

The remaining, chilled, induced cells of NS433 (600ml) were harvested by centrifugation at 8000rpm for 10min at 4°C. All the supernatant liquid was poured off from the pellet and the last traces were removed with a sterile cotton wool swab. The pellets were resuspended in a total of 1.2ml of Tris-sucrose buffer (10% sucrose in 50mM Tris.HCl pH7.4) and 80$\mu$l aliquots were distributed into Nunc vials and immediately frozen in liquid nitrogen. (All the tubes were to be processed further, but it was convenient to work with only half the quantity at a time, the other half remaining in liquid nitrogen.)

Half the tubes were thawed in a beaker of water at room temperature. This was followed by quick freezing in liquid nitrogen and rapidly thawing them again in room temperature water. To each

tube was added 4μl of lysozyme solution (1mg.ml$^{-1}$ in 0.25M Tris.HCl pH7.4) which was stirred into the extremely viscous lysed cell suspension with the end of the Gilson tip. The lysates were incubated on ice for 30min. In a separate tube and on ice, 625μl of glycerol were mixed with 200μl of buffer B (50mM spermidine; 100mM putrescine; 10mM Tris.HCl pH7.4; 15mM ATP; 16mM $MgCl_2$; 30mM 2-mercaptoethanol). 33μl of this solution was added to each of the lysates, and stirred in with the end of the pipette tip. These lysates were then stored in liquid nitrogen while the rest of the tubes were processed in the same way. One standard packaging reaction could be done with one tube of extract B.

**8.7 The Packaging Reaction**

The following ingredients were mixed at room temperature, in a polypropylene tube in the order shown;

30μl Buffer A

2μl Buffer B

20μl Extract A

5μl DNA (up to 1μg per reaction)

and incubated at room temperature for 15min, then 100μl of Extract B were added and stirred thoroughly. This mixture was incubated at 37°C for 60min after which time 150μl of DNase buffer (100mM NaCl; 10mM $MgSO_4$; 50mM Tris.HCl pH7.5; 0.01% gelatin; 10μg.ml$^{-1}$ DNase I) were added. After a further 15min at 37°C, with an occasional stir, 2-3 drops of chloroform were added and the lysate cleared by spinning for 5min in an Eppendorf centrifuge. The mixture was then treated as a conventional phage lysate, typically containing 2-10$\times 10^7$ phage per μg of DNA when wild-type λ DNA was used.

For packaging phage DNA, 60μl of the lysate were diluted to 200μl with phage buffer (3g $KH_2PO_4$; 7g $Na_2HPO_4$; 5g NaCl; 10ml of 100mM

$MgSO_4$; 10ml of 10mM $CaCl_2$; 1ml of 1% gelatin solution, per litre), then further dilutions of $10^{-2}$, $10^{-3}$ and $10^{-4}$ were made, also in phage buffer. 100$\mu$l aliquots of each dilution were mixed with 100$\mu$l of plating cells and incubated at room temperature for 10min before being top-layered in 2.5ml of L-top agar, supplemented with 10mM $MgCl_2$, onto fresh L-agar plates.

For cosmid concatemers or cosmids containing inserts, dilutions of the lysates were made in 10mM $MgCl_2$, 30$\mu$l diluted to 100$\mu$l, before being mixed with 100$\mu$l of plating cells. After incubation for 10 min at room temperature,1ml of L-broth was added and incubation continued for up to 60min at 37°C. Dilutions were made to $10^{-2}$ and $10^{-3}$ in L-broth and 100$\mu$l aliquots of each dilution were spread onto L-agar plates containing the appropriate antibiotic for selection. In both cases the plates were incubated at 37°C.

To test if the cosmids, Cos1-8, were capable of being packaged *in vitro* the first approach was to prepare concatemers of just the cosmid molecules themselves, as the packaging substrates. *Eco*RI digests were made and the linear molecules treated with T4 DNA Ligase to form multimeric units. However, agarose gel electrophoretic analysis of these reaction mixtures revealed that although ligation was occurring to a significant extent, some of the cosmid molecules had more than one *Eco*RI site and small fragments, which could contain the *cos* regions, were being preferentially converted into multimers without ligation to the large fragments. This situation suggested that no suitable substrates were being formed, and in fact no colonies were obtained from a packaging of these molecules. (In retrospect an *Sst*I digest and re-ligation for multimers may have been better since there was more chance of this site being unique).

An alternative approach to testing for packaging was to actually

attempt to produce a gene bank. Thus *M.Methylotrophus* DNA was partially digested with *Sal*I, ligated to *Sal*I cut cosmids, and the products put into an *in vitro* packaging reaction. Again, no resistant colonies were obtained that could be confirmed as containing any plasmid. At this time other cosmid molecules became available, one called Homer (P.Rigby, personal communication) and a commercial source (Boehringer Corporation Limited) of pHC79 (Hohn and Collins,1980). I was unable to obtain any evidence of packaging using Homer in either of the two tests just described above, and decided to stop any further work at this stage.

Subsequently, Jenny Broome-Smith constructed a broad host-range cosmid by introduction of a *Bgl*II fragment, containing the *cos* site from Homer, into *Bam*HI digested pGSS15 (*Bgl*II and *Bam*HI produce the same sticky ends, but when these are ligated to each other, the resulting hexanucleotide sequence is not recognised by either enzyme). This cosmid, pJBS2 (13.1kb), exemplified some of the practical limitations referred to earlier in this chapter. Its size would in theory allow 24-37kb inserts to be cloned, but a dimer or trimer could also be packaged. Also there were now fewer restriction sites that were suitable for cloning, and the *Pst*I in the $Ap^r$ gene could not be used as this would destroy the only selectable marker.

David Burt, in continuing the cosmid work did produce packaging substrates from *M.methylotrophus* DNA and pJBS2, but was unable to recover any clones from packaging experiments using the Sternberg and Enquist system, described above. He was however successful using another method (Scalenghe *et al.*,1981) and obtained approximately $10^3$-$10^4$ cosmid clones per microgram of *M.methylotrophus* DNA.

## CHAPTER 9

DNA-dependant gene expression

Once a cloning procedure is complete and a gene is carried on a suitable vector, it is often important to be able to identify any polypeptide products. For *E.coli*, there are four major gene expression systems available.

a) "UV-irradiated host" system.

*E.coli* is heavily UV-irradiated to reduce host protein synthesis and then infected with phage λ carrying the cloned gene(s). Phage-encoded proteins are labelled by the inclusion of $^{35}S$- methionine and analysed by SDS-PAGE.

b) "Maxicell" system.

*E.coli recA, uvrA* cells (ensuring extreme UV sensitivity) are transformed with a multicopy plasmid containing the cloned genes, then UV-irradiated. After 16hr incubation the chromosome is largely destroyed, but some plasmid molecules will still remain intact. Subsequent incubation with $^{35}S$-methionine specifically labels the plasmid encoded proteins, which can then be analysed as before.

c) "Minicell" system.

*E.coli minA, minB* cells (which produce large numbers of anucleate minicells) are transformed with a recombinant multicopy plasmid. This results in the formation of minicells into which plasmid DNA has segregated, but not chromosomal DNA. Labelling of the purified minicells then provides an excellent system for the specific identification of plasmid-coded proteins.

d) *In vitro* transcription/translation system.

The *in vitro* synthesis of proteins programmed by either phage or plasmid DNA in a coupled transcription/translation system (DeVries and Zubay,1969; Zubay,1973) constitutes perhaps the most versatile method. Manipulation of the DNA *in vitro* before its use as a template, can pin-point relatively small regions of DNA coding for a specific polypeptide.

For organisms other than *E.coli*, however, there have been few such systems reported. For *M.methylotrophus* in particular, there are no systems by which cloned gene products can readily be analysed. Indeed, at present, only the coupled *in vitro* transcription/ translation system looks a possible candidate, as there are no phages with which to operate a UV-irradiated system, and no suitably mutant strains to provide maxicell or minicell systems. Because *M.methylotrophus* cannot be transformed, an *in vitro* system for analysis of cloned gene products could provide information on the expression of a cloned gene relatively quickly.

The cell-free system devised by Zubay involves the preparation of a crude cell extract from *E.coli* which contains all the enzymes and factors necessary for transcription and translation. The extract however, needs to be supplemented with nucleotide triphosphates, amino acids, magnesium, an energy regenerating system and certain cofactors. I decided to prepare an *E.coli* cell-free system, then to go on and and prepare a similar system from *M.methylotrophus*, once I had established the necessary techniques.

9.1 Preparation of the components of an *E.coli* transcription/ translation system (S-30 extract). (Collins,1979; Pratt *et al.*,1981)

The procedure at first seems to be extremely complicated, but once the various solutions are made an extract can be prepared in a

day in sufficient quantity, for most requirements, that will remain active for at least a year with storage in liquid nitrogen. The method is given in precise detail such that, with care and attention to detail, an active extract will be produced every time.

**i) Chemicals.**

| Component | Manufacturer | Maximum Requirement |
|---|---|---|
| $KH_2PO_4$ | BDH AR | 200g |
| $K_2HPO_4$ | BDH AR | 1kg |
| Yeast extract | Difco | 100g |
| Thiamine HCl | Sigma | 100mg |
| Glucose | BDH AR | 500g |
| Tris | Sigma | 500g |
| Magnesium acetate | Fisons AR | 500g |
| Ammonium acetate | BDH AR | 5g |
| Potassium acetate | Fisons SLR | 5g |
| Calcium acetate | Fisons | 5g |
| Polyethylene glycol 6000 | Fisons | 5g |
| Dithiothreitol | Sigma | 3-5g |
| ATP (Sodium salt) | Sigma | 2g |
| Phosphoenolpyruvate ($Na_3$) | Sigma | 1-2g |
| Cyclic AMP | Sigma | 100mg |
| Folinic acid | Sigma | 100mg |
| CTP (Sodium salt) from yeast | Sigma | 100mg |
| GTP (Sodium salt) | Sigma | 100mg |
| UTP (Sodium salt) from yeast | Sigma | 100mg |
| Diethyl pyrocarbonate | Sigma | 25ml |
| Pyruvate kinase (rabbit muscle) | Sigma | 0.5ml |
| tRNA (*E.coli*) | Sigma (R-4251) | 100mg |
| $^{35}S$ methionine | Amersham | 1mCi |

**ii) Treatment of water and apparatus.**

Before any of the reagents can be made up some diethylpyrocarbonate (DEPC)-treated water and apparatus must be prepared.

a) Treatment of water to prepare all reagents and S30 buffer

1ml of DEPC was slowly added with stirring to each litre of water. This was allowed to stir for up to one hour, then autoclaved and stored at 4°C. For each preparation, 3 × 4 l and 8 × 500ml were usually required i.e. a total of 16 litres.

b) Treatment of apparatus

All glassware and plastic tubing was treated with DEPC by filling (or soaking) the apparatus with water to which DEPC was then added. After about an hour, the water was drained off and the apparatus autoclaved. An average requirement for equipment is listed below.

6 × one litre flasks with screw tops.

16 × SS34 tubes

6 × 25ml beakers

4 × 100ml or 50ml measuring cylinders.

6-10 × McCartney bottles.

3-4 × stirring bars.

20-30 × pasteur pipettes.

3 × 100ml Buchner flasks and one small funnel.

1 × bung for above

Several spatulas.

**iii) To pH solutions.**

Solutions which were too alkaline were brought to the correct pH with acetic acid and those which were too acidic, with 2M Tris (2.42g/10ml DEPC water). In some cases $^{1}/_{10}$th acetic acid (1ml glacial acetic acid + 9ml DEPC water) was used (e.g. to pH the

PEP). Both these solutions were autoclaved prior to use.

iv) Preparation of S-30 buffer.

This buffer was required to wash the cells and to prepare the S-30 extract. Stock solutions were prepared as follows:

A. 500ml 1M Tris acetate pH8.2

B. 500ml 6M potassium acetate

C. 500ml 1.4M magnesium acetate

D. 100ml 100mM DTT

all in DEPC-treated water.

100ml each of solutions A, B and C were mixed and made up to 1 litre, (in a pre-treated flask) with DEPC-treated water. This was a ×10 S-30 buffer solution which was stored at 4°C and diluted just prior to use. 100ml of ×10 S-30 buffer and 10ml of solution D (DTT) were made up to 1 litre with DEPC-water. This produced a buffer with the composition:

10mM Tris acetate pH8.2

14mM magnesium acetate

60mM potassium acetate

1mM DTT

v)Preparation of solutions

The following solutions were prepared in DEPC-treated water.

| Solution | weight volume | aliquots |
|---|---|---|
| 2.2M Tris acetate pH8.2 | 26.65g/100ml | |
| 3M magnesium acetate | 64.20g/100ml | |
| 0.55M DTT | 424mg/5ml | 130µl, 25µl |
| 38mM ATP (pH7.0 with Tris) | 419mg/20ml | 2.7ml, 220µl |

| Solution | | weight | volume | aliquots |
|---|---|---|---|---|
| 88mM CTP<br>GTP<br>UTP | (PH7.0<br>with<br>Tris) | 42mg<br>46mg<br>42mg | /1ml | 70μl |
| 0.42M PEP | | 983mg/10ml | | 1.6ml, 420μl |
| 40% polyethylene glycol | | 4g/10ml | | |
| Folinic acid ($Ca^{++}$ leucovorin) | | 2.7mg/1ml | | 100μl |
| 50mM cAMP pH7.0 | | 16mg/1ml | | 100μl |
| transfer RNA | | 17.4mg/1ml | | 70μl |
| 0.1M magnesium acetate | | 0.21g/10ml | | |
| Methionine | | 44mg/ml | | |

Inorganic mix

| | | |
|---|---|---|
| 1.4M ammonium acetate | 1.08g | /10ml |
| 2.8M potassium acetate | 2.74g | |
| 0.38 calcium acetate | 0.60g | |

Amino acid mixtures:- each mixture of amino acids was added to 5ml of DEPC-treated water and the suspensions frozen at -20°C in 50μl aliquots.

| | A wt(mg) | B |
|---|---|---|
| Amino acid | 50mM | 55mM |
| Alanine | 22.3 | 24.53 |
| Arginine, HCl | 52.7 | 58.0 |
| Asparagine | 33.0 | 36.3 |
| Aspartic acid | 33.3 | 36.6 |
| Cysteine | 30.3 | 33.33 |

| | A | wt(mg) | B |
|---|---|---|---|
| Amino acid | 50mM | | 55mM |
| Glutamine | 36.5 | | 40.15 |
| Glycine | 18.8 | | 20.68 |
| Histidine, HCl | 47.9 | | 52.7 |
| Isoleucine | 32.8 | | 36.1 |
| Leucine | 32.8 | | 36.1 |
| Lysine | 45.7 | | 50.3 |
| Methionine | 37.3 | | - |
| Phenylalanine | 31.3 | | 34.4 |
| Proline | 28.8 | | 31.7 |
| Serine | 26.3 | | 29.0 |
| Threonine | 29.8 | | 32.8 |
| Tryptophan | 51.1 | | 56.2 |
| Tyrosine | 4503 | | 49.8 |
| Valine | 29.3 | | 32.2 |

**vi) Treatment of dialysis tubing.**

Dialysis tubing (8 pieces of 1" wide, 12" long) was boiled in 0.1M sodium carbonate; 0.01M EDTA solution, rinsed extensively in distilled water and then in sterile DEPC water at 4°C.

vii) Growth of cells for preparation of S-30 extract.

An overnight culture of *E.coli* strain MRE600 was used to inoculate, at 1%, eleven flasks each containing 1 litre of the following medium:

Each litre contained:

| | |
|---|---|
| 5.6g | $KH_2PO_4$ |
| 28.9g | $K_2HPO_4$ |
| 1g | yeast extract |
| 40ml | 25% glucose |
| 10ml | 100mM magnesium acetate |

supplemented with thiamine and the required amino acids (see materials and methods section). The cultures were grown with shaking at 37°C for 3 hours to an $OD_{650}$ of 1.35. The cells were chilled on ice and harvested in a Sorvall SS34 rotor, fitted with a continuous flow apparatus, at 15krpm.

43.7g wet weight of cells were obtained and these were washed twice in 345ml of S30 buffer containing 0.5ml β-mercaptoethanol per litre. A third wash in 150ml of this buffer allowed the cells to be pelleted in 4x50ml SS34 polypropylene tubes. The pellets were then stored frozen at -70°C overnight.

The frozen cells were allowed to thaw slowly (1-2hr) on ice before being homogenized slowly in 437ml of S30 buffer + 0.5ml β-mercaptoethanol per litre (i.e. 100ml/10g cells). The cell pellet obtained after centrifuging this suspension for 30min at 10krpm weighed 32.2g and was resuspended anaerobically (under suction in a 100ml Buchner flask) in 40ml of S30 buffer (63.5ml/50g of cells). The cells were then passed through a French press (pre-washed with DEPC water, wrapped in foil and chilled to 4°C) at ~9000psi. Immediately 100μl of 0.1M DTT were added for every 10ml of pressed cells collected and cell debris was centrifuged down at 15.5krpm for 30min.

At this point the Preincubation Mix was prepared.

Preincubation mix.

| | |
|---|---|
| Pyruvate kinase [26mg/ml in $(NH_4)_2SO_4$] | 25$\mu$l |
| Tris acetate pH 8.2 (2.2M) | 1.0ml |
| Magnesium acetate (3M) | 23$\mu$l |
| ATP 38mM pH 7.0 | 2.63ml |
| PEP 0.42M pH 7.0 | 1.5ml |
| DTT 0.55M | 60$\mu$l |
| 20 amino acid mix (50mM) | 6$\mu$l |
| Water | to 7.5ml |

The upper 4/5ths of supernatant fluid in each tube was aspirated off into a small Erlenmeyer flask protected from the light. The centrifugation was repeated and again the upper 4/5ths of fluid removed to the flask. 7.5ml of preincubation mix were added for every 25ml of supernatant and incubated at 37°C for 80min. The mixture was then transferred to dialysis tubing and dialysed 4 × 45min against 50 volumes of S30 buffer. On completion of the dialysis the "S30 extract" was distributed into 0.5ml portions in 2ml Nunc vials and stored in liquid nitrogen. At this stage the S30 extract was ready and the components for the *in vitro* transcription/translation could be mixed from the stock solutions described in (v) above. The low molecular weight mix LMM, can be frozen and kept for a few weeks.

Low Molecular Weight Mix

| Stock solution | Volume | End conc$^{n}$. |
|---|---|---|
| 1. Tris acetate 2.2M pH 8.2 | 40$\mu$l | 77mM |
| 2. DTT 0.55M | 5$\mu$l | 2.4mM |
| 3. ATP 38mM pH7.0 | 50$\mu$l | 1.66mM |
| 4. CTP GTP each 88mM pH 7.0 UTP | 15$\mu$l | 1mM |
| 5. PEP 0.42M pH 7.0 | 100$\mu$l | 37mM |
| 6. 19-amino acid mix (55mM each) | 10$\mu$l | 480$\mu$M |
| 7. Polyethylene Glycol-6000 40% in water. | 75$\mu$l | 2.6% |
| 8. Folinic acid 2.7mg/ml | 20$\mu$l | 4.7$\mu$g/ml |
| 9. cAMP 50mM pH 7.0 | 20$\mu$l | 88$\mu$M |
| 10. tRNA *E.coli* 17.4mg/ml | 15$\mu$l | 190$\mu$g/ml |
| 11. Inorganic mix | 40$\mu$l | $NH_4^+$ 50mM |
| | | $K^+$ 100mM |
| | | $Ca^{++}$ 13mM |

The components were mixed in the order given.

**9.2 Optimisation of the Magnesium Concentration.**

Different preparations of S30 extracts appear to require different magnesium levels for optimal activity. The level is easily obtained by setting up incubations with varying amounts of 0.1M magnesium acetate solution. It may also be necessary to optimise the S30 concentration as well.

The optimum magnesium concentration for the *E.coli* extract just described was determined as follows:-

Requirements

DNA (R300B, 180μg/ml)

$^{35}S$ Methionine 7μCi/μl

Water or Tris acetate

LMM

0.1M Magnesium acetate

S30 extract (One Nunc vial removed from liquid nitrogen and allowed to thaw on ice, 1-2hr. The contents were transferred to an eppendorf tube and spun for 3min. The supernatant was aliquoted in 60-100μl lots and any not used was frozen at -20°C to be thawed once more).

The experiment could be planned out thus:-

| Tube | DNA | vol | $^{35}S$ | LMM | $Mg^{++}$ | TrisAc | S30 |
|---|---|---|---|---|---|---|---|
| 1 | R300B | 7.5μl | 2μl | 7.5μl | 1μl | 7μl | 5μl |
| 2 | " | 7.5 | 2 | 7.5 | 2 | 6 | 5 |
| 3 | " | 7.5 | 2 | 7.5 | 3 | 5 | 5 |
| 4 | " | 7.5 | 2 | 7.5 | 4 | 4 | 5 |
| 5 | " | 7.5 | 2 | 7.5 | 5 | 3 | 5 |
| 6 | " | 7.5 | 2 | 7.5 | 6 | 2 | 5 |
| 7 | " | 7.5 | 2 | 7.5 | 7 | 1 | 5 |
| 8 | " | 7.5 | 2 | 7.5 | 2 | 6 | 8 |
| 9 | " | 7.5 | 2 | 7.5 | 4 | 4 | 8 |
| 10 | " | 7.5 | 2 | 7.5 | 6 | 2 | 8 |
| 11 | - | - | 2 | 7.5 | 4 | 11.5 | 5 |
| 12 | - | - | 2 | 7.5 | 4 | 8.5 | 8 |

The DNA, $^{35}S$ methionine, LMM, Tris acetate and magnesium acetate were mixed in sterile eppendorf tubes and incubated at 37°C for 2-5min. The S30 extract was then added to each tube and incubated with shaking at 37°C for 30-60min. After this period 5μl of 44mg/ml methionine (prewarmed) was added and incubation carried out for a further 5min. 2μl samples were then removed, spotted on to strips of 3MM paper, dried, washed with cold 10% TCA + methionine, with 5% TCA + methionine at 90°C for 10min, then three times with ice-cold 5% TCA. The filters were then washed twice with acetone and dried again before counting in Fisofluor 3 (Fisons). From a plot of counts versus volume of $Mg^{++}$ added the optimum value could be readily observed as that giving the highest level of incorporation of the label, i.e. 6μl of 0.1M magnesium acetate [Fig.30].

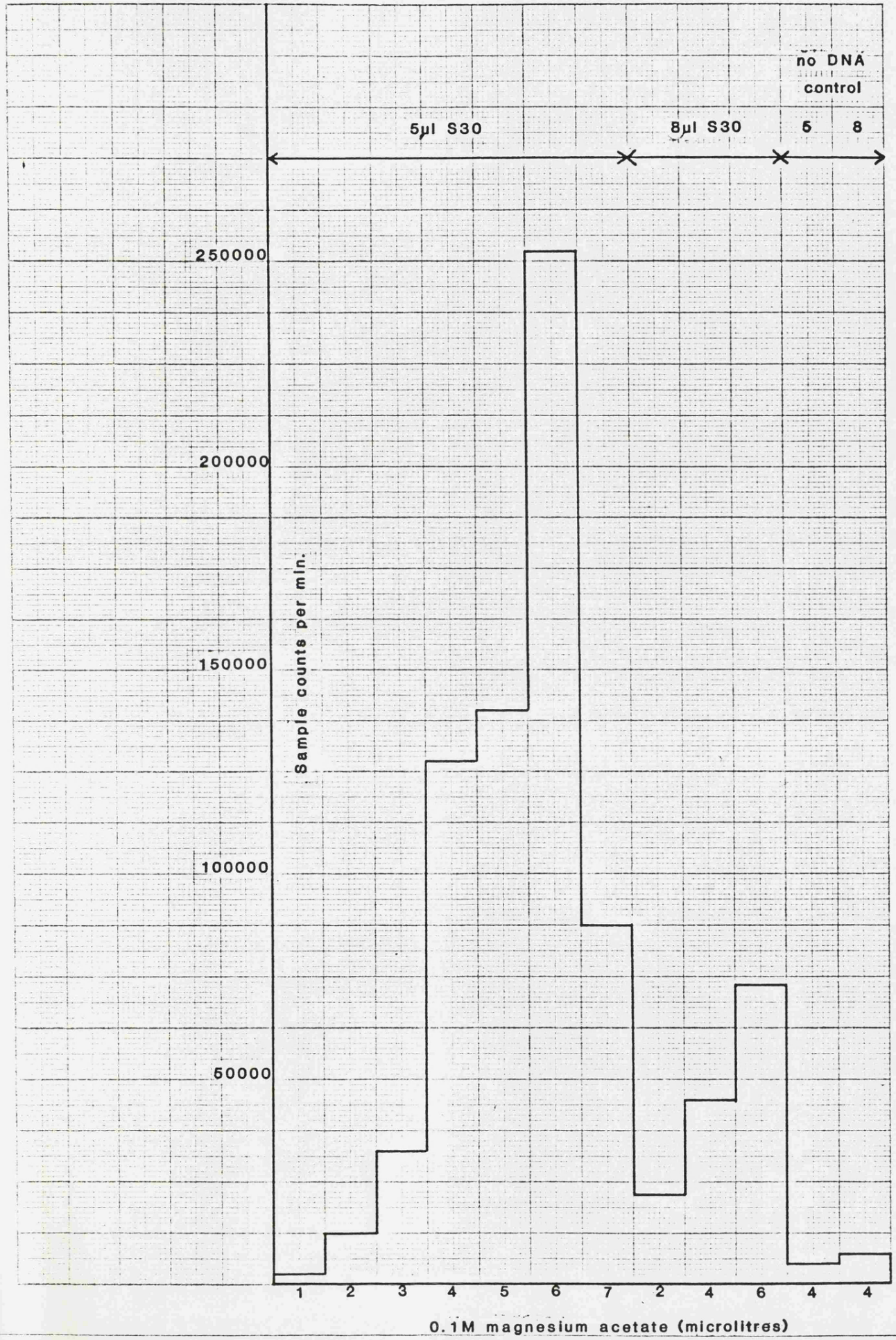
## 9.3 Use of *E.coli* S-30 extracts to investigate protein products of various plasmids.

DNAs prepared from several of the broad host range vectors described earlier were included in reaction mixtures made up as follows:-

| N$^{O}$. | DNA | VOL. | $^{35}S$-met | LMM | $Mg^{++}$ | $H_2O$ | S-30 |
|---|---|---|---|---|---|---|---|
| 1 | pBR322 | 5μl | 2μl | 7.5μl | 6μl | 4.5μl | 5μl |
| 2 | R300B | 7.5 | 2 | 7.5 | 6 | 2.0 | 5 |
| 3 | pGSS15 | 5 | 2 | 7.5 | 6 | 4.5 | 5 |
| 4 | pGSS8 | 5 | 2 | 7.5 | 6 | 4.5 | 5 |
| 5 | pGSS7 | 5 | 2 | 7.5 | 6 | 4.5 | 5 |
| 6 | pGSS11 | 5 | 2 | 7.5 | 6. | 4.5 | 5 |
| 7 | pDHFR7 | 5 | 2 | 7.5 | 6 | 4.5 | 5 |
| 8 | pDHFR2.43 | 5 | 2 | 7.5 | 6 | 4.5 | 5 |
| 9 | Control | 0 | 2 | 7.5 | 6 | 9.5 | 5 |

**Figure 30.**

Result of magnesium optimisation experiment for *E.coli* S-30 extracts. Clearly 6$\mu$l of 0.1M magnesium acetate was producing the highest incorporation of label. The chart also shows that using more S-30 extract was less efficient.


no DNA
control
5µl S30
8µl S30
5
8
250000
200000
150000
100000
50000
Sample counts per min.
1
2
3
4
5
6
7
2
4
6
4
4
0.1M magnesium acetate (microlitres)

The mixtures were incubated at 37°C for 5min before the S-30 extract was added, then the incubation continued for 1hr with shaking. 2.5μl of prewarmed methionine (44mg/ml) were added to each tube and incubated for a further 5min at 37°C. The tubes were then transferred to ice and 2μl samples removed for TCA precipitation and counting as described in the magnesium optimization experiment. To the remaining mixture in each tube were added 60μl of gel buffer and, after heating in a boiling water bath for 4min, 15μl samples were loaded onto a gel for analysis by SDS-PAGE and autoradiography [Fig.31].

**9.4 Preparation of a *M.methylotrophus* S-30 extract.**

After establishing that an active *E.coli* S-30 extract could be prepared, an identical procedure was undertaken to produce a similar extract from *M.methylotrophus*. It was not possible to freeze *M.methylotrophus* cells and thaw them again without almost complete lysis, therefore the growth of the cells was adjusted so that this step was not necessary. 11× 2litres of medium were inoculated at 0.1% and grown for 13hr to an $OD_{650}$ of 0.75. This yielded 24.2g wet weight of cells. The centrifugation step after French pressing allowed only about 10ml of supernatant to be collected with a further 5ml from a second spin. The pellet was very viscous and may need to be spun harder and for longer, in future, to recover more supernatant.

Unfortunately, the extract prepared from *M.methylotrophus* did not appear to be active when provided with a template of supercoiled R300B DNA. *E.coli* extracts with the same DNA showed good incorporation of the label. R300B could replicate in *M.methylotrophus* and it was known that DNA prepared from a *dam*$^+$ *E.coli* strain would not be restricted in *M.methylotrophus* (Boyd,1983), indicating that

host restriction was not a problem. To rule out the possibility that there was a constituent of the *M.methylotrophus* extract that was toxic to the system, an experiment was set up using mixtures of *E.coli* and *M.methylotrophus* extracts.

**9.5 Testing *M.methylotrophus* extract.**

To test *M.methylotrophus* S-30 extract for inhibitory activity.

The following reaction mixtures were set up.

| R300B DNA | $^{35}$S-met | LMM | $Mg^{++}$ | $H_2O$ | S-30 | |
|---|---|---|---|---|---|---|
| | | | | | *M.met.* | *E.coli* |
| 7.5µl | 2µl | 7.5µl | 3µl | 5µl | 0µl | 5µl |
| 7.5 | 2 | 7.5 | 3 | 5 | 5 | 0 |
| 7.5 | 2 | 7.5 | 3 | 5 | 2.5 | 2.5 |
| 7.5 | 2 | 7.5 | 3 | 5 | 3.5 | 1.5 |
| 7.5 | 2 | 7.5 | 3 | 5 | 1.5 | 3.5 |
| 0 | 2 | 7.5 | 3 | 12.5 | 1.5 | 3.5 |
| 0 | 2 | 7.5 | 3 | 12.5 | 5 | 0 |

There was no apparent inhibition of the *E.coli* extract and the suggestion from this was that the *M.methylotrophus* system was sub-optimal. In view of the number of components that could be varied, I decided that optimization of the system was beyond the scope of this thesis and terminatedthe work at this stage.

# CHAPTER 10

## 10.1 Measurement of plasmid copy number

All plasmids appear to exist within fairly well-defined numbers of copies in their respective host cells. For many, there is genetic evidence to suggest that the plasmids specify their own copy number (Nordstrom *et al.*,1972; Gelfand *et al.*,1978). In the case of the plasmid R1, the copy control mechanisms are now understood sufficiently well to allow the genes responsible to be manipulated. Derivatives of R1 which can be temperature-switched into an uncontrolled copy number mode have been constructed (Molin *et al.*, 1981), providing a facility to produce large amounts of a cloned gene product. This approach, in going from low to high copy number, may be the only means of propagating a gene which produces a toxic product. Broad host range plasmids introduce another aspect of plasmid copy control, the ability to define their copy number in different host species. It is therefore important to be able to determine, as accurately as possible, the intracellular plasmid copy number, for analysis of the factors that influence this number, or perhaps to understand how to alter the copy number artificially.

Several approaches to the determination of plasmid copy number have been reported and these can be conveniently divided into four main classes.

1. Centrifugation of:

    a) total cell lysates (Womble *et al.*,1977)

    b) cleared lysates (Clewell and Helinski,1969) in caesium chloride gradients containing intercalating dye.

    c) alkaline sucrose gradients (Uhlin and Nordstrom, 1978).

2. Electrophoresis of:

a) whole cell lysates with analysis of the bands by fluorescence densitometry (Projan *et al.*,1983)

b) *in vivo* radiolabelled cells lysed directly in the gel with analysis by cutting out the bands, digesting the agarose and counting (Hochmannova and Nesvera,1982)

c) alkaline SDS DNA preparations from known numbers of cells, staining with ethidium bromide, photographing and microdensitometer scanning of the negatives (Stueber and Bujard,1982).

3. Gene dosage-dependent production of antibiotic-inactivating enzymes has also been used (Uhlin and Nordstrom,1977), but this is only applicable to plasmids carrying the relevant genes - β-lactamase, chloramphenicol acetlytransferase or streptomycin adenylylase.

4. Hybridization analysis of total DNA from plasmid-containing cells:

a) filter saturation hybridization (Warnaar and Cohen,1966)

b) reassociation kinetic determination of plasmid DNA concentration in solution (Austin *et al.*,1978)

c) saturation hybridization in solution (Gelfand *et al.*,1978)

The wide range of techniques that have been applied to the measurement of plasmid copy number gives some indication of the importance attached to this figure, and the difficulty in obtaining accurate estimates! Also, the different approaches are evidence of the fact that plasmid copy numbers are not easy to measure accurately, and there are advantages and disadvantages for all the methods listed above.

The centrifugation methods offer minimal estimates of plasmid

copy number, requiring that all the plasmid DNA be present as supercoils (1a, above), or that a quantifiable proportion of plasmid be recoverable after high speed centrifugation (1b). Nicking of supercoils by shearing, the action of non-specific endonucleases or disruption of plasmid relaxation complexes (Lovett and Helinski,1975; Rowbury,1977) causes plasmid DNA to band with chromosomal DNA in dye-buoyant density gradients. Similar problems reduce the accuracy of copy number estimates by alkaline sucrose gradient centrifugation. However, the method of Womble *et al.* claims to give very high and reproducible recoveries of covalently closed circular (CCC) DNA and provides a convenient way to compare relative copy numbers of plasmid derivatives.

The various electrophoretic methods provide a potentially quicker route to obtaining comparative data which may be useful for screening for copy number mutants. However the apparent simplicity of these methods is often overshadowed by a requirement for specialist analytical equipment, e.g. a dual wavelength chromatogram scanner (2a), or a microdensitometer (2c). The electrophoretic separation of radiolabelled DNA from a total cell lysate prepared in the gel (2b), is really only suitable for comparing copy numbers of the same plasmid e.g. when screening for a copy mutant. The method requires complete lysis, which has to be determined empirically by adjusting the volume of the aliquots loaded on the gel. Individual bands have to be excised from the gel, then solubilized overnight before counting. The advantages of electrophoretic methods are that they are usually easy to set up and run, but, more important, they allow separation of all plasmid forms away from chromosomal DNA.

The hybridization methods represent the most theoretically precise approach, but they are difficult to perform and contain inherent flaws that can only be avoided with great care. Filter

hybridization is not very efficient and loss of DNA from filters makes this the least desirable hybridization assay. The kinetics of reassociation of simple sequence DNA are reputedly easy to determine but a great deal more work is required to process the kinetic data reliably, e.g. a computer programme is required for the rate evaluation. Consequently this method is not suitable for routine use.

The saturation hybridisation in solution is reputed to produce unambiguous results, but may not be readily applicable to all organisms. The method requires the preparation of total intracellular DNA from plasmid-containing cells, which is subsequently labelled to high specific activity *in vitro* by nick translation with DNA polymerase I and used to prepare labelled single stranded tracer DNA which is then driven into hybrid form with excess purified plasmid DNA. The fraction of labelled DNA tracer, complementary to plasmid DNA, is rapidly driven into hybrid form because of the low sequence complexity of plasmid DNA. Progress of the reaction is monitored by chromotography through hydroxyapatite, which binds the fraction of the tracer equivalent to the fraction of total DNA that is plasmid specific. I attempted to prepare total intracellular DNA form *M.methylotrophus* cells containing various vector plasmids. In all cases, the alcohol precipitation step brought down a non-DNA substance (possibly a polysaccharide) that produced a viscous solution when re-dissolved in water. Also, agarose gel electrophoretic analysis of these preparations showed bands corresponding to CCC plasmid, indicating that the ratio of plasmid:chromosomal DNA was grossly exaggerated and therefore not a suitable substrate for the determination of plasmid copy number. [A similar occurrence has been reported for a whole-cell DNA preparation that used an isopropanol precipitation step (Projan *et al.*,1983)]

My interest in plasmid copy number measurement was to be able to compare, as accurately as possible, the copy number of a plasmid in *E.coli* and *M.methylotrophus*. This was important since in the assays for mouse dihydrofolate reductase and the synthetic interferon, small but consistent increases in the levels of expresssion were observed in the methylotroph, compared to the levels in *E.coli*.

There are, however, a number of complications involved when attempting these measurements in *M.methylotrophus*. The size of the chromosome is not known and therefore any measurements would have to be related to cell number. The size of a *M.methylotrophus* cell is less than half that of *E.coli* and is consequently more difficult to measure. I did obtain some estimates of cell number directly using a Coulter counter with a twenty micron orifice, but this proved to be too impractical for general use because of the time taken to accumulate reproducible numbers (a 20$\mu$ orifice becomes blocked extremely easily!). Direct cell counting under a microscope was also impractical in my hands. Yet another complication was that *in vivo* labelling of DNA was not possible by any of the standard techniques; thymine and thymidine for example, were not taken up when the cells were grown in liquid culture.

In view of these difficulties I decided to concentrate on obtaining data from plasmids in *E.coli*. by the centrifugation method of Womble *et al.*, with the ultimate aim of comparing the figures to those obtained by some other non-radiolabelled method, which may also be suitable for *M.methylotrophus*.

## 10.2 Determination of plasmid-to-chromosome ratios from bacterial lysates by dye-buoyant density centrifugation.

The method makes use of the fact that covalently closed circular (CCC) plasmid DNA does not bind as much of an intercalating dye as

nicked circular plasmid DNA, or linear chromosomal DNA (Bauer and Vinograd,1968). The buoyant density of CCC plasmid DNA in a caesium chloride gradient is consequently not lowered to the same extent as that of linear DNA, causing the former to band separately. Very high total recoveries of DNA have been reported for this method and standardisation of optimized procedures has led to reproducible recoveries of CCC DNA (Womble *et al.*,1977).

I have used the method of Womble *et al.* to analyse the CCC:chromosome DNA ratios in several plasmid-containing strains of *E.coli*. In some cases I have been interested to observe if any amplification took place after treatment with spectinomycin. For some plasmids I have also measured the CCC:chromosome ratios (and hence plasmid copy number) at different stages during the growth cycle.

**10.2(i). Growth of cells**

To achieve maximum labelling of cellular DNA *in vivo*, I initially measured plasmid copy numbers in a *thy*$^-$ strain of *E.coli*, BW85. However this strain was resistant to streptomycin, rendering it unsuitable for use with R300B ($Su^r$,$Sm^r$) since it is not possible to use sulphonamide drugs in a *thy*$^-$ strain. I have therefore used *E.coli* strain JA221 for R300B, together with several other plasmids previously measured in BW85 to ensure that there were no gross effects caused by changing strains.

Plasmid-containing *E.coli* strain BW85 cells were grown in minimal salts plus trace elements (as described for *M.methylotrophus*) containing 0.2% glucose, 5μg/ml thiamine, 4μg/ml leucine, 0.2% casamino acids, 200μg/ml deoxyguanosine, antibiotic as required and 2μg/ml of $^{14}C$ thymine (Amersham). Normally 10ml cultures, inoculated 1/100 from an unlabelled overnight culture, were grown at

37°C and 2ml samples were taken for analysis.

**10.2(ii). Preparation of lysates and centrifugation through CsCl.**

The cells from the samples were collected by centrifugation at 6krpm for 10min, washed twice with SV buffer (0.15M NaCl, 0.1M EDTA pH10.2) then resuspended in 1ml of SV buffer. 100μl of lysozyme solution (5mg/ml in TES: TES = 50mM each of Tris.HCl, EDTA and NaCl, pH8.0) were added and the mixture incubated at 37°C for 10min. 1ml of 1% Brij 58, 0.5% deoxycholate in TES was added followed by 100μl of Pronase (20mg/ml in TES, autodigested for 10min at 80°C) and incubation continued at 37°C for 2hr. The lysates were then transferred to a 65°C water bath and incubated for 30min, after which time they went completely clear and viscous. The chromosomal DNA was sheared by twenty passes up and down through a 1ml Gilson tip. A 0.5ml amount of ethidium bromide (2.5mg/ml in TES) was added together with more TES to make a final volume of 3.7ml. 3.45g of solid CsCl were added, then the solution poured into a 13.5ml polyallomer tube and topped up with paraffin oil. Centrifugation to equilibrium was achieved in a Beckman 50Ti rotor at 45krpm for 36-40hr at 23°C.

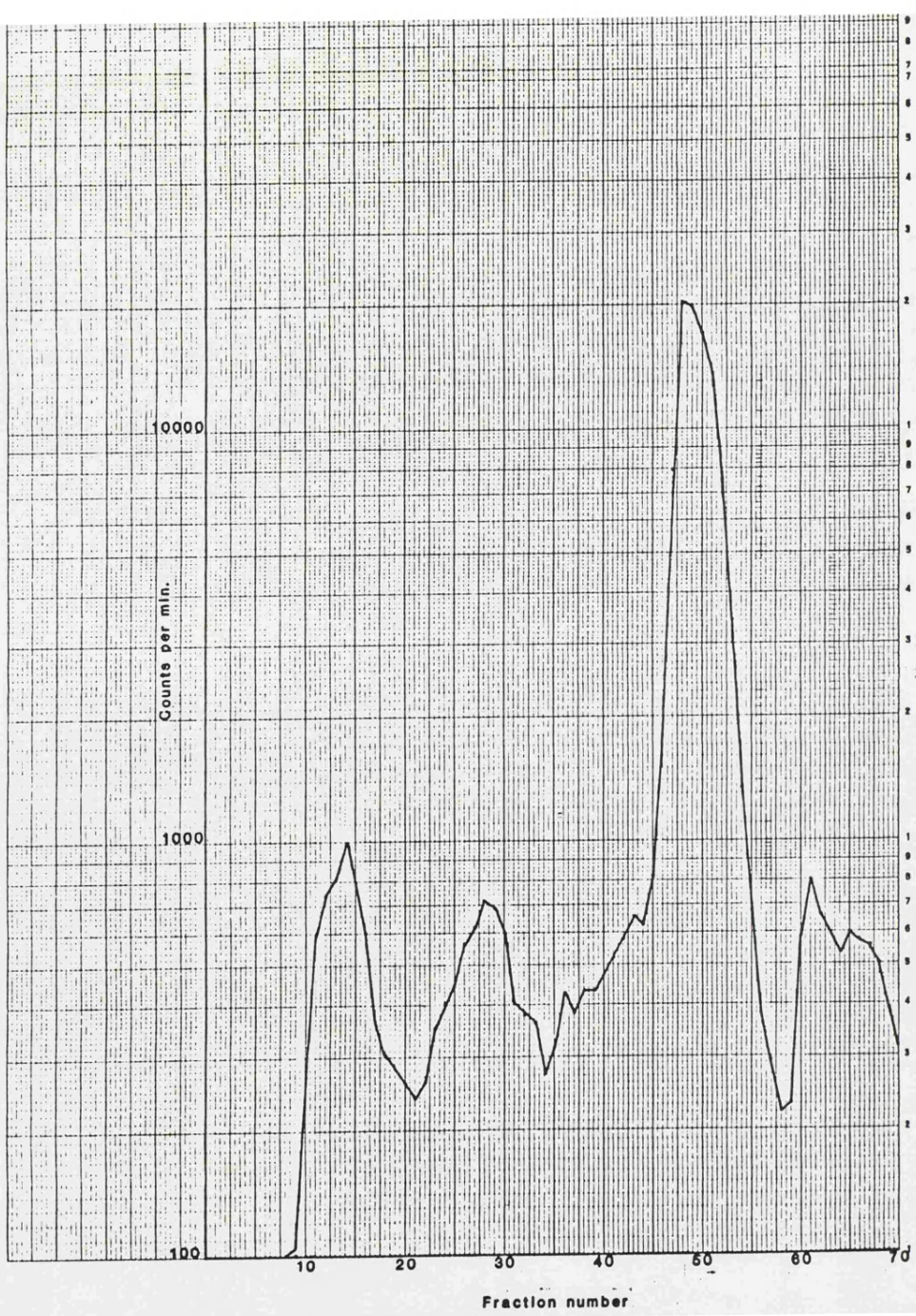
The gradient was fractionated by piercing the bottom of the tube with a fine needle and collecting 5-drop fractions onto one inch strips of Whatman 3MM filter paper marked off into one inch squares. The filter strips were dried , washed three times with ice-cold 7.5% trichloracetic acid, twice with 95% ethanol then dried. The filter squares were cut off the strips into scintillation vials containing 7ml of Fisofluor 3 (Fisons) and counted in a liquid scintillation counter.

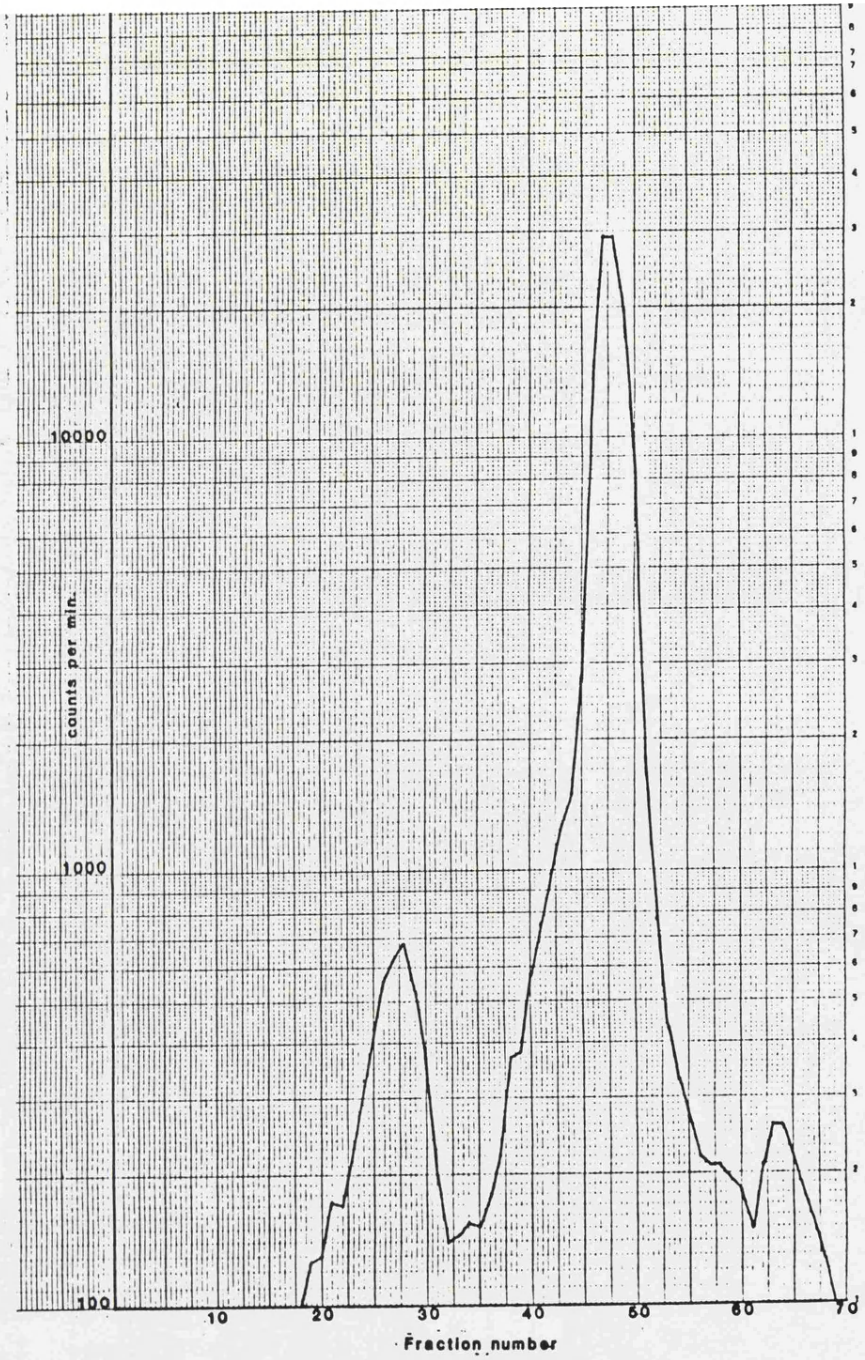
Typical profiles of the radioactivity throughout the gradients are shown in Fig.32. The total counts under each peak were determined and the number of plasmid copies per chromosome equivalent

**Figure 32.**

Fractionation profiles (70 × 5-drop fractions) of pGSS15 DNA in *E.coli* strain JA221 after labelling with $^{14}C$ thymine *in vivo* and subjecting whole cell lysates to caesium chloride/ethidium bromide density gradient centrifugation (Womble *et al.*,1977)

Fractionation was from the bottom of the tube such that CCC plasmid DNA peaked before chromosomal DNA (very large peak). However, the profiles differ in the regions corresponding to CCC DNA even though they were prepared from samples of the same culture.


10000
1000
100
Counts per min.
10
20
30
40
50
60
70
Fraction number


10000
1000
100
counts per min.
10
20
30
40
50
60
70
Fraction number

were calculated according to the formula:-

$$\text{N}^{\text{o}}\text{. of plasmid copies} = \frac{\text{cpm in CCC DNA}}{\text{cpm in chromosome}} \times \frac{\text{mol.wt. of chromosome}}{\text{mol.wt. of plasmid}}$$

The molecular weight of the *E.coli* chromosome was taken to be 2500Md (Cooper and Helmstetter,1968).

Some of the profiles showed two peaks in the region were CCC DNA would be expected to band. The appearence of these two peaks is at present unexplained, and in the following tables of the results [Tables V AND VI]I have separated the counts obtained in the two peaks and compared the copy numbers by using these different values in the calculations. The two-peak phenomenon does not seem to be plasmid, strain or growth stage dependent, and even in duplicate samples of the same culture, one sample had a single peak, the other two! [Fig.32].

Apparently this appears to be a phenomenon observed when the plasmids involved have a high copy number (D.Womble, personal communication). It is difficult to imagine what could produce a more dense band than CCC DNA, perhaps some form of RNA:DNA hybrid molecule, though incorporation of RNase into the lysis procedure did not significantly alter the overall patterns. Another possible explanation could be that the two peaks represent different levels of supercoiling, or even different forms (positive or negative) of the supercoiled plasmid.

Previously reported copy numbers for R300B have been 9-12 (Barth and Grinter,1974; Barth *et al.*,1978). However, other SuSm IncQ plasmids identical to R300B have reported copy numbers of 15-40 (Bagdasarian *et al.*,1979) and 13-25 (Rubens *et al.*,1976) for RSF1010 and 57 for R1162 (Meyer *et al.*,1982). My estimate for the copy number of R300B is 18-30 copies per chromosome equivalent. Whether

this estimate is correct as an absolute value may be open to debate, but the method is most useful for comparisons of relative copy numbers. Thus from the summary table [Table VI] it can be seen that all the pGSS series of vectors have essentially the same copy number.

Table V

Results from fractionation of CsCl/ethidium bromide density gradients.

| Plasmid | Size | cpm Plasmid | | | cpm | copies/Xsome | | Cells | Cell |
|---|---|---|---|---|---|---|---|---|---|
| | (kb) | Peak1 | Peak2 | Total | Xsome | 2 | 1+2 | | status |
| pBR322 | 4.36 | - | 48245 | 48245 | 1036312 | - | 40.4 | BW85 | Log |
| | | 7149 | 33751 | 40900 | 767479 | 38.2 | 46.3 | BW85 | Log |
| | | - | 6499 | 6499 | 147209 | - | 38.4 | JA221 | Log |
| | | - | 6750 | 6750 | 127465 | - | 46.0 | JA221 | Log |
| | | - | 969 | 969 | 11842 | - | 71.1 | JA221 | Log |
| | | - | 902 | 902 | 12395 | - | 63.2 | JA221 | Log |
| | | 52236 | 115602 | 167838 | 2358031 | 42.6 | 61.8 | BW85 | Stat |
| | | 31371 | 127517 | 158888 | 2288149 | 48.4 | 60.3 | BW85 | Stat |
| | | - | 14340 | 14340 | 71140 | - | 175.0 | JA221 | Stat |
| | | 66950 | 190256 | 257206 | 2197647 | 75.2 | 101.7 | BW85 | $Amp^d$ |
| | | 73487 | 238756 | 312243 | 1539721 | 134.7 | 176.1 | BW85 | $Amp^d$ |
| R300B | 8.68 | 2350 | 11679 | 14029 | 179030 | 28.5 | 34.2 | JA221 | Log |
| | | 3910 | 11128 | 15038 | 182509 | 26.6 | 36.0 | JA221 | Log |
| | | 19058 | 11299 | 30357 | 502152 | 9.8 | 26.4 | JA221 | Log |
| | | 19424 | 11009 | 30433 | 497737 | 9.7 | 26.7 | JA221 | Log |
| | | - | 83473 | 83473 | 1352385 | - | 27.0 | JA221 | Stat |

| Plasmid | Size | cpm Plasmid | | | cpm | Copies/Xsome | | Cells | Cell |
|---|---|---|---|---|---|---|---|---|---|
| | (kb) | Peak1 | Peak2 | Total | Xsome | 2 | 1+2 | | Status |
| pGSS8 | 10.0 | - | 21851 | 21851 | 901519 | - | 9.2 | BW85 | Log |
| | | - | 29056 | 29056 | 949747 | - | 11.6 | BW85 | Log |
| | | 15423 | 78562 | 93985 | 2226671 | 13.3 | 16.0 | BW85 | Stat |
| | | 13521 | 83747 | 97268 | 2121258 | 14.9 | 17.3 | BW85 | Stat |
| | | 32048 | 78041 | 110089 | 2013552 | 14.7 | 20.7 | BW85 | $Amp^d$ |
| | | 4209 | 43243 | 47452 | 1766149 | 9.3 | 10.2 | BW85 | $Amp^d$ |
| | | - | 696 | 696 | 18715 | - | 14.1 | JA221 | Log |
| | | - | 782 | 782 | 20491 | - | 14.4 | JA221 | Log |
| | | - | 5913 | 5913 | 122788 | - | 18.2 | JA221 | Stat |
| pGSS6 | 13.0 | 3520 | 9057 | 12577 | 135221 | 19.5 | 27.1 | JA221 | Log |
| | | 4245 | 9848 | 14093 | 129774 | 22.1 | 31.7 | JA221 | Log |
| pGSS15 | 12.0 | - | 4690 | 4690 | 111753 | - | 13.3 | BW85 | Log |
| | | 6241 | 6074 | 12315 | 95947 | 20.0 | 40.5 | BW85 | Log |
| | | 3859 | 7208 | 11067 | 130168 | 17.5 | 26.8 | JA221 | Log |
| | | 5764 | 8923 | 14687 | 128349 | 21.9 | 36.1 | JA221 | Log |
| | | 16696 | 23799 | 40495 | 226891 | 33.1 | 56.4 | BW85 | Stat |
| | | 6800 | 8646 | 15446 | 251707 | 10.8 | 19.4 | BW85 | Stat |
| | | 1808 | 22293 | 24101 | 188537 | 37.3 | 40.4 | BW85 | $Amp^d$ |
| | | 17855 | 27509 | 45364 | 185255 | 46.9 | 77.3 | BW85 | $Amp^d$ |
| pGSS26 | 12.4 | 7854 | 9829 | 17683 | 130605 | 23.0 | 41.4 | JA221 | Log |
| | | 5559 | 10122 | 15681 | 133030 | 23.3 | 36.0 | JA221 | Log |
| pGSS29 | 10.3 | - | 47214 | 47214 | 480592 | - | 35.8 | BW85 | Log |
| | | 25604 | 33092 | 58696 | 498024 | 24.2 | 43.0 | BW85 | Log |
| | | - | 51641 | 51641 | 522184 | - | 36.0 | BW85 | Log |
| | | - | 7692 | 7692 | 45827 | - | 61.2 | JA221 | Stat |

| Plasmid | Size (kb) | cpm Plasmid | | | cpm Xsome | Copies/Xsome | | Cells | Cell Status |
|---|---|---|---|---|---|---|---|---|---|
| | | Peak1 | Peak2 | Total | | 2 | 1+2 | | |
| pGSS33 | 13.4 | 3167 | 2909 | 6076 | 80669 | 10.2 | 21.3 | BW85 | Log |
| | | 3938 | 3426 | 7364 | 83877 | 11.6 | 24.8 | BW85 | Log |
| | | 7052 | 7357 | 14409 | 126828 | 16.4 | 32.1 | JA221 | Log |
| | | 7137 | 7982 | 15119 | 121756 | 18.5 | 35.1 | JA221 | Log |
| | | 4283 | 8515 | 12798 | 239633 | 10.1 | 15.1 | BW85 | Stat |
| | | 3807 | 8716 | 12523 | 228258 | 10.8 | 15.5 | BW85 | Stat |
| | | 4845 | 7658 | 12503 | 167365 | 12.9 | 21.1 | BW85 | $Amp^d$ |
| | | - | 7384 | 7384 | 169064 | - | 12.4 | BW85 | $Amp^d$ |
| | | - | 16787 | 16787 | 407308 | - | 11.7 | BW85 | Log |
| | | 13232 | 13124 | 26356 | 413380 | 9.0 | 18.0 | BW85 | Log |
| | | 17221 | 13250 | 30471 | 389613 | 9.6 | 22.1 | BW85 | Log |

Table VI

Summary of plasmid copy numbers.

| Plasmid | Average number of plasmid copies per chromosome in log phase | | | | No. of separate determinations |
|---|---|---|---|---|---|
| | Peak 2 | s.d. | Peaks 1+2 | s.d. | |
| pBR322 | 49.6 | 12.9 | 50.9 | 12.0 | 6 |
| R300B | 18.7 | 8.9 | 30.8 | 4.3 | 4 |
| pGSS6 | 20.8 | 1.3 | 29.4 | 2.3 | 2 |
| pGSS8 | 12.3 | 2.1 | 12.3 | 2.1 | 4 |
| pGSS15 | 18.2 | 3.2 | 29.2 | 10.4 | 4 |
| pGSS26 | 23.5 | 0.2 | 38.7 | 2.7 | 2 |
| pGSS29 | 32.0 | 5.5 | 38.3 | 3.3 | 3 |
| pGSS33 | 12.4 | 3.4 | 23.6 | 7.4 | 7 |

The plasmid pGSS29 is a $Cm^r$ derivative of pWS1.1 (W.Schuch personal communication) and has a copy number of 35 copies per chromosome equivalent when measured in mid-log phase cells. This copy number did not appear to correspond to the very high yields of plasmid DNA obtained (3mg from a 400ml stationary phase culture). A gel electrophoretic analysis (Stueber and Bujard,1982) of relative copy numbers of pGSS29 at different stages of the host growth cycle, revealed that the very high plasmid levels were only apparent in stationary phase cells. One measurement of the copy number of pGSS29 made from stationary phase cells by the method of Womble *et al.* only showed a slight increase to 61.2 copies per chromosome equivalent. At only about twice the value for R300B this did not account for the almost ten-fold DNA yield increase! These experiments are very preliminary and require repeating more stringently, but there is at least a slight possibility that the method of Stueber and Bujard could be used to compare plasmid DNA levels in different species and at different stages in the cell growth cycle. With some more work at careful quantification, the method may yield convincing relative plasmid copy numbers between *E.coli* and *M.methylotrophus*.

# CHAPTER 11

## Discussion and conclusions

The obligate methylotroph, *Methylophilus methylotrophus*, used by ICI for its SCP production, may represent a valuable alternative host organism to *E.coli* for the commercial production of heterologous gene products. The organism has the advantages of being safe, coupled with an ability to grow well on a cheap carbon source. Accumulated experience over many years and novel fermentation technology allows the potential scale of production to be extremely large, if required. There was, however, very little known about how well the organism could express heterologous coding sequences. The work of Windass *et al.* (1980) had provided a promising start by showing the organism to be capable of using an *E.coli* gene.

To be able to assess how well *M.methylotrophus* could express, or be made to express, heterologous coding sequences, particularly eukaryotic coding sequences, cloning vectors equivalent to those available for *E.coli* needed to be made. Since it had proved impossible to transform *M.methylotrophus* with plasmid DNA by any method currently available, a mobilizable plasmid was the best candidate for development of a cloning vector. R300B was an obvious choice as a basic replicon because of its unique broad host range and relatively small size.

A series of cloning vectors has been constructed by a combination of R300B's broad host range and several unique cloning sites derived from sequences of the *E.coli* plasmids pBR322 and pBR328. During these constructions, features and properties of R300B have been found that were novel, that confirmed observations of other workers or that raised questions for which, as yet, there are no clear answers.

The very early experiments to introduce the $Ap^r$ gene from pBR322 into R300B demonstrated that while it was possible to create $Ap^r$ plasmids in *E.coli*, genes essential for mobilization to, or maintenance in, other organisms may have been disrupted. This was beginning to suggest that essential functions for the broad host range of R300B were distributed around the plasmid, and confirmed the observation of Bagdasarian *et al.* (1981) for RSF1010. More recent evidence from the same group (Scherzinger *et al.*,1984) demonstrating the presence of three *trans*-acting proteins required for replication and/or maintenance of RSF1010 but remote from the replication origin, complicates the use of random insertions to add cloning sites.

However, my approach of using *Hae*II partial digests to form the cointegrate plasmid pGSS6 may have led to a more stable arrangement of the two parental plasmids. Plasmid stability in general is becoming increasingly important as plasmid vectors find applications in commercial processes. R300B is exceptionally stable in *E.coli* and *P.aeruginosa*, no plasmid loss being detectable in the absence of antibiotic-selection. Stability measurements in *M.methylotrophus* are not yet complete but the indications are that R300B is also stable in this organism. High stability of this type would be the ideal for any plasmid vector used in a commercial process, since it eliminates the need, for example, of an antibiotic-selection.

None of the pGSS series of vectors retain the extreme stability of R300B, though in *E.coli* they are not lost after twenty generations of exponential phase growth in the absence of selection. However, longer-term subculturing, including periods in stationary phase, did show plasmid loss at a level comparable to pBR322. In *P.aeruginosa* differences in the stability of some plasmids were observed; pGSS6 was found to be more stable than pGSS8, which was more stable than pGSS15 (J. Lewington, personal communication).

The plasmid pGSS8 represents a derivative of R300B with the least disruption to essential regions and the least amount of inserted foreign DNA sequences. However, this plasmid is less stable than R300B in both *E.coli* and *P.aeruginosa*, and thus raises the question of whether it is the point of insertion of the DNA sequence, or something in the DNA sequence itself, which causes the reduced stability. There is very recent evidence that the point of insertion, between the *Bst*EII site and the *Pst*I site just outside the $Su^r$ gene, could possibly disrupt the start of an essential gene (Scherzinger *et al.* 1984). They have reported that there are three genes, remote from the origin of replication, that produce *trans*-acting proteins essential for replication and/or maintenance. One of these genes, *rep*C, has its start mapped very close to the point of insertion of pBR322 sequence into R300B in pGSS6 and its derivatives. However, the *rep*C gene-product is thought to play a role in initiating replication of R300B (Scherzinger *et al.*,1984) and since the pGSS plasmids appear to replicate satisfactorily, this may not be the main cause of the instability.

Inserting a DNA fragment into different positions on the R300B genome, especially positions known to be non-essential, would provide a test for such "insertional instability". If, however, stability is not restored by this procedure, then different DNA sequences could be cloned into identical positions on the R300B genome and their stabilities compared. The extreme stability of R300B in *E.coli* and *P.aeruginosa* also raises the question of whether this plasmid carries some form of *par* sequence (Meacock and Cohen,1980) which guarantees at least one plasmid copy going to every daughter cell at division. There is as yet no evidence for such a sequence, though I feel sure one will be found.

The fact that R300B has sequences essential for replication

remote from its vegetative origin is unique in a plasmid of this size, and the property appears to be closely linked with the plasmid's broad host range. Similar features are normally associated with the much larger IncP group plasmid RP4/RK2 (57kb), where the replication and maintenance determinants are scattered over a 20-kb portion of the genome (Thomas *et al.*,1980). It has been speculated that this scattering of essential regions in RK2 results in this plasmid encoding more functions for its replication than narrow host range plasmids, making it less dependent on host replication functions and allowing a greater degree of promiscuity (Thomas *et al.*,1980). Work described in this thesis and in Barth *et al.* (1981) for R300B, and the recent paper of Scherzinger *et al.* (1984) for RSF1010, support this view.

Since the host range of R300B (and hence many of the pGSS vectors) is so broad, many more organisms (besides *E.coli* and *M. methylotrophus*) have the potential to become hosts for cloning experiments. The extended host range also provides the potential to study expression of cloned coding sequences in several different organisms while eliminating the need for further manipulation of the cloned genes.

The mobilization of these broad host range vectors by IncIα plasmids is interesting and may be more widely applicable. Two instances of the transfer of non-mobilizable plasmids by I-like R factors, to *Providentia* and *Proteus mirabilis* (Coetzee,1964; Datta and Hedges,1972) have been reported, in which the I factor was not detected in the recipients. In four cases reported here, Iα plasmids were able to promote transfer of IncQ plasmids into organisms outside the normal Iα host range, but again the conjugative plasmids themselves were never detected in the recipients. Thus Iα plasmids appear to be capable of achieving conjugal transfer to many

organisms, but only capable of stable maintenance in a few. For the purpose of transferring the pGSS series of vectors however, this is a positive advantage, providing a clean delivery of vector plasmid with a degree of containment, since subsequent mobilization would require introduction of another conjugative plasmid.

As observed for *M.methylotrophus*, the IncIα plasmid pLG221 not only allowed clean delivery of a pGSS vector plasmid but also provided a means of *in vivo* transposon mutagenesis. In matings, the Tn*5* transposon carried by pLG221 could be readily detected ($Km^r$) in the recipient strain while the Iα plasmid itself could not be detected. This observation was extended to use in *P.aeruginosa* where it was clearly demonstrated that the Tn*5* transposon from pLG221 was present in the recipient chromosome without trace of the Iα plasmid (Boulnois *et al.*,1984). Thus while the host range of the IncIα plasmids for maintenance may be relatively narrow, these plasmids appear to possess valuable "cell-accessing" properties which may be possible to exploit for other organisms.

There has been some confusion in the literature about whether or not IncQ plasmids require DNA polymerase I for replication. My data confirm the data of Grindley and Kelly (1974) that they do not. This was most clearly shown with pGSS4 which contained sequences of pBR322 and R300B, though it was not clear from which origin replication was being initiated. Nevertheless the plasmid could transform and replicate in a *polA* strain of *E.coli*. When the plasmid was cut and religated to form pGSS7, this plasmid could no longer replicate in the *polA* strain and was shown to have the pBR322 *ori*. Clearly a region of R300B that had been providing DNA polymerase I-independence had been lost. Gautier and Bonewald (1980) drew the opposite conclusion, but this may have been due to some instability of their plasmids in their *polA* strain (Grindley and Kelly,1974).

The plasmid pGSS15 has been the most widely used of the pGSS vectors so far. Having similarity to pBR322 and pAT153, it allows easy sub-cloning of sequences previously cloned in these *E.coli* vectors. It has been used to demonstrate the expression of mouse DHFR (this thesis), chicken ovalbumin (Hennam *et al.*,1981) and synthetic human α-1 interferon (DeMaeyer *et al.*,1982) in *M.methylotrophus*. The latter case has also demonstrated that increased expression from strong promoters is possible in *M.methylotrophus* and that a promoter which is strong in *E.coli* usually behaves similarly in the methylotroph. Analysis of pGSS15 revealed several features that were undesirable in such a vector (two replication origins, two *Sal*I sites and an inverted repeat sequence) but it did prove very convenient to have the antibiotic-resistance genes arranged exactly as in pBR322 and this led to the construction of pGSS33.

The plasmid pGSS33 represents a significant improvement in the cloning capacity of any broad host range vector presently available. However as with any development process there are always improvements which could be made. At 13.4kb pGSS33 is considerably larger than the *E.coli* plasmids, though as discussed above, most of its size is necessary to retain its broad host range. (Plasmid size is less of a problem if conjugal transfer, rather than transformation, is chosen as the means of introduction of recombinants into the host organisms). A size-reduction of up to 10% could be achieved by removal of extraneous sequences, derived from pBR322, which occur at the ends of the $Tc^r$ and $Ap^r$ genes, but the present size is probably close to the practical minimum.

It has been reported (Prentki *et al.*,1981) that the plasmid pBR325 contained a 482-bp inverted duplication of DNA sequence derived from pBR322. The restriction map of pBR328 (Soberon *et*

*al.*,1980) suggests that this plasmid has retained the inverted repitition from pBR325. The region in question lies between the *Eco*RI site and the start of the $Tc^r$ gene, and has therefore also been transferred to all the $Cm^r$ pGSS plasmids, including pGSS33. The implications of this repeat (and the similar one generated in the construction of pGSS15) are not very serious and at worst could sometimes complicate the analysis of recombinant molecules. The formation of snap-back structures (see for example Fig.18b) could be troublesome in heteroduplex analysis, but could be turned to advantage in some cases by providing a convenient internal marker in electron microscopic studies.

Gene-amplification is often a desirable property of a cloning vector, to be able to increase yields of a particular product. For plasmids like R300B the only possibility for amplification is to find an up-copy mutant. The plasmids pWS1.1 and the $Cm^r$ version, pGSS29, are both high copy number derivatives of R300B, but they are not lethal, indicating that there may still be some form of copy number control. Thus the deletion up to the *Pvu*II site at 2.1 on the R300B map results in the high copy phenotype and is probably a good place to begin looking for regions or genes involved in copy control, with a view to manipulating them.

Interestingly, pGSS7 which also yields very high levels of plasmid DNA, though from a pBR322 origin of replication, carries the whole region between the *Eco*RI site and the *Pvu*II site that was deleted from R300B to yield pWS1.1. Conversely, pWS1.1 carries the *rop* gene from pBR322 which is known to affect initiation of replication and hence copy number of pBR322. The two plasmids, pGSS7 and pWS1.1 should be compatible, and it would be interesting to observe if there were any effects on the copy number of either plasmid due to the presence of *trans*-acting products in the cell. A

reduced copy number of pWS1.1 in the presence of pGSS7 would confirm that the 2.1-kb *Eco*RI-*Pvu*II region of R300B should be a fruitful starting point to investigate copy control of this plasmid. The DNA sequence of this region may provide information of possible protein products and may also be of use in designing a strategy of directed mutagenesis. The fact that removal of this 2.1-kb fragment causes a higher copy number of R300B without cell-death is significant, suggesting that there could be other regions, either immediately adjacent or some distance away, that may also be involved in the plasmid's copy control mechanisms.

It could be this type of study which leads to a level of understanding of copy control mechanisms similar to that which now exists for the plasmid RI. This has allowed sophisticated manipulations of the system to produce, for example, a temperature -switchable runaway replication mode (Molin *et al.*,1981), which makes even materials toxic to a cell possible to prepare from a batch culture. Thus the use of strong promoters, together with high copy number or runaway replication broad host range plasmid vectors, could provide the basis to produce high yields of cloned gene products in *M.methylotrophus* or other Gram-negative bacteria not previously amenable to such processes.

While considering the use of strong promoters in cloning vectors for *M.methylotrophus* the obvious choice was to try readily available and proven *E.coli* promoters, e.g. *trp* and *lac*. However some thought was given to obtaining a strong *M.methylotrophus* promoter should this prove necessary.

Production of large quantities of a particular protein could be associated with a strong promoter and in *M.methylotrophus* approximately 5-10% of the soluble protein is methanol dehydrogenase (MDH) (D.Byrom, personal communication). The *mdh* gene may therefore

be associated with a srong promoter whose isolation and characterisation could be important in developing a homologous *M.methylotrophus* host/vector expression system.

Consequently there has been much interest, by several of my collegues, in isolating and characterising the *mdh* gene but to date the results are inconclusive. It was therefore fortuitous that the *E.coli* promoters worked as well as they did in *M.methylotrophus*, though for other organisms the homologous promoter system may be essential.

In summary, this thesis has described the construction and characterisation of a series of vectors, based on R300B, which maintain the broad host range but which have increased capacity for easy-to-use cloning sites. Many of the sites allow detection of recombinants by providing insertional inactivation of an antibiotic-resistance gene. Of the series, pGSS33 in particular, carries many of the most useful cloning sites of pBR328 together with *Sst*I and *Sst*II sites derived from R300B, providing alternatives to *Pst*I as sites useful for insertion of homopolymer-tailed fragments. Having four antibiotic-resistance genes, all with restriction sites capable of producing insertional inactivation, means that pGSS33 is an extremely versatile vector, and makes the convenience and simplicity of the well known and well used *E.coli* vectors, pBR322 and pBR328, available for use with most Gram-negative bacteria.

The first demonstration of expression of a eukaryotic coding sequence, murine dihydrofolate reductase, in *M.methylotrophus* has been described. This was followed by expression of chicken ovalbumin, *E.coli* B-galactosidase and synthetic human α-1 interferon, all making use of pGSS vectors. The plasmid pDHFR2.43 may turn out to be a valuable test plasmid for studying the stability of cloned eukaryotic coding sequences in both *E.coli* and *M.methylotrophus*; it

is readily detected ($Tp^r$) and can be grown with or without selection.

A start has been made towards providing increased expression from vectors carrying strong *E.coli* promoters which have been demonstrated to work well in *M.methylotrophus*.

Broad host range cosmid vectors have been constructed, though it may now be easier to sub-clone the *cos* region from a well-proven cosmid into e.g. pGSS33, and a start has also been made to demonstrate that broad host range cosmids could be used for the production of gene-banks in Gram-negative organisms other than *E.coli*. A point to consider here is that mobilization of recombinants from, for example *E.coli* to *M.methylotrophus*, should be as efficient as possible. This would probably require the use of an IncP group plasmid (e.g. RP4 or R751), rather than an IncIα plasmid which I have preferred to provide delivery of just the vector plasmid.

Although the *in vitro* DNA-dependent transcription/translation system from *M.methylotrophus* did not reach expectations, there is plenty of scope for development and some lessons may be learnt from this work and from systems developed for other organisms.

A lead has also been given in the search for a high copy number plasmid, which coupled with a strong promoter, could provide the basis for a very efficient batch production process.

Thus with the availability of easy-to-use cloning vectors, convenient delivery systems and the accumulated evidence of strong *E.coli* promoters working efficiently in *M.methylotrophus*, this organism can now seriously be considered as a safe alternative host to *E.coli*.

REFERENCES

**Austin,S.**, Sternberg,N. and Yarmolinsky,M.: Miniplasmids of bacteriophage P1 I. Stringent plasmid replication does not require elements that regulate the lytic cycle. J. Mol. Biol. 120 (1978) 297-309.

**Baccanari,D.**, Phillips,A., Smith,S., Sinski,D. and Burchall,J.: Purification and properties of *Escherichia coli* dihydrofolate reductase. Biochemistry 14 (1975) 5267-5273.

**Bagdasarian,M.**, Bagdasarian,M.M., Coleman,S. and Timmis,K.N.: New vector plasmids for gene cloning in *Pseudomonas*, in Timmis,K.N. and Pühler,A. (Eds.), Plasmids of Medical, Environmental and Commercial Importance, Elsevier/North Holland Biomedical Press, Amsterdam, 1979, pp. 411- 422.

**Bagdasarian,M.**, Lurz,R., Ruckert,B., Franklin,F.C.H., Bagdasarian, M.M., Frey,J. and Timmis,K.N.: Specific-purpose plasmid cloning vectors, II. Broad host range, high copy number RSF1010-derived vectors and a host-vector system for gene-cloning in *Pseudomonas*. Gene 16 (1981) 237-247.

**Barth,P.T.** and Grinter,N.J.: Comparison of the deoxyribonucleic acid molecular weights and homologies of plasmids conferring linked resistance to streptomycin and sulfonamides. J.Bacteriol. 120 (1974) 618-630.

**Barth,P.T.**: RP4 and R300B as wide host range plasmid cloning vehicles, in Timmis,K.N. and Pühler,A. (Eds.), Plasmids of Medical, Environmental and Commercial Importance, Elsevier/North Holland Biomedical Press, Amsterdam, 1979, pp. 399-410.

Barth,P.T., Richards,H. and Datta,N.: Copy numbers of coexisting plasmids in *E.coli* K12. J.Bacteriol. 135 (1978) 760-765.

**Barth,P.T.**, Tobin,L. and Sharpe,G.S.: Development of broad host range plasmid vectors, in Levy,S.B., Clowes,R.C. and Koenig,E.L. (Eds.),

Molecular Biology, Pathogenicity and Ecology of Bacterial Plasmids. Plenum Press, New York, 1981, pp. 439-448.

**Bauer,W.** and Vinograd,J.: The interaction of closed circular DNA with intercalating dyes. I. The superhelix density of SV40 DNA in the presence and absence of dye. J. Mol. Biol. 33 (1968) 141-171.

**Becker,A.** and Gold,M.: Isolation of the bacteriophage lambda *A*-gene protein. Proc. Natl. Acad. Sci. USA 72 (1975) 581-585.

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

**Bolivar,F.**, Rodriguez,R.L., Greene,P.J., Betlach,M.C., Heyneker,H.L., Boyer,H.W., Crosa,J.H. and Falkow,S.: Construction and characterization of new cloning vehicles, II. A multipurpose cloning system. Gene 2 (1977) 95-113.

Bolivar,F.: Construction and characterization of new cloning vehicles. III Derivatives of plasmid pBR322 carrying unique *Eco*RI sites for selection of *Eco*RI generated recombinant DNA molecules. Gene 4 (1978) 121-136.

**Boulnois,G.J.**: Colicin Ib does not cause plasmid-promoted abortive phage infection of *Escherichia coli* K12. Mol. Gen. Genet. 182 (1981) 508-510.

**Boulnois,G.J.**, Varley,J.M., Timmis,K.N., Sharpe,G.S. and Franklin,F.C.H.: Transposon donor plasmids, based on ColIb-P9, for use in *Pseudomonas putida* and a variety of other Gram negative bacteria. 1984 (Submitted for publication).

**Boyd,A.C.**: Restriction Systems in *Methylophilus methylotrophus*, Ph.D thesis, University of Leicester, 1983.

Broome-Smith,J.: *Rec*A independent, site-specific recombination between ColE1 or ColK and a miniplasmid they complement for mobilization and relaxation: implications for the mechanism of DNA

transfer during mobilization. Plasmid 4 (1980) 51-63.

**Bukhari,A.I.**, Shapiro,J.A. and Adhya,S.L. (Eds.) in DNA Insertion Elements, Plasmids and Episomes. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1977).

**Burchall,J.J.**: Comparative biochemistry of dihydrofolate reductase. Ann. N.Y. Acad. Sci. 186 (1971) 143-152.

**Burkhardt,H.**, Reiss,G. and Pühler,A.: Relationship of group P1 plasmids revealed by heteroduplex experiments: RP1, RP4, R68 and RK2 are identical. J. Gen. Microbiol. 114 (1979) 341-348.

**Burt,D.**: Gene-Fusion and the Analysis of Genetic Control Signals. Ph.D thesis, University of Leicester, 1981.

**Burt,D.** and Brammar,W.J.: The *cis*-specificity of the *Q*-gene product of bacteriophage lambda. Mol. Gen. Genet. 185 (1982) 468-472.

**Chakrobarty,A.M.**, Mylroie,J.R., Friello,D.A. and Vacca,J.G.: Transformation of *Pseudomonas putida* and *Escherichia coli* with plasmid-linked drug-resistance factor DNA. Proc. Natl. Acad. Sci. USA 72 (1975) 3647-3651.

**Chang,A.C.Y.** and Cohen,S.N.: Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. 134 (1978) 1141-1156.

**Chang,A.C.Y.**, Hunberg,J.N., Kaufman,R.J., Erlich,H.A., Shimke,R.T. and Cohen,S.N.: Phenotypic expression in *E.coli* of a DNA sequence coding for mouse dihydrofolate reductase. Nature (London) 275 (1978) 617-624.

**Clarke,L.** and Carbon,J.: Functional expression of cloned yeast DNA in *Escherichia coli*: specific complementation of argininosuccinate lyase (*arg*H) mutations. J. Mol. Biol. 120 (1978) 517-532.

Clewell,D.B.: Nature of ColE1 plasmid replication in *E.coli* in the presence of chloramphenicol. J. Bacteriol. 110 (1972) 667-676.

Clewell,D.B. and Helinski,D.R.: Supercoiled circular DNA-protein

complex in *E.coli* purification and induced conversion to an open circular DNA form. Proc. Natl. Acad. Sci. USA 62 (1969) 1159-1166.

**Coetzee,J.N.**: Transmission of colicinogeny to *Providence* strains. Nature (London) 203 (1964) 897-898.

**Cohen,S.N.**, Chang,A.C.Y. and Hsu,L.: Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. Proc. Natl. Acad. Sci. USA 69 (1973) 2110-2114.

**Collins,J.** and Hohn,B.: Cosmids: a type of plasmid gene-cloning vector that is packagable *in vitro* in bacteriophage λ heads. Proc. Natl. Acad. Sci. USA 75 (1978) 4242-4246.

**Collins,J.**, Yanofsky,S. and Helinski,D.R.: Altered DNA-protein relaxation complex in a replication mutant of plasmid ColE1. Mol. Gen. Genet. 167 (1978) 21-28.

**Cooper,S.** and Helmstetter,C.E.: Chromosome replication and the division cycle of *Escherichia coli* B/r. J. Mol. Biol. 31 (1968) 519-540.

**Covarrubias,L.**, Cervantes,L., Covarrubias,A., Soberon,X., Vichido,I., Blanco,A., Kupersztoch-Portnoy,Y.M. and Bolivar,F.: Construction and characterization of new cloning vehicles V. mobilization and coding properties of pBR322 and several deletion derivatives including pBR327 and pBR328. Gene 13 (1981) 25-35.

Dagert,M. and Ehrlich,S.D.: Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. Gene 6 (1979) 23-28.

**Datta,N.**, Hedges,R.W., Shaw,E.J., Sykes,R.B. and Richmond,M.H.: Properties of an R factor from *Pseudomonas aeruginosa*. J. Bacteriol. 108 (1971) 1244-1249.

**Datta,N.** and Hedges,R.W.: Host ranges of R factors. J. Gen. Microbiol. 70 (1972) 453-460.

**David,M.**, Tranchet,M. and Denarie,J.: Transformation of *Azotobacter*

*vinelandii* with plasmids RP4 (IncP-1 group) and RSF1010 (IncQ group). J. Bacteriol. 146 (1981) 1154-1157.

**DeGraaff,J.**, Crosa,J.H., Heffron,F. and Falkow,S.: Replication of the nonconjugative plasmid RSF1010 in *Escherichia coli* K-12. J. Bacteriol. 134 (1978) 1117-1122.

**DeMaeyer,E.**, Skup,D. Prasad,K.S.N., DeMaeyer-Guignard,J., Williams,B., Meacock,P., Sharpe,G.S., Pioli,D., Hennam,J., Schuch,W. and Atherton, K.: Expression of a chemically synthesized human α-1 interferon gene. Proc. Natl. Acad. Sci. USA. 79 (1982) 4256-4259.

**DeVries,J.K.** and Zubay,G.: Characterization of a β-galactosidase formed between a complementary protein and a peptide synthesized *de novo*. J. Bacteriol. 97 (1969) 1419-1425.

**Ditta,G.**, Stanfield,S., Corbin,D. and Helinski,D.R.: Broad host range DNA cloning system for Gram negative Bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA 77 (1980) 7347-7351.

**Dougan,G.** and Sherratt,D.: The transposon Tn*1* as a probe for studying ColE1 structure and function. Mol. Gen. Genet. 151 (1977) 151-160.

Dretzen,G., Bellard,M., Sassone-Corsi,P. and Chambon,P.: A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. Anal. Biochem. 112 (1981) 295.

**Edman,J.C.**, Hallewell,R.A., Valenzuela,P., Goodman,H.M. and Rutter,W.J.: Synthesis of hepatitis B surface and core antigens in *E.coli*. Nature (London) 291 (1981) 503-506.

**Edge,M.D.**, Greene,A.R., Heathcliffe,G.R., Meacock,P.A., Schuch,W., Scanlon,D.B., Atkinson,T.C., Newton,C.R. and Markham,A.F.: Total synthesis of a human leukocyte interferon gene. Nature (London) 292 (1981) 756-762.

**Emtage,J.S.**, Tacon,W.C.A., Catlin,G.H., Jenkins,B., Porter,A.G. and Carey, N.H.: Influenza antigenic determinants are expressed from

haemagglutinin genes cloned in *Escherichia coli*. Nature (London) 283 (1980) 171-174.

**Everett,R.** and Willetts,N.: Cloning, mutation, and location of the F origin of conjugal transfer. EMBO. J. 1 (1982) 747-753.

Finnegan,J., and Sherratt,D.: Plasmid ColE1 conjugal mobility: the nature of *bom* a region required in *cis* for transfer. Mol. Gen. Genet. 185 (1982) 344-351.

**Fornari,C.S.**, and Kaplan,S.: Genetic transformation of *Rhodopseudomonas sheroides* by plasmid DNA. J. Bacteriol. 152 (1982) 89-97.

**Franklin,F.C.H.**, Bagdasarian,M. and Timmis,K.N.: Manipulation of degradative genes of soil bacteria, in Hütter,R. and Leisinger,T. (Eds.), Microbial Degradation of Xenobiotics and Recalcitrant Compounds, Academic Press, London 1981.

Gautier,F. and Bonewald,R.: The use of the plasmid R1162 and derivatives for gene cloning in the methanol-utilizing *Pseudomonas* AM1. Molec. Gen. Genet. 178 (1980) 375-380.

**Gelfand,D.H.**, Shepherd,H.M., O'Farrell,P.H. and Polisky,B.: Isolation and characterization of a ColE1-derived plasmid copy number mutant. Proc. Natl. Acad. Sci. USA 75 (1978) 5869-5873.

**Goeddel,D.V.**, Heyneker,H.L., Hozumi,T., Arentzen,R., Itakura,K., Yansura,D.G., Crea,R. and Seeberg,P.H.: Direct expression in *Escherichia coli* of a DNA sequence coding for human growth hormone. Nature (london) 281 (1979a) 544-548.

Goeddel,D.V., Kleid,D.G., Bolivar,F., Heyneker,H.L., Yansura, D.G., Crea,R., Hirose,T., Krzenwski,A., Itakura,K. and Riggs,A.D.: Expression in *Escherichia coli* of chemically synthetised genes for human insulin. Proc. Natl. Acad. Sci. USA 76 (1979b) 106-110.

Grindley,N.D.F. and Kelly,W.: Effects of different alleles of the *E.coli* K12 *polA* gene on the replication of non-transferring plasmids.

Mol. Gen. Genet. 143 (1974) 311.

**Grinter,N.J.** and Barth,P.T.: Characterization of SmSu plasmids by restriction endonuclease cleavage and compatibility testing. J.Bacteriol. 128 (1976) 394-400.

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

**Guerry,P.**, van Embdon,J. and Falkow,S.: Molecular nature of two nonconjugative plasmids carrying drug resistance genes. J. Bacteriol. 117 (1974) 619-630.

**Heffron,F.**, Rubens,C. and Falkow,S.: Translocation of a plasmid DNA sequence which mediates ampicillin resistance: molecular nature and specificity of insertion. Proc. Natl. Acad. Sci. USA 72 (1975) 3623-3627.

**Hennam,J.F.**, Cunningham,A.E., Sharpe,G.S. and Atherton,K.T.: Expression of eukaryotic coding sequences in *Methylophilus methylotrophus*. Nature (London) 297 (1982) 80-82.

**Hille,J.** and Schilperoort,R.: Behaviour of IncQ plasmids in *Agrobacterium tumefaciens*. Plasmid 6 (1981) 360-362.

Hochmannova,J and Nesvera,J.: A rapid and economic method for estimation of the number of plasmid copies in *Escherichia coli* cells. Folia. Microbiol. 27 (1982) 138-141.

**Hohn,B.**: DNA sequences necessary for packaging of bacteriophage λ DNA. Proc. Natl. Acad. Sci. USA 80 (1983) 7456-7460.

Hohn,B. and Collins,J.: A small cosmid for efficient cloning of large DNA fragments. Gene 11 (1980) 291-298.

Holmes,D.S. and Quigley,M.: A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114 (1981) 193-197.

Inselberg,J.: Replication of colicin E1 plasmid DNA in minicells from a unique replication initiation site. Proc. Natl. Acad. Sci. USA 71

(1974) 2256-2259.

**Itakura,K.**, Hirose,T., Crea,R., Riggs,A.D., Heyneker,H.L., Bolivar,F and Boyer,H.W.: Expression in *Escherichia coli* of a chemically synthetised gene for the hormone somatostatin. Science N.Y. (1977) 1056-1063.

**Jacoby,G.A.** and Shapiro,J.A.: Plasmids studied in *Pseudomonas aeruginosa* and other enteric bacteria, in Bukhari,A.J., Shapiro,J.A. and Adhya,S.L. (Eds.) DNA Insertion Elements, Plasmids and Episomes, Cold Spring Harbor Laboratory, Cold Spring Harbor N.Y. 1977.

Jacoby,G.A., Rogers,J.E., Jacob,A.E. and Hedges,R.W.: Transposition of *Pseudomonas* toluene-degrading genes and expression in *Escherichia coli*. Nature (London) 274 (1978) 179-180.

**Kahn,M** and Helinski,D.R.: Construction of a novel plasmid-phage hybrid: use of the hybrid to demonstrate ColE1 DNA replication *in vivo* in the absence of a ColE1-specified protein. Proc. Natl. Acad. Sci. USA 75 (1979) 2200-2204.

Kahn,M., Kolter,R., Thomas,C., Figurski,D., Meyer,R., Remaut,E. and Helinski,D.R.: Plasmid cloning vehicles derived from plasmids ColE1, F, R6K and RK2, in Wu,R. (Ed.) Methods in Enzymology Vol. 68 pp268-280 Academic Press N.Y. 1979.

Kingsbury,D.T. and Helinski,D.R.: DNA polymerase as a requirement for the maintenance of the bacterial plasmid colicinogenic factor E1. Biochem. Biophys. Res. Commun. 41 (1970) 1538-1544.

**Lovett,M.A.** and Helinski,D.R.: Relaxation complexes of plasmid DNA and protein II. Characterization of the proteins associated with the unrelaxed and relaxed complexes of plasmid ColE1. J. Biol. Chem. 250 (1975) 8790-8795.

**MacLennan,D.G.**, Gow,J.S. and Stringer,D.A.: Methanol - bacterium process for SCP. Process Biochem. June 1973.

Mandel,M. and Higa,A.: Calcium dependent bacteriophage DNA infection.

J. Mol. Biol. 53 (1970) 159-162.

**Mann,M.B.**, Rao,R.N. and Smith,H.O.: Cloning of restriction and modification genes in *E.coli*: the*Hha*II system from *Haemophilus haemolyticus*. Gene 3 (1978) 97-112.

**Martial,J.A.**, Hallewell,R.A., Baxter,J.D. and Goodman,H.M.: Human growth hormone: complementary DNA cloning and expression in bacteria. Science 205 (1979) 602-607.

**M$^{C}$Donnel,M.W.**, Simon,M.N. and Studier,F.W.: Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J. Mol. Biol. 110 (1977) 119-146.

**Meacock,P.A.** and Cohen,S.N.: Partitioning of bacterial plasmids during cell division: a *cis*-acting locus that accomplishes stable plasmid inheritance. Cell 20 (1980) 529-542.

**Mercereau-Puijalon,O.**, Royal,A., Cami,B., Garapin,A., Krust,A., Cannon,F. and Kourilsky,P.: Synthesis of an ovalbumin-like protein by *Escherichia coli* K12 harouring a recombinant plasmid. Nature (London) 275 (1978) 505-510.

**Meyer,R.**, Boch,G. and Shapiro,J.: Transposition of DNA inserted into deletions of the Tn*5* kanamycin resistance element. Molec. Gen. Genet. 171 (1979) 7-13.

**Meyer,R.**, Laux,R., Boch,G., Hinds,M., Bayly,R. and Shapiro,J.: Broad host range IncP4 plasmid R1162: effects of deletions and insertions on plasmid maintenance and host range. J. Bacteriol. 152 (1982) 140-150.

**Miller,J.H.**: Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1972.

Miller,J.H., and Reznikoff,W.S. (Eds.) in The Operon, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. 1978.

Nagahari,K. and Sakaguchi,S.: RSF1010 plasmid as a potentially useful

vector in *Pseudomonas* species. J. Bacteriol. 133 (1978) 1527-1529.

**O'Farrell,P.H.**, Kutter,E. and Nakanishi,M.: A restriction map of the bacteriophage T4 genome. Mol. Gen. Genet. 179 (1980) 421-435.

**Olson**,K.C., Fenno,J., Lin,N., Harkin,R.N., Snider,C., Kohr,W.H., Ross,M.J., Fodge,D., Prender,G. and Stebbing,N.: Purified human growth hormone from *E.coli* is biologically active. Nature (London) 293 (1981) 408-411.

**Pasek,M.**, Goto,T., Gilbert,W., Zink,B., Schaller,H., MacKay,P., Leadbetter,G. and Murray,K.: Hepatitis B virus genes and their expression in *E.coli*. Nature (London) 282 (1979) 575-579.

**Peden,K.W.C.**: Revised sequence of the tetracyline-resistance gene of pBR322. Gene 22 (1983) 277-280.

Perlman,D. and Stickgold,R.: Selective amplification of genes on the R plasmid, NR1, in *Proteus mirabilis*: an example of the induction of selective gene amplification. Proc. Natl. Acad. Sci. USA 74 (1977) 2518-2522.

**Prentki,P.**, Karch,F., Iida,S. and Meyer,J.: The plasmid cloning vector pBR325 contains a 482 base-pair-long inverted duplication. Gene 14 (1981) 289-299.

**Projan,S.J.**, Carleton,S. and Novick,R.P.: Determination of plasmid copy number by fluorescence densitometry. Plasmid 9 (1983) 182-190.

Roberts,R.J.: Restriction and modification enzymes and their recognition sequences. Nucl. Acids Res. 12 (1984) r167-r204.

Roberts,T.M., Kacich,R. and Ptashne,M.: A general method for maximising the expression of a cloned gene. Proc. Natl. Acad. Sci. USA 76 (1979) 760-764.

**Rowbury,R.J.**: Bacterial plasmids with particular reference to their replication and transfer properties. Prog. Biophys. Mol. Biol. 31 (1977) 271.

**Rubens,C.**, Heffron,F. and Falkow,S.: Transposition of a plasmid

deoxyribonucleic acid sequence that mediates ampicillin resistance: Independence from host *Rec* functions and orientation of insertion. J. Bacteriol. 1218 (1976) 425-434.

**Scalenghe,F.**, Turco,E., Edström,J.E., Pirrotta,V. and Melli,M.: Microdissection and cloning of DNA from a specific region of *Drosophila melanogaster* polytene chromosomes. Chromosoma 32 (1981) 205-216.

**Scherzinger,E.**, Bagdasarian,M.M., Scholz,P., Lurz,R., Rückert,B. and Bagdasarian,M.: Replication of the broad host range plasmid RSF1010: requirement for three plasmid-encoded functions. Proc. Natl. Acad. Sci. USA 81 (1984) 654-658.

**Sharp,P.A.**, Sugden,B. and Sambrook,J.: Detection of two restriction endonuclease activities in *Haemophilus parainfluenzea* using analytical agarose. Biochemistry 12 (1973) 3055-3063.

**Sherratt,D.J.**: Plasmid vectors for genetic manipulation *in vitro*. Biochem. Soc. Symp. 44 (1979) 29-36.

**Shine,J.** and Dalgarno,L.: Determinant of cistron specificity in bacterial ribosomes. Nature (London) 254 (1975) 34-38.

Smith,H.R., Humphreys,G.O. and Anderson,E.S.: Genetic and molecular characterization of some non-transferring plasmids. Molec. Gen. Genet. 129 (1974) 229

**Soberon,X.**, Covarrubias,L. and Bolivar,F.: Construction and characterization of new cloning vehicles, IV. Deletion derivatives of pBR322 and pBR325. Gene 9 (1980) 287-305.

**Sternberg,N.**, Tiemeier,D. and Enquist,L.: *In vitro* packaging of a λD*am* vector containing *Eco*RI DNA fragments of *Escherichia coli* and phage P1. Gene 1 (1977) 255-.

**Stüber,D.** and Bujard,H.: Organization of transcriptional signals in plasmids pBR322 and pACYC184. Proc. Natl. Acad. Sci. USA 78 (1981) 167-171.

**Stueber,D.** and Bujard,H.: Transcription from efficient promoters can interfere with plasmid replication and diminish expression of plasmid specified genes. EMBO J. 1 (1982) 1399-1404.

**Sutcliffe,J.G.**: Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. Cold Spring Harb. Symp. Quant. Biol. 43 (1978a) 77-90.

**Sutcliffe,J.G.**: pBR322 restriction map derived from the DNA sequence: accurate size markers up to 4361 nucleotide pairs long. Nucl. Acids Res. 5 (1978b) 2721-2728.

**Tait,R.C.**, Close,T.J., Lundquist,R.C., Hagiya,M., Rodriguez,R.L. and Kado,C.I.: Construction and characterisation of a versatile broad host range DNA cloning system for Gram-negative bacteria. Biotechnology 1 (1983) 269-275.

Tanaka,T. and Weisblum,B.: Construction of a colicin E1-R factor composite plasmid *in vitro*: means for amplification of deoxyribonucleic acid. J. Bacteriol. 121 (1975) 354-362.

**Thomas,C.M.**, Meyer,R. and Helinski,D.R.: Regions of broad-host-range plasmid RK2 which are essential for replication and maintenance. J. Bacteriol. 141 (1980) 213-222.

**Twigg,A.J.** and Sherratt,D.J.: *Trans*-complementable copy number mutants of plasmid ColE1. Nature (London) 283 (1980) 216-218.

Tyler,B.: Regulation of the assimilation of nitrogen compounds. Ann. Rev. Biochem. 47 (1978) 1127-1162.

**Uhlin,B.E.** and Nordström,K.: Plasmid incompatibility and control of replication: copy mutants of the R-factor R1 in *Escherichia coli* K12. J. Bacteriol. 124 (1975) 641-649.

**Uhlin,B.E.** and Nordström,K.: R plasmid gene dosage effects in *Escherichia coli* K12: copy mutants of the R plasmid R1*drd*19. Plasmid 1 (1977) 1-7.

**Uhlin,B.E.** and Nordström,K.: A runaway-replication mutant of plasmid

R1*drd*19: temperature-dependent loss of copy number control. Mol. Gen. Genet. 165 (1978) 165-179.

**Villa-Komaroff,L.**, Efstratiadis,A., Broome,S., Lomedico,P., Tizard,R., Naber,S.P., Chick,W.L. and Gilbert,W.: A bacterial clone synthesising proinsulin. Proc. Natl. Acad. Sci. USA 75 (1978) 3727-3731.

**Warnaar,S.O.** and Cohen,J.A.: A quantitative assay for DNA-DNA hybrids using membrane filters. Biochem. Biophys. Res. Comm. 25 (1966) 554-558.

**Warren,G.J.**, Twigg,A.J. and Sherratt,D.J.: ColE1 plasmid mobility and relaxation complex. Nature (London) 274 (1978) 259-261.

Weisblum,B., Graham,M.Y., Gryczan,T. and Dubnau,D.: Plasmid copy number control: isolation and characterization of high-copy-number mutants of plasmid pE194. J. Bacteriol. 137 (1979) 635-643.

**Weiss,B.**, Jacquemin-Sablon,A., Live,T.R., Fareed,G.C. and Richardson,C.C.: Enzymatic breakage and joining of deoxyribonucleic acid. VI Further purification and properties of polynucleotide ligase from *Escherichia coli* infected with bacteriophage T4. J. Biol. Chem. 243 (1968) 4543.

**Westmoreland,B.C.**, Szybalski,W. and Ris,H.: Mapping of deletions and substitutions in heteroduplex DNA molecules of bacteriophage lambda by electron microscopy. Science 163 (1969) 1343-1348.

**Wetzel,R.** Perry,L.J. Estwell,D.A., Lin,N., Levine,H.L., Slinker,B., Fields,F., Ross.M.J. and Shively,J.: Properties of a human alpha interferon purified from *E.coli* extracts. J. Interferon Res. 1 (1981) 381-390.

**Willetts,N.** and Crowther,C.: Mobilization of the non-conjugative IncQ plasmid RSF1010. Genet. Res. 37 (1981) 311-316.

Windass,J.D., Worsey,M.J., Pioli,E.M., Pioli,D., Barth,P.T., Atherton,K.T., Dart,E.C., Byrom,D., Powell,K. and Senior,P.J.:

Improved conversion of methanol to single-cell protein by *Methylophilus methylotrophus*. Nature (London) 287 (1980) 396-401.

**Womble,D.D.**, Dean,P.T. and Rownd, R.H.: Method for obtaining more-accurate covalently closed circular plasmid-to-chromosome ratios from bacterial lysates by dye-buoyant density centrifugation. J. Bacteriol. 130 (1977) 148-153.

Wood,D., Hollinger,M.F. and Tindol,M.B.: Versatile cloning vector for *Pseudomonas aeruginosa*. J. Bacteriol. 145 (1981) 1448-1451.

Zubay,G.: Cell-free studies on the regulation of the *lac* operon. Ann. Rev. Genet. 7 (1973) 267-287.

**ADDITIONAL REFERENCES**

**Collins,J.**: Cell-free synthesis of proteins coding for mobilisation functions of ColE1 and transposition functions of Tn3. Gene 6 (1979) 29-42.

**Kirby,K.S.**: A new method for the isolation of nucleic acids from mammalian tissues. Biochem. J. 64 (1956) 405.

Lowry,O.H., Rosebrough,N.J., Farr,A.L. and Randall,R.J.: Protein measurement with the Folin phenol reagent. J. Biol. Chem. 193 (1951) 265-270.

**Madin,S.H.** and Darby,N.B.: Proc. Soc. Exp. Biol. Med. 98 (1958) 574.

**Pratt,J.M.**, Boulnois,G.J., Darby,V., Orr,E., Wahle,E. and Holland, I.B.: Identification of gene products programmed by restriction endonuclese fragments using an *E.coli in vitro* system. Nucleic Acid Res. 9 (1981) 4459-4474.

**Sharpe,G.S.**: Broad host range cloning vectors for Gram-negative bacteria. Gene 29 (1984) 93-102.

Reprinted from Nature, Vol. 297, No. 5861, pp. 80-82, 6 May 1982


# Expression of eukaryotic coding sequences in *Methylophilus methylotrophus*

**John F. Hennam*, Angela E. Cunningham*, Geoffrey S. Sharpe† & Keith T. Atherton***

* Corporate Research Laboratory, ICI Ltd, Runcorn Heath, Cheshire WA7 4QE, UK
† ICI/University Joint Laboratory, Department of Biochemistry, University of Leicester, Leicester LE1 7RH, UK

**Although there have been many reports of the expression of eukaryotic genes in *Escherichia coli* K-12 (refs 1–4), very little attention has focused on the expression of these sequences in other bacteria. The use of other organisms may prove advantageous, especially with regard to industrial processes. For example, work is in progress to develop alternative host/vector systems based on *Bacillus subtilis*[5] and *Saccharomyces cerevisiae*[6]. The obligate methylotroph, *Methylophilus methylotrophus*, is a fermentation organism which grows very well on cheap substrates, methanol and ammonia. Moreover, obligate methylotrophs are unlikely to infect man and hence may be considered relatively safe. Here we report the expression of two eukaryotic cDNAs encoding chicken ovalbumin and mouse dihydrofolate reductase in *M. methylotrophus*, hence demonstrating the potential of this organism for the commercial production of mammalian peptides.**

Most cloning vectors developed to date are plasmids or phages specific for *E. coli*. Wide host-range cloning vectors based on RP4 and other IncP group plasmids have been developed[7], but these vectors are large, have a low copy number and are therefore not suitable for cloning when increased expression due to copy number is desired. For ease of manipulation and high gene dosage, small, high copy-number, broad host-range plasmids are needed. IncQ group plasmids, including R300B[8], meet these criteria and although they are themselves non-conjugative, they are mobilized efficiently by conjugative plasmids of various groups[8]. Despite its advantages, however, R300B has relatively few restriction enzyme sites suitable for cloning. The composite plasmid, pGSS15[9], has both the ampicillin ($Ap^r$) and tetracycline ($Tc^r$) resistance from pBR322 while retaining the broad host-range characteristics of R300B and hence contains a number of unique restriction enzyme sites for cloning. Moreover, cloning of eukaryotic cDNA into the *PstI* site of the $Ap^r$ gene by dG.dC tailing has been demonstrated to lead to the expression of either fused[10] or native[11] eukaryotic polypeptides at least in *E. coli*.

It has been reported previously that using broad host-range cloning vectors, the *E. coli gdh* gene has been expressed in *M. methylotrophus*[12]. Here we have used pGSS15 to demonstrate expression of the mouse dihydrofolate reductase and chicken ovalbumin genes in *M. methylotrophus*.

Copy DNA derived from the mRNA for mouse dihydrofolate reductase (DHFR) has been cloned and expressed in *E. coli*[13]; bacteria which synthesize mouse DHFR can be selected because they are resistant to levels of trimethoprim that would ordinarily inhibit growth.

The plasmid pDHFR7 (ref. 13) was constructed by annealing deoxy-C-'tailed' cDNA to pBR322 DNA which had been cleaved in the β-lactamase gene by *PstI* and 'tailed' with deoxy-G residues. The *PstI* sites were regenerated on both ends of the insert as a result of such recircularization. The cDNA insert of pDHFR7 was isolated after *PstI* cleavage of the plasmid and 1 μg of this DNA was mixed with 0.5 μg of pGSS15 DNA which had been linearized by *PstI* digestion. The mixture was ligated and used to transform *E. coli* ED8654 (*met*B, $rk^-mk^+$, *sup*F). Out of 221 of the trimethoprim-resistant ($Tp^r$) clones, 216 were

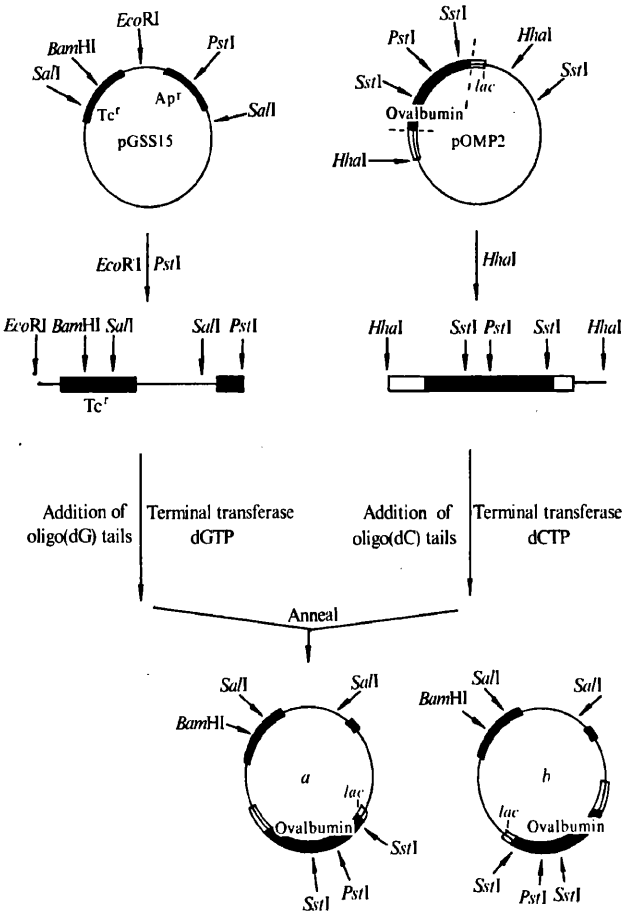



**Fig. 1** Outline of the construction of pGSS15 derivatives carrying ovalbumin cDNA sequences fused to the *lac* UV5 promoter. pGSS15 DNA (1 μg) was digested with *Eco*RI and *PstI* and the larger fragment isolated from an agarose gel using hydroxyapatite[25]. A *Hha*I DNA fragment containing the ovalbumin sequence and *lac* UV5 promoter from pOMP2 was isolated using the same technique. In the case of pOMP2, several other *Hha*I sites are present outside the *lac*–ovalbumin region (not shown). Oligo (dC) or (dG) tails were added using calf thymus terminal transferase, in the presence of 1 mM $CaCl_2$, and the two DNAs were mixed in equimolar amounts, at a total concentration of 1 μg $ml^{-1}$, for annealing[26]. Aliquots were used to transform $CaCl_2$-treated strain CL499 and transformants selected on nutrient plates containing (10 μg $ml^{-1}$) tetracycline. Colonies were tested for the presence of ovalbumin DNA sequence by hybridization to the *Taq*I ovalbumin fragment labelled with $^{32}P$, and finally for the production of immunoreactive ovalbumin.

ampicillin-sensitive, showing that the DHFR fragment had been inserted into the *bla* gene. Gel electrophoretic analysis showed that all the plasmids in the $Tp^rAp^s$ clones were the same size and were larger than pGSS15. The DHFR cDNA has a single *Bgl*II restriction site which allows orientation of the fragment. *Eco*RI/*Bgl*II digests of plasmids from $Tc^rAp^sTp^s$ and $Tc^rAp^sTp^r$ clones established that only one orientation led to functional expression of the DHFR sequence, indicating that productive transcription was from the β-lactamase promoter.

The plasmid R64drd11 (ref. 14) was introduced into one $Tp^r$ clone, pDHFR 2.43, and used to mobilize this hybrid plasmid into *M. methylotrophus*. Table 1 shows the levels of dihydrofolate reductase activity in extracts of *E. coli* (pDHFR 2.43) and *M. methylotrophus* (pDHFR 2.43). When trimethoprim (100 μg $ml^{-1}$) was added to the *M. methylotrophus* extracts, complete inhibition of the bacterial enzyme activity was observed, but the activity in the pDHFR 2.43 extract was reduced to only ~25%, consistent with observations that the mammalian enzyme is less prone to inhibition by trimethoprim[15]. Furthermore, changing the substrate to folate instead of dihydrofolate abolished the activity of the bacterial enzyme but the mam-

malian enzyme retained a low level of activity (Table 1 and ref. 16). These two observations confirm that the trimethoprim resistance of *M. methylotrophus* carrying pDHFR 2.43 is a consequence of the expression of the mouse dihydrofolate reductase gene.

As pGSS15 retains the broad host-range characteristics of R300B[8], we have been able to transfer pDHFR 2.43 into another methylotroph, *Methylomonas methylovora*, and into an autotrophic bacterium, *Alcaligenes eutrophus*, and have demonstrated that these strains are also rendered trimethoprim-resistant (our unpublished results).

Chicken ovalbumin, the major egg white protein, is a polypeptide consisting of 386 amino acids. The DNA sequence corresponding to the entire ovalbumin structural gene has been cloned[17,18] and ovalbumin-like protein has been detected in bacterial extracts by immunological methods[19,20].

A *Taq*I fragment of ~2,200 base pairs, containing the entire ovalbumin structural gene, can be excised from the plasmid pOV230 (ref. 19). The *Taq*I ovalbumin fragment was extended with oligo(dC) tails and annealed to pGSS15 DNA to which oligo(dG) sequences had been added after cleavage with *Pst*I. The annealed DNA was used to transform *E. coli* CL499 (*E. coli* CB100 P1 transduced to $rk^-mk^+$). $Ap^sTc^r$ transformants were screened for the presence of ovalbumin cDNA using a $^{32}P$-labelled probe generated by nick-translation[21] of the *Taq*I ovalbumin fragment. Several positive clones from the hybridization were screened for the production of ovalbumin using an immunoradiometric assay. The plasmid, R751 (ref. 22), was then used to mobilize the hybrid plasmids from ovalbumin-producing *E. coli* strains into *M. methylotrophus*. Table 2 shows the levels of ovalbumin expressed in *E. coli* and *M. methylotrophus*. Restriction enzyme mapping of the inserts of clones which did or did not express ovalbumin confirmed that expression was in one orientation only and was probably initiated at the β-lactamase promoter.

In the plasmid pOMP2 (ref. 20), ovalbumin sequences have been fused to part of the *E. coli* β-galactosidase gene. Expression of this hybrid plasmid yields a product in which the first four amino acids of ovalbumin are replaced by the first eight of β-galactosidase and the system is under the control of the strong *lac* UV5 (ref. 23) promoter. This system leads to high levels of expression in *E. coli* and it was of interest to determine whether similar high levels of expression could be achieved in *M. methylotrophus*. The fused β-galactosidase–ovalbumin sequence was isolated from pOMP2 as a single *Hha*I fragment of 2.7 kilobases. pGSS15 was cleaved with *Eco*RI and *Pst*I and the large fragment isolated. The *Hha*I *lac*–ovalbumin and *Eco*RI–*Pst*I pGSS15 fragments were then joined by oligo (dG.dC) tailing. (The construction of these hybrid plasmids is outlined in Fig. 1.) The annealed plasmids were used to transform *E. coli* CL499 and transformants were isolated and analysed in a similar manner to that used for the *Taq*I ovalbumin fragment. Table 2 shows the levels of expression of the *lac*–ovalbumin hybrid sequence in both *E. coli* and *M. methylotrophus*. Restriction enzyme analysis of the plasmid DNA from the hybrids expressing ovalbumin-like protein show that the inserted sequence is expressed in both orientations. However, one orientation (shown in Fig. 1*a*) is favoured (evident in 9 out of 10 isolates); the significance of this unknown. Expression in both orientations, together with the comparable levels of expression of the *lac*–ovalbumin fragment in pGSS15 and in the parental plasmid is evidence that the expression remains under control of the *lac* UV5 promoter.

The levels of immunoreactive ovalbumin detected in *E. coli* from both the β-lactamase and *lac* promoters are ~50-fold lower than expected. The expression of a penicillinase–proinsulin fused protein produced in similar conditions[24] was ~100 molecules per cell and expression of ovalbumin-like protein under *lac* control has been reported at 30,000 molecules per cell[19,20], compared with 2.8 and 760 from the present study. The most likely explanation of these differences concerns the assay systems and in particular the specificity, affinity and avidity of the antibodies used. We have tested this hypothesis using a

**Table 1** Dihydrofolate reductase activities of pDHFR 2.43 in extracts of *E. coli* ED 8654 and *M. methylotrophus*

| Strain | Substrate | Addition | Activity |
|---|---|---|---|
| *E. coli* ED 8654 | Dihydrofolate | — | 0.70±0.14 (4) |
| *E. coli* ED 8654 (pDHFR 2.43) | Dihydrofolate | — | 1.58±0.26 (5) |
| *M. methylotrophus* | Dihydrofolate | — | 2.34±0.68 (5) |
| *M. methylotrophus* | Dihydrofolate | Trimethoprim* | 0 (2) |
| *M. methylotrophus* | Folate† | — | 0 (2) |
| *M. methylotrophus* (pDHFR 2.43) | Dihydrofolate | — | 12.6±1.7 (4) |
| *M. methylotrophus* (pDHFR 2.43) | Dihydrofolate | Trimethoprim* | 2.57±0.53 (4) |
| *M. methylotrophus* (pDHFR 2.43) | Folate† | — | 0.50 (1) |

*E. coli* strains used were grown on medium containing (per l): $(NH_4)_2SO_4$, 2 g; $K_2HPO_4$, 14 g; $KH_2PO_4$, 6 g; trisodium citrate, 1 g; $MgSO_4.7H_2O$, 0.2 g (*p*H 7.2) supplemented with methionine (50 μg ml$^{-1}$) and glucose added to a final concentration of 0.2%. *M. methylotrophus* strains were grown on methanol minimal medium as described previously[12]. Cultures were grown with shaking at 37 °C overnight, collected by centrifugation and the pellets resuspended in 3 vol 50 mM potassium phosphate buffer (*p*H 7). The cells were disrupted by sonication on ice in the presence of 10 μM phenylmethylsulphonyl fluoride (PMSF) and 10 μM benzamidine and cell debris removed by centrifugation at 10,000*g* for 15 min and 100,000*g* for 60 min. Enzyme assays were performed at 37 °C. The standard assay mixture contained 50 μM dihydrofolate, 60 μM NADPH, 12 mM 2-mercaptoethanol, 50 mM potassium phosphate buffer *p*H 7 and enzyme (normally 100 μl of bacterial extract) in a final volume of 1.0 ml. The oxidation of NADPH was followed at 340 nm for ~10 min. Enzyme activity is expressed as nmol of NADPH oxidized per min per mg protein (mean ± s.e.m.). Numbers in parentheses are the number of separate experiments. Protein concentration was determined by the method of Lowry *et al.*[27].

* Trimethoprim (100 μg ml$^{-1}$) was added to the enzyme assays.

† Folate (70 μM) was included in the assay mixture instead of dihydrofolate.

**Table 2** Synthesis of ovalbumin-like protein in bacterial extracts

| Strain | Plasmid | Immunoreactive ovalbumin (molecules per cell) |
|---|---|---|
| *E. coli* CL499 | pOV230 | ND |
| *E. coli* CL499 | pGSS15: *Taq*-ov | 2.8 |
| *M. methylotrophus* | pGSS15: *Taq*-ov | 3.6 |
| *E. coli* CL499 | pOMP2 | 740 |
| *E. coli* CL499 | pGSS15: *lac*-ov | 760 |
| *M. methylotrophus* | pGSS15: *lac*-ov | 150 |

Cultures (5 ml) of *E. coli* and *M. methylotrophus* were grown with shaking at 37 °C overnight. *E. coli* cultures were grown in Luria broth containing 10 μg ml$^{-1}$ tetracycline and *M. methylotrophus* in methanol minimal medium[12] containing 1 μg ml$^{-1}$ tetracycline. Cells were collected by centrifugation and washed in 5 ml of phosphate-buffered saline (PBS). The cells were resuspended in 5 ml of ice-cold PBS and disrupted by sonication. Cell debris was removed by centrifugation at 3,000*g* for 10 min and the cell-free supernatant was used for the assay. Antibodies against ovalbumin were raised in rabbits by injecting into multiple sites 1 mg of ovalbumin (99%; Sigma) emulsified in Freund's complete adjuvant. Injections were repeated once a week for 4 weeks and blood was collected 10–14 days after the last injection. Antibodies were purified by $(NH_4)_2SO_4$ precipitation (15–40% saturation fraction) and dialysis. Polystyrene wells (7 mm diameter × 12 mm deep) were coated with anti-ovalbumin immunoglobin; 300 μl of antibody (10 μg ml$^{-1}$) in 0.2M $NaHCO_3$ (*p*H 9.2) was pipetted into each well and incubated at 37 °C for 5 h. Wells were washed three times with wash buffer (PBS containing 0.5% normal rabbit serum and 0.1% bovine serum albumin). In triplicate, cell extract (300 μl) was pipetted into wells and the released antigen was allowed to bind overnight at 4 °C. After incubation, wells were washed three times with wash buffer and 300 μl of $^{125}$I-anti-ovalbumin immunoglobulin in wash buffer was added (25 ng ml$^{-1}$ $^{125}$I-anti-ovalbumin, specific activity 18 μCi μg$^{-1}$). Wells were incubated at 37 °C for 3 h, washed three times and individual wells counted in a gamma counter. A standard curve was constructed using ovalbumin diluted in PBS or *E. coli* cell extract. As a measure of the immunoreactive ovalbumin-like material produced in *E. coli* and *M. methylotrophus*, the dilution of each strain giving 50% of the maximum binding was taken to be equivalent to the mass of authentic material giving the same counts. The viable cell count per ml of culture and the mass of ovalbumin detected were used to calculate the number of molecules of ovalbumin-like protein produced per cell. ND, not detected.

published radioimmunoassay[20]. The ovalbumin levels detected were similar to those of the immunoradiometric assay described above, indicating that the differences from other published results are probably due to the antisera used.

Although less than in *E. coli*, the level of expression of the two eukaryotic cDNAs in *M. methylotrophus* is significant, indicating that reasonable levels of expression may be achieved in the latter using promoter sequences previously investigated only in *E. coli*. It remains to be determined how closely related these *E. coli* promoter sequences are to the natural *M. methylotrophus* control sequences.

We thank Dr S. N. Cohen for providing pDHFR7, Dr B. W. O'Malley for pOV230, Dr O. Mercereau-Puijalon for pOMP2, Dr B. J. Bachman for CB100 and E. M. Pioli for CL499. We thank Drs M. S. Rose and E. C. Dart and Professor W. J. Brammar for their interest, support and helpful discussions.

Received 30 December 1981; accepted 18 March 1982.

1. Kupper, H. *et al. Nature* **289**, 555–559 (1981).
2. Yelverton, E. *et al. Nucleic Acids Res.* **9**, 731–741 (1981).
3. Keshet, E. *et al. Nucleic Acids Res.* **9**, 19–30 (1981).
4. Goeddel, D. V. *et al. Nature* **287**, 411–416 (1980).
5. Hardy, K., Stahl, S. & Kupper, M. *Nature* **293**, 431–483 (1981).
6. Hitzman, R. A. *et al. Nature* **293**, 719–722 (1981).
7. Meyer, R. J., Figurski, D. & Helinski, D. R. in *DNA Insertion Elements Plasmids and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S. L.) 559–566 (Cold Spring Harbor Laboratory, New York, 1977).
8. Grinter, N. J. & Barth, P. T. *J. Bact.* **128**, 394–400 (1976).
9. Barth, P. T., Tobin, L. & Sharpe, G. S. in *Molecular Biology, Pathogenicity and Ecology of Bacterial Plasmids* (eds Levy, S. B., Clowes, R. C. & Koenig, E. L.) 439–448 (Plenum, New York, 1981).
10. Itakura, K. *et al. Science* **198**, 1056–1063 (1977).
11. Pasek, M. *et al. Nature* **282**, 575–579 (1979).
12. Windass, J. D. *et al. Nature* **287**, 396–401 (1980).
13. Chang, A. C. Y. *et al. Nature* **275**, 617–624 (1978).
14. Meynell, E. & Datta, N. *Nature* **214**, 885–887 (1967).
15. Burchall, J. J. *Ann. N. Y. Acad. Sci.* **186**, 143–152 (1971).
16. Baccanari, D. *et al. Biochemistry* **14**, 5267–5273 (1975).
17. McReynolds, L. A. *et al. Nature* **273**, 723–728 (1978).
18. McReynolds, L. A., Caterall, J. F. & O'Malley, B. W. *Gene* **2**, 217–231 (1977).
19. Fraser, T. H. & Bruce, B. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5936–5940 (1978).
20. Mercereau-Puijalon, O. *et al. Nature* **275**, 505–510 (1978).
21. Rigby, P. J. W. *et al. J. molec. Biol.* **113**, 237–251 (1977).
22. Bennett, P. M. & Richmond, M. H. *J. Bact.* **126**, 1–6 (1976).
23. Backman, K. & Ptashne, M. *Cell* **13**, 65–71 (1978).
24. Villa-Komaroff, L. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3727–3731 (1978).
25. Tabati, H. F. & Flavell, R. A. *Nucleic Acids Res.* **5**, 2321–2332 (1978).
26. Roychoudhury, R., Jay, E. & Wu, R. *Nucleic Acids. Res.* **3**, 101–116 (1976).
27. Lowry, N. J., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–270 (1951).

*Proc. Natl. Acad. Sci. USA*
Vol. 79, pp. 4256–4259, July 1982
Biochemistry

# Expression of a chemically synthesized human $\alpha$1 interferon gene

[expression in *Escherichia coli* and *Methylophilus methylotrophus*/antiviral activity/cell specificity/stimulation of (2′-5′)oligoadenylate synthetase]

EDWARD DE MAEYER*, DANIAL SKUP*, KOLLI S. N. PRASAD*, JAQUELINE DE MAEYER-GUIGNARD*, BRYAN WILLIAMS†, PETER MEACOCK‡, GEOFF SHARPE‡, DAVE PIOLI§, JOHN HENNAM§, WOLFGANG SCHUCH§, AND KEITH ATHERTON§

*Institut Curie, Universite de Parid-Sud, Bât. 110, Campus d'Orsay, 91405 Orsay, France; †Division of Infectious Diseases, Hospital for Sick Children, Toronto, Ontario, Canada; ‡ICI/University Joint Laboratory, Biochemistry Department, University of Leicester, Leicester LE1 7RH, England; and §Biosciences Group, Corporate Research Laboratory, ICI Limited, The Heath, Runcorn WA7 4QE, Cheshire, England

*Communicated by Sydney Brenner, March 24, 1982*

ABSTRACT **Cells of *Escherichia coli* containing a chemically synthesized human $\alpha_1$ interferon (IFN-$\alpha_1$) gene, under control of the *lac* promoter, make a product with biological properties indistinguishable from those of the natural IFN-$\alpha_1$ [antiviral activity, acid stability, species crossreactivity, inactivation by antisera directed against leukocyte or Namalwa cell interferon, and stimulation of (2′-5′)oligoadenylate synthetase activity]. Similar levels of IFN synthesis were obtained when the expression unit (*lac* promoter plus synthetic IFN-$\alpha_1$ gene) was transplanted into the obligate methylotroph *Methylophilus methylotrophus*.**

Information on the biological properties of interferons (IFNs) has come predominantly from studies of material produced by cell culture techniques. Recent results with recombinant DNA techniques have shown that human leukocyte interferon (IFN-$\alpha$) is a family of at least eight different molecules (1). Furthermore, Allen and Fantes (2) have demonstrated that microheterogeneities in IFN preparations from Namalwa cells can be ascribed to the presence of at least five homologous proteins with different amino acid sequences and, hence, to different structural genes. Despite these advances, very little is known about the distinct physiological properties of individual human IFN species. Such work is now in progress (3) with cloned genes and may lead to the identification of species of IFN with preferred biological and pharmacological profiles for clinical use.

The availability of a number of cloned IFN genes also has facilitated the production of structural variants of IFN by the construction of hybrid genes. Cleavage at a common restriction site (4) within the coding sequence of two different IFN genes allows fusion of the $NH_2$-terminal sequence from one with the COOH-terminal sequence of the other. However, the number and type of molecules that can be synthesized by this approach is limited. In contrast, the recent report of the total synthesis of the IFN-$\alpha_1$ gene (5) provides a starting base for the construction of a large number of structural variants.

In this report we describe the expression of the synthetic gene in cells of *Escherichia coli*, yielding a product having properties characteristic of IFN-$\alpha_1$. Although *E. coli* is most commonly used for the expression of eukaryotic genes (6), the use of other organisms may provide advantages, especially with regard to industrial processes. For instance, work is in progress to develop alternative host/vector systems based on *Bacillus subtilis* (7) and *Saccharomyces cerevisiae* (8). We have demonstrated expression of a mouse dihydrofolate reductase cDNA and a chicken ovalbumin cDNA in the obligate methylotroph *Methylophilus methylotrophus* (9). We now report the expression of the synthetic IFN-$\alpha_1$ gene in *M. methylotrophus* when the expression unit (*lac* promoter plus synthetic gene) was introduced into this organism.

## MATERIALS AND METHODS

**DNA Construction.** The construction of the pIFS1–14 plasmids has been described (5). In the original constructions, the wild-type *lac* promoter sequence was used. The synthetic IFN-$\alpha_1$ gene has also been cloned in an identical manner into a pAT153 vector (pPM60) carrying the *lac UV5* promoter (10). This series of plasmids is referred to as pIFS101–106. In order to study the expression of the synthetic IFN-$\alpha_1$ gene in *M. methylotrophus*, the synthetic sequence fused to the *lac* promoter was recloned from pIFS1 into the plasmid pGSS15 (11) as shown in Fig. 1.

**Transfer of Plasmids to *M. methylotrophus*.** A wide host range-mobilizing plasmid, R751 (13), was introduced into *E. coli* carrying the pGSS15–sIFN hybrid plasmids. These were then used as donors in genetic crosses with *M. methylotrophus*, as described (14).

**Preparation of Bacterial Extracts.** *E. coli* strains were grown in Luria broth containing tetracycline (10 $\mu$g/ml) or ampicillin (25 $\mu$g/ml), and *M. methylotrophus* strains were grown in methanol minimal medium (14) with tetracycline (1 $\mu$g/ml). Overnight cultures were inoculated into fresh medium and grown to early stationary phase ($\approx 1 \times 10^9$ cells per ml). The cells were harvested and resuspended in 0.01 vol of 50 mM Tris·HCl, pH 8.0/50 mM EDTA/15% sucrose. Lysozyme was added to a final concentration of 2.5 mg/ml, and the suspension was kept on ice for 30 min before freezing at −20°C. The frozen lysate was thawed at room temperature and sonicated for 15 sec. Aliquots of the sonicated lysate were stored at −20°C or below until thawed for bioassay.

**IFN Bioassay.** Serial 1:2 dilutions were made in 96-well microtiter plates (Falcon plastics). T98G cells (a human neuroblastoma cell line) or MDBK (bovine) cells were added to each well, and the cultures were incubated overnight at 37°C in a $CO_2$ incubator. T98G and MDBK cultures, respectively, were challenged with encephalomyocarditis virus and vesicular stomatitis virus (0.5 plaque-forming unit per cell) and were scored for cytopathic effect 24–48 hr later. The titer of a sample (expressed as units/0.15 ml of bacterial extract) was defined as the reciprocal of the maximum dilution reducing the cytopathic effect by at least 50%. In this assay, the National Institutes of Health human reference standard G-023-901-527 (20,000 in-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: IFN, interferon; IFN-$\alpha$, human leukocyte IFN; IFN-$\beta$, fibroblast IFN; IPTG, isopropyl thiogalactoside.

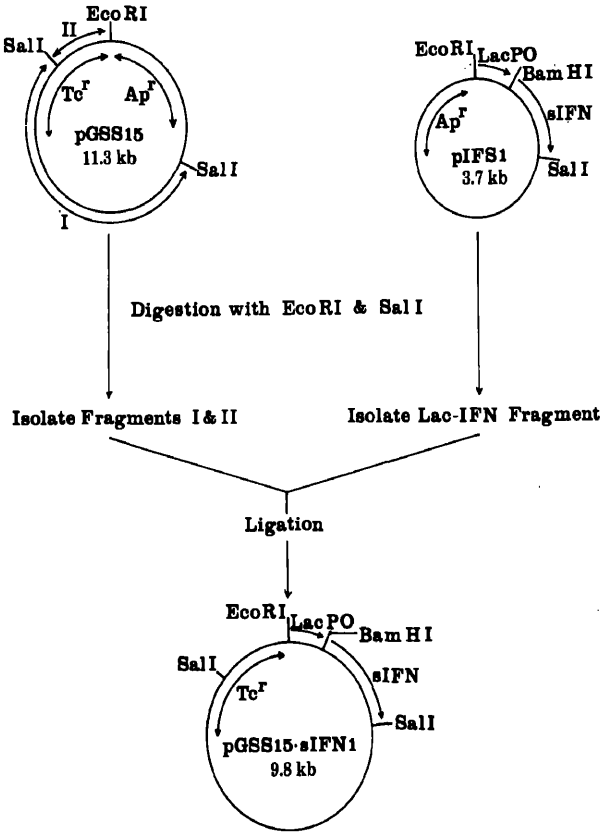



FIG. 1. Construction of pGSS15–sIFN plasmids. pGSS15 was digested to completion with *Eco*RI and *Sal* I, and the digestion products were separated by agarose gel electrophoresis. The *lac* promoter (*lac* PO)–synthetic gene (sIFN) fragment was obtained from an *Eco*RI/*Sal* I double digest of pIFS1 DNA and was ligated to fragments I and II of pGSS15 DNA. Recombinant plasmid transformants of *E. coli* were selected by the following criteria: tetracycline resistance ($Tc^r$); ampicillin sensitivity ($Ap^r$); a blue coloration in the presence of 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside, indicating the presence of the *lac* promoter/operator sequence (12); and hybridization to an oligonucleotide from within the synthetic gene sequence. kb, Kilobases.

ternational units/ml) was diluted 1:50 and, containing by definition 400 international units/ml, gave a titer of 120 units/0.15 ml in T98G cells. The MDBK cell line is ≈20-fold more sensitive than the T98G cells are for the IFN-$\alpha_1$ species (see Table 1 and ref. 15).

**Antiserum Neutralization.** This was performed in an IFN assay as described above. For each serum tested, 10 μl of the appropriate dilution was added to two rows of the microtiter plate.

**(2′-5′)Oligoadenylate Synthetase Activity.** Assay of (2′-5′)-oligoadenylate synthetase was performed as described (16).

## RESULTS

**IFN Activity Produced from the Synthetic Gene.** The construction of recombinant plasmids (pIFS1–14) containing a chemically synthesized DNA fragment that included a coding sequence for human IFN-$\alpha_1$ has been reported (5). In these plasmids the IFN-$\alpha_1$ gene was arranged to be under the transcriptional control of a *lac* promoter sequence and was positioned so that the ATG initiation codon formed a hybrid "ribosome-binding" site with the *lac* promoter "Shine–Dalgarno" (SD) sequence. Thus, when these plasmids are introduced into *E. coli* cells, the *lac* promoter should give both active transcription and translation of the IFN-$\alpha_1$ coding sequence.

In order to determine whether an IFN-like polypeptide is produced, we assayed extracts of *E. coli* cells containing the pIFS plasmids for the production of an antiviral activity. The results, presented in Table 1 for two of these clones, clearly show that a significant level of antiviral activity was detected in cells containing the recombinant plasmids. No such activity was found in cells containing only the vector plasmid pPM50 (pAT153 carrying the wild-type *lac* promoter).

Because the nucleotide sequence of the chemically synthesized IFN-$\alpha_1$ gene was altered from that of the natural gene, as defined by the sequence of the IFN-$\alpha_1$ cDNA (17), it was important to establish that the polypeptide conferring the antiviral activity detected had biological properties compatible with those of interferon produced in *E. coli* cells expressing IFN-$\alpha_1$ cDNA (15). Therefore, the following properties have been investigated.

*Stability at pH 2.* Stability at pH 2 was one of the first described characteristics of IFN-α and fibroblast IFN (IFN-β) (18), and we tested whether the synthetic IFN-$\alpha_1$ gene product displayed this property. For this purpose, 10 μl of the bacterial extract was diluted either in Sörensen buffer at pH 2 or in phosphate-buffered saline at pH 7.4 and kept at 4°C for 24 hr, after which time the antiviral activity was determined. Like natural interferon, the synthetic gene product was acid stable (Table 1).

*Neutralization of the antiviral activity.* The antiviral activity was further characterized in a neutralization assay in the presence of five different anti-IFN antisera. The activity of the product from the two clones tested was neutralized by the anti-Namalwa IFN antiserum and two different anti-IFN-α antisera but not by two different anti-IFN-β antisera (Table 1). This demonstrates that the antiviral activity could be correlated with the synthesis of an IFN-α species.

*Antiviral activity on bovine and murine cells.* The effect of various human IFNs on heterologous cells has been extensively documented (19). The bacterially derived synthetic gene product was assayed on human, bovine, and murine cells. The spectrum of activity is shown in Table 1 and corresponds in profile

Table 1. Characterization of antiviral activity extracted from *E. coli* containing pIFS plasmids*

| Plasmid | Stability to acid† | | Neutralization by antisera† | | | | Activity on cells of different species | | |
|---|---|---|---|---|---|---|---|---|---|
| | pH 7.4 | pH 2.0 | None | Anti-Namalwa IFN‡ | Anti-IFN-α§ | Anti-IFN-β¶ | Human T98G cells | Murine L cells | Bovine MDBK cells |
| pIFS1 | 520 | 1,200 | 1,200 | 75 | 75 | 1,200 | 1,200 | 600 | 25,600 |
| pIFS11 | 1,040 | 1,040 | 1,200 | 75 | 75 | 1,200 | — | — | — |

* Expressed as antiviral units/0.15 ml of cell extract. The three experiments summarized in this table were carried out independently.
† Assayed on T98G cells.
‡ Anti-Namalwa IFN antiserum was from K. Fantes (Wellcome, England).
§ Anti-IFN-α antisera were from K. Berg (Aarhus, Denmark) and E. Mogensen (Villejuif, France).
¶ Anti-IFN-β antisera were from J. Vilcek (New York) and A. Billiau (Louvain, Belgium). The two anti-IFN-α antisera behaved indistinguishably in the assays as did the two anti-IFN-β antisera.

to that reported for the IFN-$\alpha_1$ produced by *E. coli* containing IFN-$\alpha_1$ cDNA (15).

**Stimulation of (2′-5′)Oligoadenylate Synthetase Activity.** IFN treatment of cells has been shown to stimulate the activity of the enzyme (2′-5′)oligoadenylate synthetase (16). Therefore, we investigated if the synthetic gene product also showed this characteristic. For this experiment the interferon was first purified as follows.

The globulin fraction of an anti-IFN-$\alpha$ antiserum, obtained from K. Berg, was fixed to agarose through a spacer arm (Affigel 10, Bio-Rad). This material was then used as an antibody affinity column for the partial purification of the bacterially derived product; 1 ml of bacterial extract was diluted with 9 ml of phosphate-buffered saline (pH 7.4) and loaded onto a 0.4-ml column. After extensive washing, desorption was carried out with 0.1 M sodium citrate (pH 2). With this procedure, no antiviral activity was found in the flow through or wash fluid, whereas 48% of the original antiviral activity was recovered with the desorption buffer. A comparable result was obtained by incubating the bacterial extract (diluted 1:10 in phosphate-buffered saline) with the antibody-coated agarose beads overnight on a rotating wheel, followed by desorbing with 0.1 M sodium citrate (pH 2). T98G cells in microtiter wells were treated with this partially purified synthetic IFN gene product or with control lymphoblastoid IFN for 20 hr. Triplicate cultures were then pooled and assayed for (2′-5′)oligoadenylate synthetase activity. The synthetic gene product stimulated production of the enzyme activity (Table 2).

**Regulation of Expression of the IFN-$\alpha_1$ Gene.** The construction of the original recombinant plasmids pIFS1–14 utilized a vector pPM50 that contained a DNA fragment derived from the wild-type *lac* operon. Maximal expression from this promoter is dependent upon stimulation of the promoter–RNA polymerase interaction by the catabolite-activating protein (10). However, the 95-base-pair *Alu* I-generated *lac* promoter fragment present in pPM50 does not contain the complete recognition sequence for the catabolite-activating protein effector. Therefore, this promoter fragment can only function with reduced efficiency. The *UV5* mutation, which maps within the *lac* promoter, allows high-level expression without the need for catabolite-activating protein stimulation (10). Therefore, we transplanted the *Bam*HI–*Sal* I IFN-coding synthetic oligonucleotide to a second vector pPM60, which is identical to pPM50 in all respects except that the *lac* promoter fragment carries the *UV5* mutation. The relative positioning of the IFN-$\alpha_1$ gene and the *lac* promoter in these plasmids, pIFS101–106, was identical to that in the original (pIFS1–14) plasmids. Antiviral assay on extracts of cells containing these plasmids showed that use of the *lac* promoter carrying the *UV5* mutation gave a significant increase in the level of IFN gene expression (Table 3).

Expression of the IFN gene could be further regulated by the *lac* repressor. The copy number of the pIFS plasmids in *E. coli* is sufficiently high that they effectively titrate out the *lac* repressor, the net result being almost constitutive levels of expression from all *lac* promoters in the cell. In the case of the pIFS plasmids, high levels of IFN synthesis were obtained even in the absence of the *lac* operon inducer isopropyl thiogalactoside (IPTG; Table 3). However, when an F′*lac* episome carrying the *lac* $I^Q$ mutation [which results in overproduction of *lac* repressor (10)] was introduced into these cells by conjugal mating, the level of IFN synthesis from the pIFS plasmid was dramatically reduced. Addition of IPTG to the culture resulted in induction of the repressed *lac* promoters, restoring high-level IFN synthesis (Table 3).

Table 2. Stimulation of (2′-5′)oligoadenylate synthetase activity

| Sample | Dose, IFN units per well | Activity, pmol/hr per $A_{260}$ |
|---|---|---|
| None | — | 15 |
| pIFS1 | 16 | 40 |
| Namalwa IFN | 1 | 30 |
| | 3 | 37 |
| | 10 | 105 |
| | 30 | 170 |
| | 100 | 209 |
| | 300 | 378 |

Table 3. Regulation of IFN expression

| Plasmid | Promoter | IPTG induction* | Antiviral activity† |
|---|---|---|---|
| | Effect of *lac* promoter | | |
| pPM50 | wt | – | <300 |
| pIFS1 | wt | – | 19,000 |
| pPM60 | *UV5* | – | <300 |
| pIFS101 | *UV5* | – | 1,300,000 |
| | Effect of *lac* repressor | | |
| pIFS101 | *UV5* | – | 190,000 |
| pIFS101 | *UV5* | + | 140,000 |
| pIFS101 with F′*lacI*$^Q$ | *UV5* | – | 1,200 |
| pIFS101 with F′*lacI*$^Q$ | *UV5* | + | 380,000 |

wt, Wild type.
* IPTG was added (final concentration, 1 mM) during growth of the bacterial culture.
† Expressed as antiviral units/0.15 ml of bacterial extract as assayed on MDBK cells.

Thus, we can conclude that the human IFN-$\alpha_1$ expression is being regulated by the *lac* promoter/operator on these plasmids.

**IFN-$\alpha_1$ Expression in *M. methylotrophus*.** We have demonstrated that the *E. coli lac* promoter can be used to express cloned DNA sequences in the obligate methylotroph *M. methylotrophus* (9). Therefore, it was of interest to test whether the IFN expression unit (*lac* promoter plus synthetic gene) could be introduced into the organism such that it directed synthesis of IFN-$\alpha_1$ as in *E. coli*. Because the pIFS plasmids are unable to replicate in *M. methylotrophus*, it was necessary to transplant the entire expression unit on a *Eco*RI–*Sal* I fragment to a broad host-range vector plasmid pGSS15 (11) that is capable of stable maintenance in both organisms (Fig. 1).

IFN activity was detected in extracts of both *E. coli* and *M. methylotrophus* containing this recombinant hybrid plasmid pGSS15–sIFN1. Moreover, the level of expression was found to be equivalent in these two hosts (Table 4).

## DISCUSSION

We have demonstrated expression of a chemically synthesized human IFN-$\alpha_1$ gene in *E. coli*. The synthetic gene product has antiviral activity, is acid stable, and is neutralized by antisera prepared against Namalwa cell IFN or IFN-$\alpha$ but not by anti-IFN-$\beta$ antisera. Furthermore, the results of the affinity chro-

Table 4. IFN activity in *M. methylotrophus*

| Plasmid | IFN activity* in *E. coli* | IFN activity* in *M. methylotrophus* |
|---|---|---|
| pGSS15 | <300 | <300 |
| pGSS15–sIFN1 | 19,000 | 38,000 |

* Expressed as antiviral activity/0.15 ml of bacterial extract as assayed on MDBK cells.

matography on an anti-IFN-$\alpha$ antibody column indicate that the synthetic IFN-$\alpha_1$ gene product binds to antibodies prepared against the natural human IFN-$\alpha$. The product has ≈20-fold higher antiviral activity on bovine cells as compared to human cells and a significant activity on murine cells. These results show conclusively that the synthetic gene product has the biological activities previously reported for IFN and also correlates well with those properties already observed for the natural IFN-$\alpha_1$ gene product expressed in *E. coli* (15). It now will be possible to synthesize a number of structural variants of this gene and to evaluate the properties of their products against those described here.

The synthetic gene product also stimulates the production of (2′-5′)oligoadenylate synthetase activity. Our results indicate that there is a difference in the degree of enzyme stimulation by similar antiviral doses of bacterially produced IFN and a natural lymphoblastoid IFN preparation. This may reflect a difference in specific activity of the two preparations because the bacterial product is probably a single polypeptide whereas the lymphoblastoid IFN is a mixture of polypeptide species (2).

The production of interferon is under control of the *lac* promoter, the system being sensitive to *lac* repressor and inducible by IPTG. The ability to modulate synthesis as required may have a number of technical advantages—for example, in preferentially labeling IFN or in containing the production of IFN to a particular stage in the cells' growth cycle.

Synthesis of IFN from the synthetic gene also has been demonstrated in *M. methylotrophus*. This organism has many attractive features as a host. It is able to grow efficiently on the cheap substrates methanol and ammonia. Additionally, the toxicology of the organism has been well studied and, being an obligate methylotroph, it is unlikely to infect man.

We thank Jacqueline Hurst, Annie Cachard, Liliane Eusebe, and Simon Withe for able technical assistance; John Windass, Ed Dart, Tom McKillop, Mike Rose, Mike Edge, and Alex Markham for their interest and discussions; Kathy Zoon for the MDBK cell line; and Michael Tovey for the T98G cell line. Bryan Williams acknowledges the Canadian MRC for financial support.

1. Goeddel, D. V., Leung, D. W., Dull, T. J., Gross, M., Lawn, R. M., McCandliss, R., Seeburg, P. H., Ullrich, A., Yelverton, E. & Gray, P. W. (1981) *Nature (London)* **290,** 20–26.
2. Allen, G. & Fantes, K. H. (1980) *Nature (London)* **287,** 408–411.
3. Yelverton, E., Leung, D., Weck, P., Gray, P. W. & Goeddel, D. V. (1981) *Nucleic Acids Res.* **9,** 731–741.
4. Streuli, M., Hall, A., Boll, W., Stewart, W. E., II, Nagata, S. & Weissmann, C. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 2848–2852.
5. Edge, M. D., Greene, A. R., Heathcliffe, G. R., Meacock, P. A., Schuch, W., Scanlon, D. B., Atkinson, T. C., Newton, C. R. & Markham, A. F. (1981) *Nature (London)* **292,** 756–762.
6. Keshet, E., Rosner, A., Bernstein, Y., Gorrecki, M. & Aviv, H. (1981) *Nucleic Acids Res.* **9,** 19–30.
7. Hardy, K., Stahl, S. & Küpper, M. (1981) *Nature (London)* **293,** 481–483.
8. Hitzeman, R. A., Hagie, F. E., Levine, H. L., Goeddel, D. V., Ammerer, G. & Hall, B. D. (1981) *Nature (London)* **293,** 719–722.
9. Hennam, J. F., Cunningham, A. E., Sharpe, G. S. & Atherton, K. T. (1982) *Nature (London)* **297,** 80–82.
10. Miller, J. H. & Reznikoff, W. S. (1978) *The Operon* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
11. Barth, P. T., Tobin, L. & Sharpe, G. S. (1981) in *Molecular Biology, Pathogenicity and Ecology of Bacterial Plasmids,* eds. Levy, S. B., Clowes, R. C. & Keonig, E. L. (Plenum, New York), pp. 439–448.
12. Marians, K. J., Wu, R., Stavinsky, J., Hozumi, T. & Narang, S. A. (1976) *Nature (London)* **263,** 744–748.
13. Bennet, P. M. & Richmond, M. H. (1976) *J. Bacteriol.* **126,** 1–6.
14. Windass, J. D., Worsey, M. J., Pioli, E. M., Pioli, D., Barth, P. T., Atherton, K. T., Dart, E. C., Byrom, D., Powell, K. & Senior, P. J. (1980) *Nature (London)* **287,** 396–401.
15. Stewart, W. E., II, Sarkar, F. H., Taira, H., Hall, A., Nagata, S. & Weissmann, C. (1980) *Gene* **11,** 181–186.
16. Williams, B. R. G. & Read, F. E. (1981) in *The Biology of the Interferon System,* eds. De Maeyer, E., Galasso, G. & Schellekens, H. (Elsevier/North-Holland), pp. 111–114.
17. Mantei, N., Schwarzstein, M., Streuli, M., Panem, S., Nagata, S. & Weissmann, C. (1980) *Gene* **10,** 1–10.
18. Isaacs, A., Lindenmann, J. & Valentine, R. C. J. (1957) *Proc. R. Soc. London Ser. B* **147,** 268–273.
19. Stewart, W. E., II (1979) *The Interferon System* (Springer, Vienna).

*Gene*, 29 (1984) 93–102
Elsevier



GENE 1021

# Broad host range cloning vectors for Gram-negative bacteria

(Recombinant DNA; plasmids; R300B; restriction analysis; multiple cloning sites; insertional inactivation; *Escherichia coli*)

**Geoffrey S. Sharpe**

*I.C.I./University Joint Laboratory, School of Biological Sciences, University of Leicester, Leicester LE1 7RH (U.K.) Tel. (0533) 551234 (Ext. 395)*

(Received February 8th, 1984)
(Revision received and accepted March 16th, 1984)

SUMMARY

A series of cloning vectors has been constructed based on the broad-host-range plasmid R300B. One of these vectors, pGSS33, has a size of 13.4 kb and carries four antibiotic resistance genes [ampicillin ($Ap^r$), chloramphenicol ($Cm^r$), streptomycin ($Sm^r$) and tetracycline ($Tc^r$)], all of which have restriction sites for insertional inactivation. The derivation, structure and uses of the plasmids are described.

## INTRODUCTION

The IncQ plasmids are very commonly found (Smith et al., 1974) over a broad geographical and species range (Barth and Grinter, 1974). Important members of this group include R300B (Barth and Grinter, 1974), NTP2 (Smith et al., 1974), RSF1010 (Guerry et al., 1974), which is identical to NTP2, (DeGraaff et al., 1978) and R1162 (Meyer et al., 1979); these appear indistinguishable at present (Barth and Grinter, 1974; Heffron et al., 1975).

The host range of the IncQ group plasmids, as represented by R300B, includes the majority of Gram-negative bacteria (Barth et al., 1981) making them important for developing broad-host-range cloning vectors.

In addition to its host range, several other properties of R300B make it suitable as the starting point for development of a cloning vector; it is relatively small (8.68 kb; Meyer et al., 1982), multicopy (9–12 copies per *Escherichia coli* genome; Barth and Grinter, 1974) and nonconjugative, although efficiently mobilized by a variety of conjugative plasmids (Grinter and Barth, 1976). In common with most other broad-host-range plasmids, there are relatively few target sites for cleavage by restriction enzymes (Fig. 1) (Barth et al., 1981; Meyer et al., 1982), with only *Sst*I and *Sst*II sites in the $Sm^r$ gene, and the *Pst*I sites in the $Su^r$ gene, giving insertional inactivation.

In contrast, the laboratory-constructed *E. coli* plasmids pBR322 (Bolivar et al., 1977) and pBR328 (Soberon et al., 1980) have many more useful restriction sites, and several workers have made hybrid plasmids from various combinations of pBR322 and an IncQ plasmid (Bagdasarian et al., 1979; 1981; Gautier and Bonewald, 1980; Wood et al., 1981).

Abbreviations: Ap, ampicillin; bp, base pairs; Cm, chloramphenicol; kb, kilobase pairs; Sm, streptomycin; Su, sulfonamide; Tc, tetracycline; superscripts $^r$ and $^s$, resistance and sensitivity, resp., to specified antibiotic.

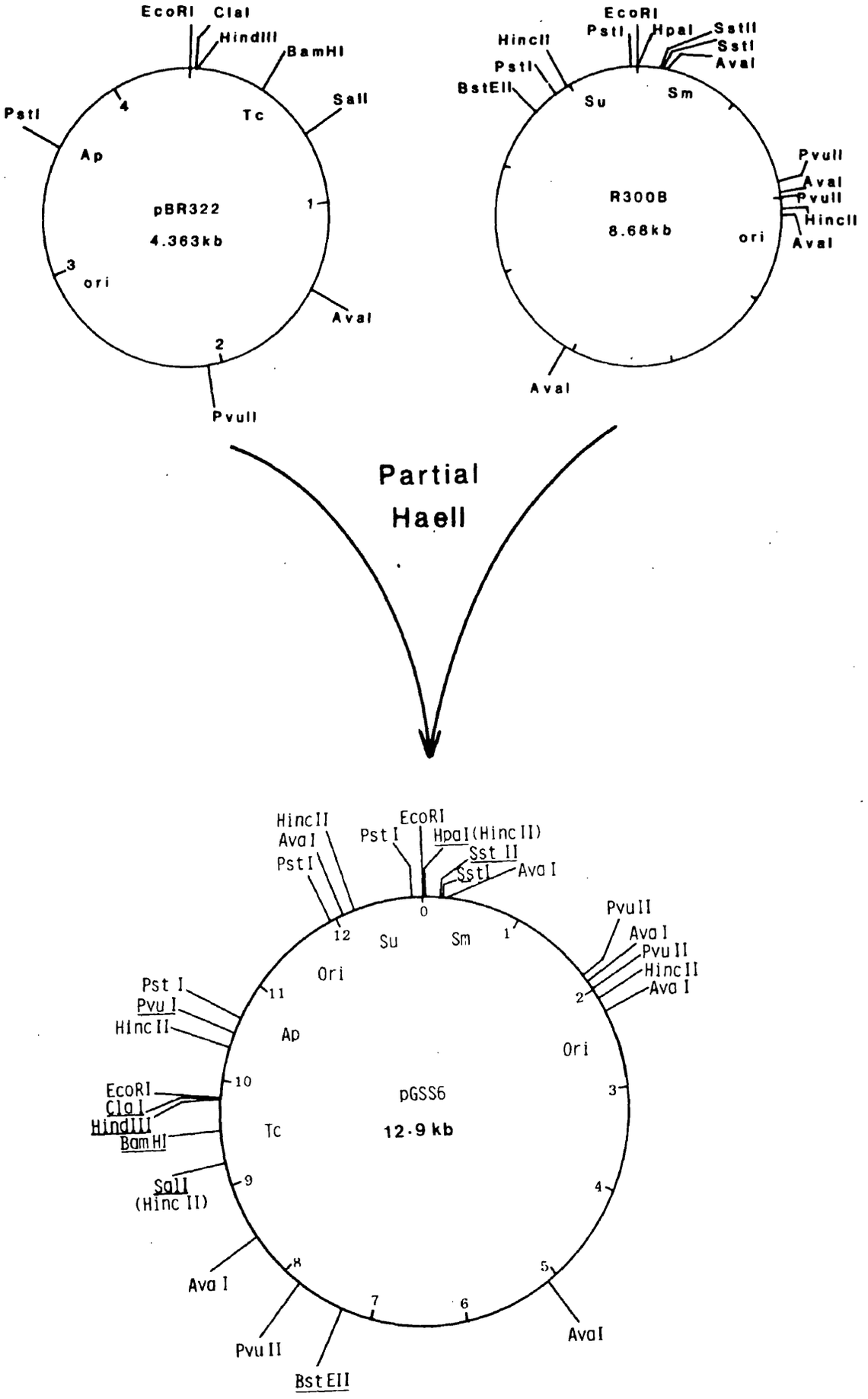



Fig. 1. Construction of a total co-integrate plasmid, pGSS6, made from pBR322 and R300B following ligation of full length linear molecules obtained by partial *Hae*II digestion. In pGSS6 unique restriction sites are shown underlined.

However, there have been reports of instability (Gautier and Bonewald, 1980) or reduced host range (Bagdasarian et al., 1979) of such chimeric plasmids.

This paper describes the development of a series of plasmid-cloning vectors that combines the broad host range of R300B with increased cloning capability in sequences derived from pBR322 and pBR328. Some of these vectors, especially pGSS15, have been used to demonstrate expression of heterologous coding sequences in organisms other than *E. coli* (Hennam et al., 1982; DeMaeyer et al., 1982), but this is the first full description of their derivation and structure. Features of pGSS15, which could occasionally cause problems, have been identified and this has led to the development of a new vector, pGSS33. The pGSS33 vector is a mobilizable, 13.4-kb plasmid carrying genes for resistance to four antibiotics which each have at least one site for insertional inactivation.

## MATERIALS AND METHODS

### (a) Plasmids, bacterial strains and their growth conditions

*E. coli* strains NM36 (*polA*1, *sup*°) and ED8654 (*met*, *hsdR*$^-$, *supE*, *supF*) were obtained from W.J. Brammar, C600 (*F*$^-$, *thr*1, *leuB*6, *thi*1, *glnV*, *lacY*1, *tonA*21) from B. Wilkins, and JA221 (*leuB*6, *trpE*5, *recA*, *lacY*1, *hsdR*$^-$) (Clarke and Carbon, 1978), from J.D. Windass.

*Methylophilus methylotrophus* strain AS1 (wild type) was obtained from D. Byrom and CMB13 (a non-restricting derivative) from C. Boyd (1983).

*E. coli* cells were grown in L-broth or on L-agar plates, in M9 medium (Miller, 1972) supplemented with amino acids as appropriate at 40 μg/ml, thiamine at 4 μg/ml, and thymine at 20 μg/ml, or on minimal plates using the above medium solidified with 15 mg/ml of agar.

*M. methylotrophus* strains were grown in media as previously described (Windass et al., 1980).

Antibiotics were used at the following concentrations; Ap 50–100, Cm or Sm 25, Su 500–1000, and Tc 10 μg/ml.

Plasmids R300B and RP4 were obtained from P.T. Barth, and R64*drd*11 from B. Wilkins.

### (b) Plasmid manipulation and analysis

All enzymes were obtained from commercial sources and used in accordance with the manufacturer's instructions.

Small-scale plasmid analysis was by the rapid boiling method (Holmes and Quigley, 1981).

Plasmid copy-number determinations were by the method of Womble et al. (1977) using Brij 58/deoxycholate for cell lysis followed by treatment with RNase and pronase, then shearing.

Mobilization of plasmids was exactly as described previously (Barth et al., 1981).

## RESULTS

### (a) Construction of a co-integrate between pBR322 and R300B

Introduction of the antibiotic-resistance genes from pBR322 into R300B to expand its cloning capability was approached by ligation of partial *Hae*II digests of the two plasmid DNAs. There are eleven sites for this enzyme in pBR322 (Fig. 2) and about 30 in R300B. As it was not known how much of R300B was required to retain its broad-host-range characteristics, digestion conditions were chosen which produced the highest proportion of full-length linear molecules as observed by agarose gel electrophoresis. [It has been reported by Bagdasarian et al. (1981), that only about 25% of the RSF1010 genome, principally the segment of antibiotic-resis-

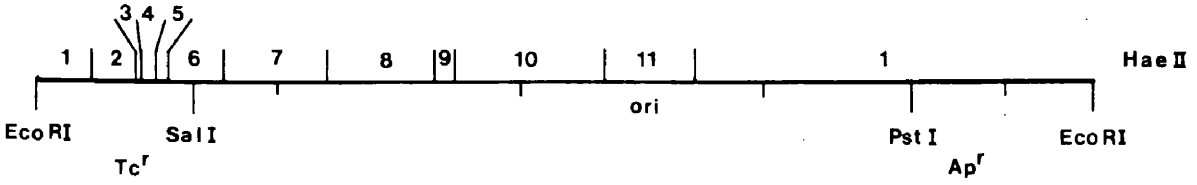



Fig. 2. *Hae*II restriction map of pBR322. The plasmid is shown linearized at its *Eco*RI site, which lies within the large *Hae*II fragment 1.

tance genes, could be removed without deleterious effects.]

Following ligation of the plasmid DNAs from the different digests of pBR322 and R300B in various combinations, transformation of *E. coli* strain ED8654 yielded three Ap$^r$, Tc$^r$, Sm$^r$, Su$^r$ colonies. One of these contained a plasmid, designated pGSS6, with a size of approx. 13 kb (consistent with its being a total co-integrate of both parental plasmids).

Plasmid pGSS6 had two *Eco*RI sites producing fragments of approx. 10 kb and 3 kb. The smaller fragment contained three *Pst*I sites while the larger fragment had single *Sst*I, *Sal*I and *Bst*EII sites. This suggested that the Ap$^r$ and Su$^r$ genes were located together on the small *Eco*RI fragment and the Sm$^r$ and Tc$^r$ genes were on the larger fragment (Fig. 1). The plasmid could transform and be maintained in *E. coli* strain NM36 [which lacks the DNA polymerase I activity, essential for replication from a ColE1 origin (Kingsbury and Helinski, 1970)], and could be mobilized into *M. methylotrophus* by the IncP plasmid RP4, both characteristics inherited from the parental R300B plasmid (Barth et al., 1981). Thus, while pGSS6 possesses features required for a broad-host-range vector, it also has potential as a starting point for development of new vectors.

After digestion of pGSS6 with *Eco*RI and transformation of *E. coli* strain ED8654 with the self-ligated fragments, two plasmid species were isolated, pGSS8 (10.1 kb, Tc$^r$ and Sm$^r$) and pGSS9 (2.94 kb, Ap$^r$ and Su$^r$) (Fig. 3). Maintenance of pGSS9 was found to be DNA polymerase I-dependent and the plasmid was not mobilizable into e.g. *M. methylotrophus*, whereas pGSS8 was DNA polymerase I-independent and had the broad host range of R300B. This result indicated that both parental replication origins were present and functional on pGSS6 and that they have been separated to produce two smaller and very different plasmids, pGSS8 and pGSS9.

### (b) Structure of pGSS8

The unique *Eco*RI site of pGSS8 at the junction of pBR322 and R300B-derived sequences lies between the Tc$^r$ and Sm$^r$ genes. The fragment carrying the Tc$^r$ gene from pBR322 has introduced unique *Cla*I, *Hin*dIII, *Bam*HI and *Sal*I sites together with additional *Ava*I and *Pvu*II sites, and has significantly expanded the cloning capability of R300B. An *Eco*RI-partial *Hae*II fragment containing these sites would need to be 2353 bp long and would comprise all the *Hae*II fragments except No. 11 (containing the pBR322 *ori*) and most of No. 1 (containing the Ap$^r$ gene) (Fig. 2). Since pGSS8 contains a *Bst*EII site but no *Pst*I sites, the second junction between pBR322 and R300B sequences must lie at a *Hae*II site located within the 270-bp region between the *Bst*EII and *Pst*I sites of R300B. Confirmation of this structure for pGSS8 came from electron microscopy of heteroduplex DNA molecules (Fig. 4, a and b).

The Su$^r$ and Sm$^r$ genes of RSF1010 comprise a single transcriptional unit (Heffron et al., 1975; Rubens et al., 1976) with the Sm$^r$ gene distal to the promoter, and a similar arrangement has been confirmed for R300B (Barth et al., 1981). Construction of pGSS8 and pGSS9 has shown that these genes, in R300B, can be separated at the *Eco*RI site without inactivation, but that in pGSS8 transcription of the Sm$^r$ gene must be dependent on a promoter located near the start of the Tc$^r$ gene (Stüber and Bujard, 1981). This may cause a problem when cloning into the *Hin*dIII site, for example, as both selectable markers may be inactivated. To provide an alternative genetic selection, as well as extra cloning sites, the Ap$^r$ gene from pBR322 was cloned into pGSS8 to give pGSS15.

### (c) Structure of pGSS15

Transformation of *E. coli* strain NM36 with a ligated mixture of partially *Hae*II-digested pGSS8 DNA and totally *Hae*II-digested pBR322 DNA, led to the isolation of an Ap$^r$, Tc$^r$, Sm$^s$ colony containing a plasmid which was designated pGSS15. The *Hae*II fragment 1 from pBR322, containing the Ap$^r$ gene, is 1867 bp long and includes *Eco*RI, *Cla*I and *Hin*dIII sites. However, upon digestion with these three enzymes, pGSS15 produced only linear bands of 11.5 kb. There was no longer a site for *Sst*I, but there were now two sites for *Sal*I approximately equidistant from a *Pst*I site, and the small *Pst*I-*Sal*I fragments from pGSS15 appeared to be the same size as the small *Pst*I-*Sal*I fragment of pBR322. *Hae*II digestion of pGSS15 followed by electrophoresis on a 5% polyacrylamide gel, further revealed the presence of the pBR322 fragment 11 and an

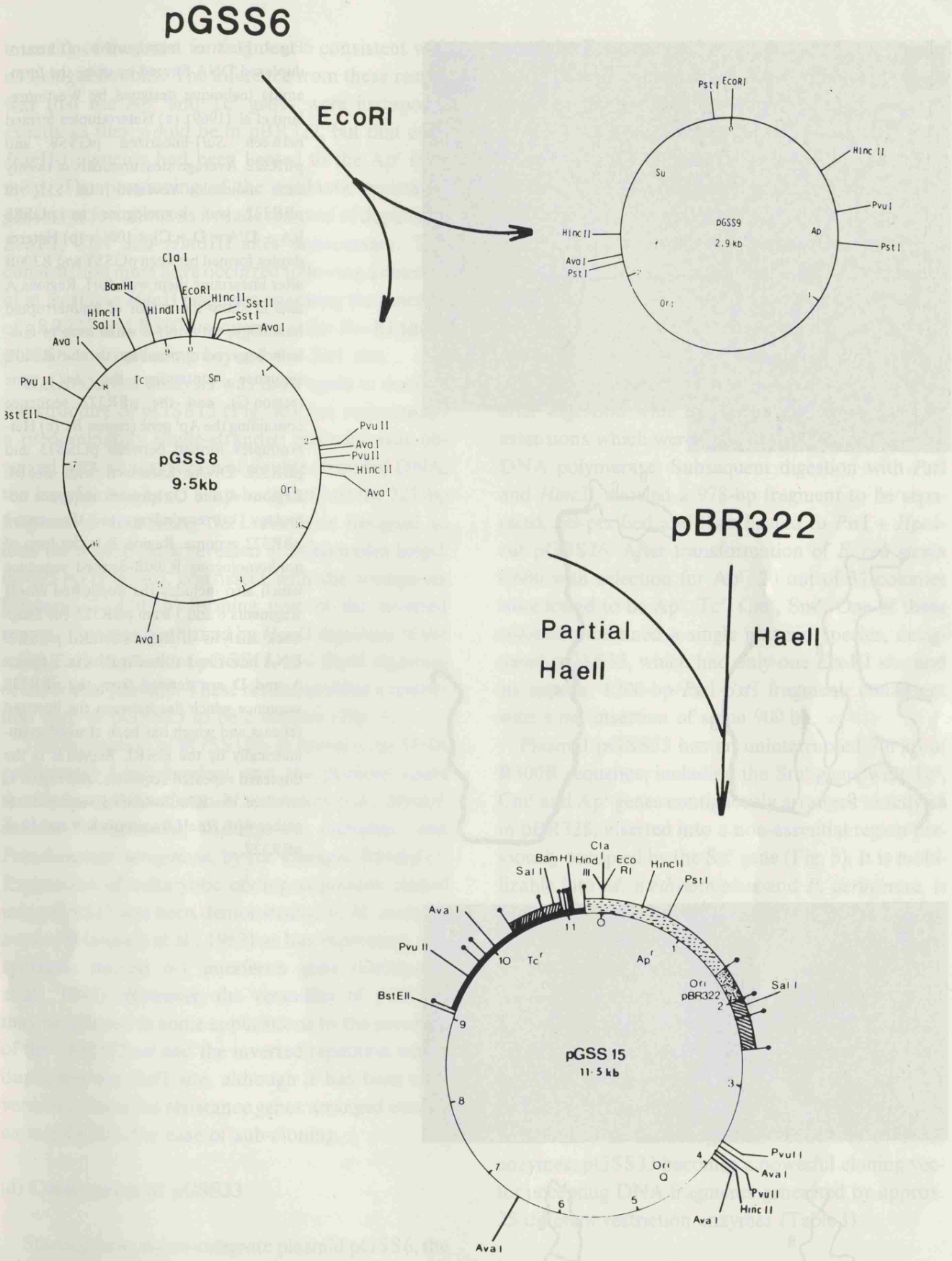



Fig. 3. Construction of plasmid pGSS15. Plasmids pGSS8 and pGSS9 were constructed by *Eco*RI digestion of pGSS6 and self-ligation of the two fragments. Plasmid pGSS15 was derived by ligation of partially *Hae*II-digested pGSS8 and *Hae*II-digested pBR322. The *Hae*II sites, marked as black dots, are only shown for the pBR322-derived sequence of pGSS15. The heavy black line represents pBR322 sequence originating from pGSS8 and the broadened portion of the plasmid shows the extent of newly acquired *Hae*II fragments from pBR322; number 1, containing the $Ap^r$ gene (horizontal hatching); number 11, containing *ori* (stippled); and numbers 6 and 7, forming an inverted repeat (diagonal hatching).

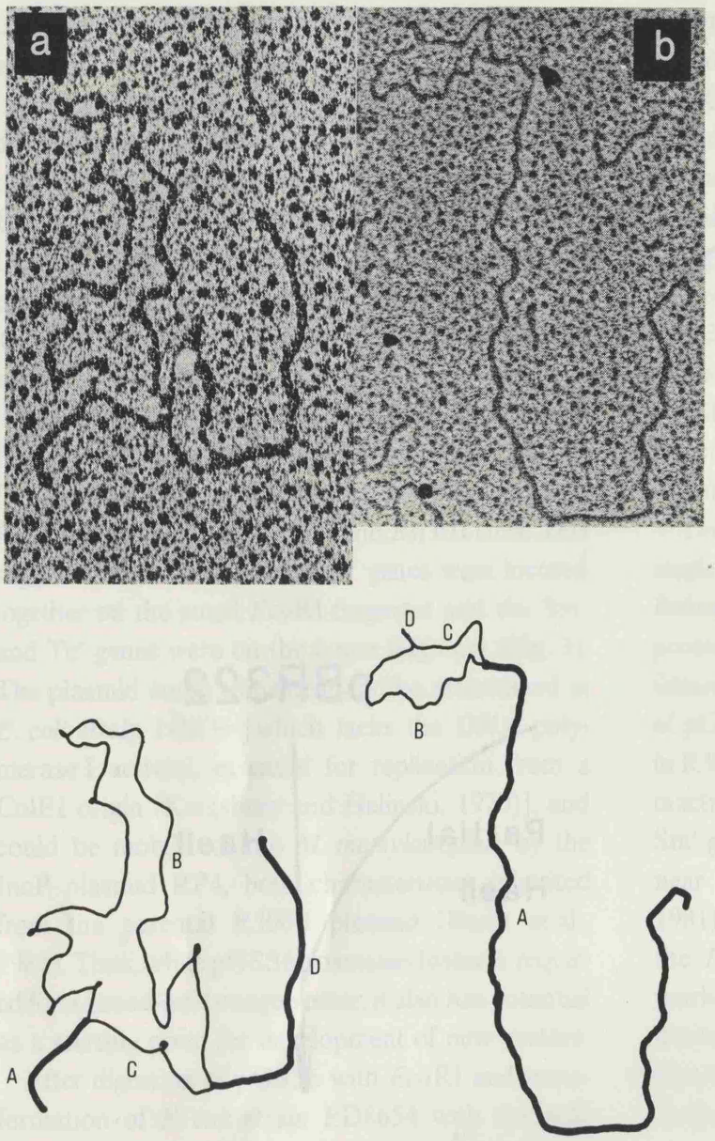



Fig. 4. Electron micrographs of heteroduplexed DNA formed by using the formamide technique designed by Westmoreland et al. (1969). (a) Heteroduplex formed between *Sal*I-linearized pGSS8 and pBR322. Average measurements of twenty such molecules revealed that 53% of pBR322 was homologous to pGSS8 [(A + D/A + D + C) × 100]. (b) Heteroduplex formed between pGSS8 and R300B after linearizing them with *Sst*I. Regions A and D show the extent of uninterrupted homology, with only a small loop of nonhomology corresponding to the R300B sequence containing the $Su^r$ gene (region C), and the pBR322 sequence containing the $Ap^r$ gene (region B). (c) Heteroduplex formed between pGSS15 and pBR322 DNAs linearized with *Eco*RI. Regions A and C represent sequence homology corresponding to the entire pBR322 genome. Region B is the loop of nonhomologous R300B-derived sequence which also includes the duplicated *Hae*II fragments 6 and 7 from pBR322. (d) Snapback analysis of single-stranded pGSS15 DNA linearized at its *Eco*RI site. Regions A and D are derived from the pBR322 sequence which lies between the inverted repeats and which has been cleaved asymmetrically by the *Eco*RI. Region B is the duplexed repeated sequence, and region C is all of the R300B-derived sequence together with *Hae*II fragments 8, 9 and 10 of pBR322.

intensity of the band for fragment 6 consistent with its being a doublet. The inference from these results was that the $Ap^r$ and $Tc^r$ genes were juxtaposed exactly as they would be in pBR322, but that extra *Hae*II fragments had been ligated to the $Ap^r$ fragment. The positioning of the antibiotic resistance genes was fortuitous as it made removal of duplicated *Eco*RI, *Cla*I and *Hin*dIII sites unnecessary. This construction must have occurred following a deletion of at least one *Hae*II fragment spanning the junction of pBR322 and R300B sequence at the *Eco*RI site in pGSS8 and extending beyond the *Sst*I site.

Heteroduplex analysis was used again to confirm the structure of pGSS15 (Fig. 4c), but occasionally a predominantly single-stranded molecule was observed with a short region of double-stranded DNA, too long for it to be a "snap-back" of just the 227-bp fragment 6 containing *Sal*I. Analysis designed to look for a snap-back revealed a homoduplex length of 655 bp (Fig. 4d), consistent with the contiguous fragment 7 (430 bp) forming part of the inverted repeat. This was confirmed by *Hae*II digestion of the small *Sal*I fragment of pGSS15, and *Hin*fI digestion of the whole plasmid. These results enabled a restriction map of pGSS15 to be compiled (Fig. 3).

The replication of pGSS15 was found to be DNA polymerase I-independent and the plasmid could readily be mobilized into *M. methylotrophus*, *Methylomonas methylovora*, *Alcaligenes eutrophus* and *Pseudomonas aeruginosa*, by for example R64*drd*11. Expression of eukaryotic coding sequences cloned into pGSS15 has been demonstrated in *M. methylotrophus* (Hennam et al., 1982) as has expression of a synthetic human α-1 interferon gene (DeMaeyer et al., 1982). However, the versatility of pGSS15 may be affected in some applications by the presence of the pBR322 *ori* and the inverted repetition which duplicates the *Sal*I site, although it has been convenient to have the resistance genes arranged exactly as in pBR322, for ease of sub-cloning.

### (d) Construction of pGSS33

Starting from the co-integrate plasmid pGSS6, the pBR322 *ori* and the $Su^r$ gene were deleted after digestion with *Bam*HI + *Pst*I, and replaced with the *Bam*HI-*Pst*I fragment from pBR328, which carries the $Cm^r$ gene, to give a $Cm^r$, $Tc^r$, $Sm^r$, $Ap^s$ plasmid, pGSS26 (Fig. 5). The incoming fragment reconstituted the $Tc^r$ gene and re-introduced part of the $Ap^r$ gene, containing the β-lactamase promoter, which was then responsible for transcription of the $Sm^r$ gene. Plasmid pGSS26 possessed two *Eco*RI sites, one within the $Cm^r$ gene (which would be useful for cloning), and one derived from R300B with a *Hpa*I site immediately adjacent (Barth, 1979; Barth et al., 1981). Thus, insertion into pGSS26, between *Pst*I and *Hpa*I, of a fragment from pBR322 extending from the *Pst*I site to beyond the end of the β-lactamase gene and terminating in a blunt end, would restore $Ap^r$ and remove the second *Eco*RI site.

A 1857-bp fragment was isolated from pBR322 after digestion with *Bst*NI, having single-base 5′-extensions which were filled in using large-fragment DNA polymerase. Subsequent digestion with *Pst*I and *Hin*cII allowed a 978-bp fragment to be separated, gel-purified and then ligated to *Pst*I + *Hpa*I-cut pGSS26. After transformation of *E. coli* strain C600 with selection for $Ap^r$, 20 out of 31 colonies were found to be $Ap^r$, $Tc^r$, $Cm^r$, $Sm^r$. One of these colonies contained a single plasmid species, designated pGSS33, which had only one *Eco*RI site and an approx. 1200-bp *Pst*I-*Sst*I fragment, consistent with a net insertion of some 900 bp.

Plasmid pGSS33 has an uninterrupted 7.6 kb of R300B sequence, including the $Sm^r$ gene, with $Tc^r$, $Cm^r$ and $Ap^r$ genes contiguously arranged exactly as in pBR328, inserted into a non-essential region previously occupied by the $Su^r$ gene (Fig. 5). It is mobilizable into *M. methylotrophus* and *P. aeruginosa*, is DNA polymerase I-independent, and exhibits Q-type incompatibility. The plasmid was found to be multicopy, 20 copies per chromosome equivalent when measured in exponentially grown *E. coli* cells, but was not amplifiable on treatment with spectinomycin. The antibiotic resistance genes all carry unique restriction sites that will give insertional inactivation by cloned fragments, and because there is redundancy in the recognition sequences of some enzymes, pGSS33 becomes a powerful cloning vector accepting DNA fragments generated by approx. 25 different restriction enzymes (Table I).

## DISCUSSION

The use of partial digests to construct the prototype co-integrate plasmid, pGSS6, followed by ge-

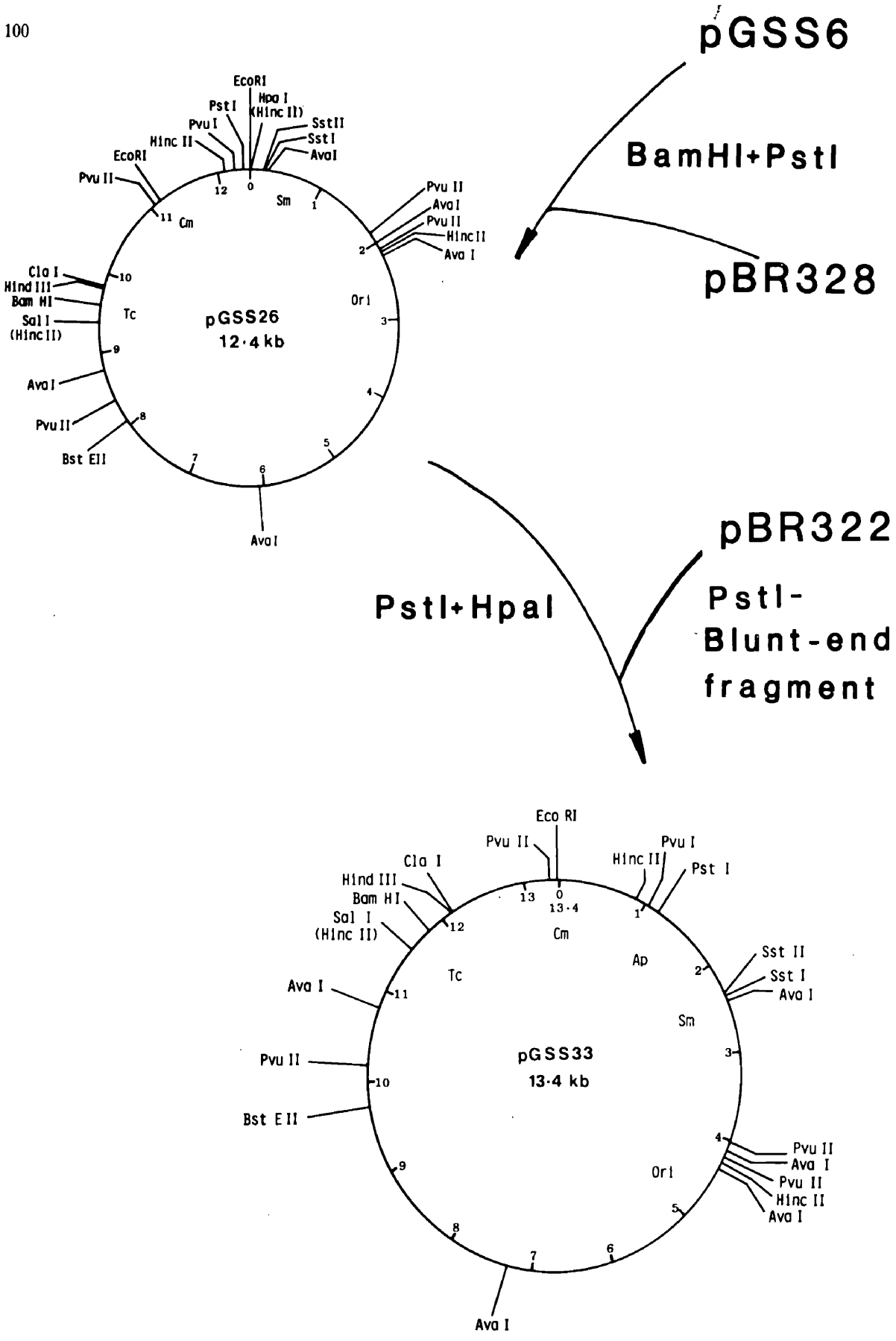



Fig. 5. Construction of pGSS26 and pGSS33. Plasmid pGSS26 was isolated from the ligation of *Bam*HI + *Pst*I-digested pGSS6 and pBR328 DNAs after transformation of *E. coli* strain C600 with selection for $Cm^r$ coupled with $Tc^r$, $Sm^r$ and $Ap^s$. In the construction of pGSS33 from pGSS26 the extra *Eco*RI site has been removed and the $Ap^r$ gene has been re-constituted. The arrangement of the $Ap^r$, $Cm^r$ and $Tc^r$ genes in pGSS33 is analogous to that of pBR328. For pGSS6 see Fig. 1.

TABLE I

Restriction enzymes which can be used for cloning DNA fragments into pGSS33

| Plasmid cut with restriction endonuclease | Will accept fragments generated by restriction endonuclease | Inactivating |
|---|---|---|
| *Bam*HI | *Bam*HI, *Bcl*I, *Bgl*II, *Mbo*I, *Sau*3A, *Xho*II | Tc |
| *Cla*I | *Cla*I, *Acy*I, *Asu*II, *Hin*P$_1$I, *Hpa*II, *Msp*I, *Nar*I, *Taq*I | (Tc)[a] |
| *Eco*RI | *Eco*RI, *Eco*RI* | Cm |
| *Hin*dIII | *Hin*dIII | (Tc)[a] |
| *Pst*I | *Pst*I | Ap |
| *Pvu*I | *Pvu*I, (*Rsh*I, *Xor*II)[b] | Ap |
| *Sal*I | *Sal*I, *Xho*I, *Ava*I | Tc |
| *Sst*I | *Sst*I, *Hgi*AI, *Hgi*JII, (*Sac*I)[b] | Sm |
| *Sst*II | *Sst*II, (*Sac*II)[b] | Sm |
| *Bst*EII | *Bst*EII | – |

[a] Insertion at these sites does not always cause inactivation.

[b] Isoschizomers.

netic selection, may have led to the isolation of a favourably stable arrangement of the parental plasmid sequences. Certainly, the plasmids in this series appear very stable, as loss of plasmid could not be detected when JA221 plasmid-carrying cultures were maintained in exponential phase for twenty generations without antibiotic selection (G.S.S., unpublished results). Also, all insertions affecting sections of R300B have been into regions inessential for replication, mobilization or maintenance (Bagdasarian et al., 1981; Meyer et al., 1982) and a region essential for an extended host range beyond *E. coli* (Barth et al., 1981) is also intact, suggesting that the host range for these vectors is unlikely to be altered from that of R300B.

The ability of these vectors to be mobilized by a variety of different conjugative plasmids means that many more organisms (besides *E. coli*) can now be used as hosts for cloning experiments. Industrially important organisms, for which there may be no transformation procedures, such as *M. methylotrophus*, become amenable to genetic and physiological modification. The extended host range provides the potential to study expression of cloned coding sequences in several different organisms while eliminating the need for further manipulation of the cloned genes.

The mobilization of these broad host range vectors by IncIα plasmids is interesting and may be more widely applicable. Two instances of the transfer of non-mobilizable plasmids by I-like R factors, to *Providentia* and *Proteus mirabilis* (Coetzee, 1964; Datta and Hedges, 1972) have been reported, in which the I factor was not detected in the recipients. In four cases reported here, Iα plasmids were able to promote transfer of IncQ plasmids into organisms outside the normal Iα host range, but again the conjugative plasmids themselves were never detected in the recipients (G.S.S., unpublished results). Thus Iα plasmids appear to be capable of achieving conjugal transfer to many organisms, but only capable of stable maintenance in a few. For the purpose of transferring the pGSS series of vectors however, this is a positive advantage, providing a clean delivery of vector plasmid with a degree of containment, since subsequent mobilization would require introduction of another conjugative plasmid.

This paper describes the development and analysis of a series of cloning vectors based on the broad host range IncQ plasmid R300B. Of the series, pGSS33 in particular, carries many of the most useful cloning sites of pBR328 together with *Sst*I and *Sst*II sites derived from R300B, providing alternatives to *Pst*I as sites useful for insertion of homopolymer-tailed fragments. Having four antibiotic-resistance genes, all with restriction sites capable of producing insertional inactivation, means that pGSS33 is an extremely versatile vector, and makes the convenience and simplicity of the well known and well used *E. coli* vectors, pBR322 and pBR328, available for use with most Gram-negative bacteria.

ACKNOWLEDGEMENTS

I would like to express my thanks to Pam McTurk for her skills in producing the electron micrographs, and to Prof. W.J. Brammar, Dr. P.T. Barth and many other colleagues in Leicester and Runcorn, for their helpful discussions.

REFERENCES

Bagdasarian, M., Bagdasarian, M.M., Coleman, S. and Timmis, K.N.: New vector plasmids for gene cloning in *Pseudomonas*, in Timmis, K.N. and Puhler, A. (Eds.), Plasmids of Medical, Environmental and Commercial Importance. Elsevier/North-Holland Biomedical Press, Amsterdam, 1979, pp. 411–422.

Bagdasarian, M., Lurz, R., Rückert, B., Franklin, F.C.H., Bagdasarian, M.M., Frey, J. and Timmis, K.N.: Specific-purpose plasmid cloning vectors, II. Broad host range, high copy number RSF1010-derived vectors and a host-vector system for gene cloning in *Pseudomonas*. Gene 16 (1981) 237–247.

Barth, P.T. and Grinter, N.J.: Comparison of the deoxyribonucleic acid molecular weights and homologies of plasmids conferring linked resistance to streptomycin and sulfonamides. J. Bacteriol. 120 (1974) 618–630.

Barth, P.T.: RP4 and R300B as wide host range plasmid cloning vehicles, in Timmis, K.N. and Puhler, A. (Eds.), Plasmids of Medical, Environmental and Commercial Importance. Elsevier/North-Holland Biomedical Press, Amsterdam, 1979, pp. 399–410.

Barth, P.T., Tobin, L. and Sharpe, G.S.: Development of broad host range plasmid vectors, in Levy, S.B., Clowes, R.C. and Koenig, E.L. (Eds.), Molecular Biology, Pathogenicity and Ecology of Bacterial Plasmids. Plenum Press, New York, 1981, pp. 439–448.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S.: Construction and characterization of new cloning vehicles, II. A multipurpose cloning system. Gene 2 (1977) 95–113.

Boyd, A.C.: Restriction Systems in *Methylophilus methylotrophus*, Ph.D. thesis, University of Leicester, 1983.

Clarke, L. and Carbon, J.: Functional expression of cloned yeast DNA in *Escherichia coli*: specific complementation of arginosuccinate lyase (argH) mutations. J. Mol. Biol. 120 (1978) 517–532.

Coetzee, J.N.: Transmission of colicinogeny to *Providence* strains. Nature (London) 203 (1964) 897–898.

Datta, N. and Hedges, R.W.: Host ranges of R factors. J. Gen. Microbiol. 70 (1972) 453–460.

DeGraaff, J., Crosa, J.H., Heffron, F. and Falkow, S.: Replication of the nonconjugative plasmid RSF1010 in *Escherichia coli* K-12. J. Bacteriol. 134 (1978) 1117–1122.

DeMaeyer, E., Skup, D., Prasad, K.S.N., DeMaeyer-Guignard, J., Williams, B., Meacock, P., Sharpe, G.S., Pioli, D., Hennam, J., Schuch, W. and Atherton, K.: Expression of a chemically synthesized human α-1 interferon gene. Proc. Natl. Acad. Sci. USA 79 (1982) 4256–4259.

Gautier, F. and Bonewald, R.: The use of the plasmid R1162 and derivatives for gene cloning in the methanol-utilizing *Pseudomonas* AM1. Mol. Gen. Genet. 178 (1980) 375–380.

Grinter, N.J. and Barth, P.T.: Characterization of SmSu plasmids by restriction endonuclease cleavage and compatibility testing. J. Bacteriol. 128 (1976) 394–400.

Guerry, P., Van Embden, J. and Falkow, S.: Molecular nature of two nonconjugative plasmids carrying drug resistance genes. J. Bacteriol. 117 (1974) 619–630.

Heffron, F., Rubens, C. and Falkow, S.: Translocation of a plasmid DNA sequence which mediates ampicillin resistance: molecular nature and specificity of insertion. Proc. Natl. Acad. Sci. USA 72 (1975) 3623–3627.

Hennam, J.F., Cunningham, A.E., Sharpe, G.S. and Atherton, K.T.: Expression of eukaryotic coding sequences in *Methylophilus methylotrophus*. Nature 297 (1982) 80–82.

Holmes, D.S. and Quigley, M.: A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114 (1981) 193–197.

Kingsbury, D.T. and Helinski, D.R.: DNA polymerase as a requirement for the maintenance of the bacterial plasmid colicinogenic factor E1. Biochem. Biophys. Res. Commun. 41 (1970) 1538–1544.

Meyer, R., Boch, G. and Shapiro, J.: Transposition of DNA inserted into deletions of the Tn5 kanamycin resistance element. Mol. Gen. Genet. 171 (1979) 7–13.

Meyer, R., Laux, R., Boch, G., Hinds, M., Bayly, R. and Shapiro, J.: Broad host range IncP4 plasmid R1162: effects of deletions and insertions on plasmid maintenance and host range. J. Bacteriol. 152 (1982) 140–150.

Miller, J.H.: Experiments in Molecular Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1972.

Rubens, C., Heffron, F. and Falkow, S.: Transposition of a plasmid deoxyribonucleic acid sequence that mediates ampicillin resistance: Independence from host *Rec* functions and orientation of insertion. J. Bacteriol. 1218 (1976) 425–434.

Smith, H.R., Humphreys, G.O. and Anderson, E.S.: Genetic and molecular characterization of some non-transferring plasmids. Mol. Gen. Genet. 129 (1974) 229.

Soberon, X., Covarrubias, L. and Bolivar, F.: Construction and characterization of new cloning vehicles, IV. Deletion derivatives of pBR322 and pBR325. Gene 9 (1980) 287–305.

Stüber, D. and Bujard, H.: Organization of transcriptional signals in plasmids pBR322 and pACYC184. Proc. Natl. Acad. Sci. USA 78 (1981) 167–171.

Westmoreland, B.C., Szybalski, W. and Ris, H.: Mapping of deletions and substitutions in heteroduplex DNA molecules of bacteriophage lambda by electron microscopy. Science 163 (1969) 1343–1348.

Windass, J.D., Worsey, M.J., Pioli, E.M., Pioli, D., Barth, P.T., Atherton, K.T., Dart, E.C., Byrom, D., Powell, K. and Senior, P.J.: Improved conversion of methanol to single-cell protein by *Methylophilus methylotrophus*. Nature 287 (1980) 396–401.

Womble, D.D., Dean, P.T. and Rownd, R.H.: Method for obtaining more-accurate covalently closed circular plasmid-to-chromosome ratios from bacterial lysates by dye-buoyant density centrifugation. J. Bacteriol. 130 (1977) 148–153.

Wood, D., Hollinger, M.F. and Tindol, M.B.: Versatile cloning vector for *Pseudomonas aeruginosa*. J. Bacteriol. 145 (1981) 1448–1451.

Communicated by S. Brenner.

From: MOLECULAR BIOLOGY, PATHOGENICITY, AND ECOLOGY OF BACTERIAL PLASMIDS
Edited by Stuart B. Levy, Royston C. Clowes, and Ellen L. Koenig
(Plenum Publishing Corporation, 1981)

# DEVELOPMENT OF BROAD HOST-RANGE PLASMID VECTORS

Peter T. Barth*, Lyn Tobin* and Geoffrey S Sharpe**

*ICI Corporate Laboratory, The Heath, Runcorn, Cheshire England
**ICI Joint Laboratory, University of Leicester Leicester, England

## INTRODUCTION

The majority of cloning vectors developed so far are based on plasmids or phages with a narrow host-range such as ColEI(pBR322[1]), P15A(pACYC184[2]),pSC101[3] or phage λ[4]. These are limited to *Escherichia coli* or closely related enterobacterial species. For the genetic analysis and manipulation of a wider range of micro-organisms, including those of agricultural, medical, environmental or industrial importance, use has been made of plasmids belonging to the IncP group such as RP4 or RK2. For example RP4 has unique cloning sites for *Eco*RI, *Bam*HI, *Hind*III,*Hpa*I and *Bgl*II[5,6] and has been converted into a *Sal*I cloning vector pRP301[7,8]. From RK2 a two vehicle system: pRK290 plus pRK2013, has been constructed[9]. But such vectors remain relatively large, have a low copy number and are therefore unsuitable for some cloning purposes. For ease of manipulation and high gene dosage of the cloned material we need small, high copy number, broad host-range plasmids.

The non-conjugative plasmids have mostly not been given incompatibility group designations, although several types have been reported to be compatible with one another[10,11]. These are shown in Table 1 although many more groups probably exist. Only one of these groups, as represented by R300B, has been given an Inc designation viz IncQ[12]. This seems to be the only one of these plasmid groups with an extended host range as judged by attempts to mobilize members of these groups from *E.coli* into *Pseudomonas aeruginosa* or *Methylophilus methylotrophus* using RP4 (unpublished observations). We used plasmid mobilization because, in contrast to transformation, it is little affected by restriction systems in the host. IncQ plasmids are very commonly found[10], over a broad geographical and bacterial

Table 1. Incompatibility Groups of Non-conjugative Plasmids

| Plasmid | Size (kb) | Phenotype[a] | Inc group | References |
|---|---|---|---|---|
| P15A | 2.0 | cryptic | - | 2 |
| ColE1-K30 | 5.3 | $ColE1^+$ | - | 10,11 |
| ColE2-P9 | 6.9 | $ColE2^+$ | - | 10 |
| ColK-235 | 6.9 | $ColK^+$ | - | 13 |
| pHH509 | 8.4 | $Ap^R$ | - | 11 |
| NTP2,R300B | 8.5 | $SmSu^R$ | Q | 10,11 |
| pSC101 | 9.3 | $Tc^R$ | - | 11 |
| R831a | 13.8 | $SmKm^R$ | - | 11 |

[a]Phenotypic symbols are according to Novick et al[14]. The references give information on incompatibility relationships.

species range[15] and thus represent a particularly successful plasmid type. Members of this group that have been studied in recent years are R300B[15], NTP2[10], RSF1010 (which is identical to NTP2[16]) and R1162[17]. All these plasmids are indistinguishable at present[15,18].

Apart from their broad host-range, the other characteristics that might make IncQ plasmids suitable as cloning vectors are: 1) they are relatively small (the majority of those studied were 8.5kb[15]), 2) they have a relatively high copy number (in *E.coli*, measured as supercoiled DNA, we found 8-12 copies per chromosome[15,11]) and 3) they are non-conjugative but efficiently mobilized by conjugative plasmids of various groups (see below).

In this paper we describe further studies on R300B including the derivation of a restriction map, a genetic map by Tn3 mutagenesis, polypeptide synthesis in mini-cells and the creation of various derivatives useful in cloning.

## MATERIALS AND METHODS

Growth media. Growth media were as previously described[19]. For *M.methylotrophus* we used M9 salts medium supplemented with 1% methanol plus $FeCl_3$,10 mg/l.

Plate-mating. Plasmids were generally transferred on solid media: a colony of the recipient strain was spread with a little sterile saline onto half the surface of a selective medium plate and colonies of the donor strains were then streaked in one direction across the plate into the recipient. For transfer into a variety of different (usually prototrophic) species, we used a *thy* donor which can be counterselected on Isosensitest (Oxoid) medium.

Plasmid DNA Isolation. Plasmid DNA was isolated after sarkosyl lysis

of cells in ethidium bromide-CsCl gradients in a scaled-up version of the previously described method[15].

DNA Restriction Analysis. Plasmid DNA was restricted and analysed by gel electrophoresis as previously described[7]. Bands were cut out of gels and the DNA isolated by electroelution into dialysis tubing. DNA ligations overnight at 10°C and subsequent transformation were carried out by standard methods.

Analysis of Polypeptides in Minicells. Plasmids were mobilized or transformed into the minicell producing strain χ1411 *trp*EΔ. Minicells were isolated in freeze-thaw generated sucrose gradients[5], radiolabelled, the polypeptides electrophoresed through polyacrylamide gels and autoradiographed approximately as described by Dougan et al[20].

Isolation of Tn3 derivatives of R300B. Cultures of C600 (R1*drd*19/R300B) were grown overnight at 30°C to facilitate transposition[21] and then plated out on media containing 1 mg ampicillin/ml. Clones containing R300B::Tn3 are selected because the increase in copy number of the *bla* gene produced by transposition from the low copy number R1*drd*19 to the high copy number R300B leads to a corresponding increase in ampicillin resistance[22].

## RESULTS

### The Host-range of IncQ Plasmids

We have previously reported[15] that the host-range of IncQ plasmids includes *E.coli*, *Salmonella typhimurium*, *S.senftenberg*, *S.dublin*, *Proteus mirabilis*, *P.morganii*, *Providencia sp.* and *Pseudomonas aeruginosa*. Since then they have also been reported to replicate stably in *Ps. phaseolicola*[23] *Rhizobium meliloti*[24], *Rhodopseudomonas sp*[25] and *Acinetobacter calcoaceticus*[26]. Experiments in this laboratory have shown that R300B or its derivatives can also stably inhabit *Methylophilus methylotrophus*, *Alcaligenes eutrophus*, *Ps.putida*, *Klebsiella aerogenes* and *Serratia marcescens*. We have used conjugative plasmids of the following groups to mobilize IncQ plasmids into some of the above species:- IncFI: F10415; IncFII: R1*drd*19, JR72; IncIα: R144*drd*312, R483, JR66a; IncN: N3T; IncP: RP4, R751, R702; IncW: R7K. Not all of these plasmids are effective in all crosses except for RP4. This host-range list is not complete, merely those tested so far. However we have not yet found a Gram negative species that IncQ plasmids do not inhabit.

### A Restriction Map of R300B

The construction of broad host-range cloning vectors from R300B depends upon knowing its restriction and genetic maps. A simple map of R300B was published recently[7]. Figure 1 shows our present restriction map. Although the *Eco*R1 and *Hpa*I sites are not distinguishable

by restriction analysis we have shown that deletion of DNA to the left of the *Eco*R1 site (as in the construction of pGSS8, see below) does not remove the *Hpa*I site. A genetic map of RSF1010 relative to the *Eco*RI has been produced in Falkow's laboratory[16,18]. We have superimposed it onto our map in the orientation shown because removal of the 0.8kb *Pst*I fragment from R300B led to loss of sulphonamide resistance[7] whereas the cloning of random *Sst*I-cut fragments of *M.methylophilus* chromosomal DNA into the *Sst*I site always led to loss of streptomycin resistance.

In addition to the sites shown in Figure 1 there are also several sites on R300B susceptible to *Bgl*I[27], *Hae*II, *Hae*III, *Hinf*I[27] and *Sau*3A and a single site for *Sst*II[28]. But we found no cleavage sites for *Bam*HI, *Bgl*II, *Cla*I, *Hind*III, *Kpn*I, *Pvu*I, *Sal*I, *Sma*I, *Xba*I, *Xho*I, or *Xma*I. A similar map has been published for RSF1010[29] and two simple maps for R1162 are also consistent with it[17,30].

## A Genetic Map of R300B

We have begun a genetic analysis of R300B using Tn3 mutagenesis, Figure 2 shows some of the Tn3 insertions we have mapped using restriction endonucleases. These data confirm the hypothesis first proposed by Heffron et al[18,31] that the two genes giving sulphonamide *(sul)* and streptomycin *(aph*C*)* resistance are in a single operon with *sul* proximal and *aph*C distal to the promoter : thus, pLT108 has lost

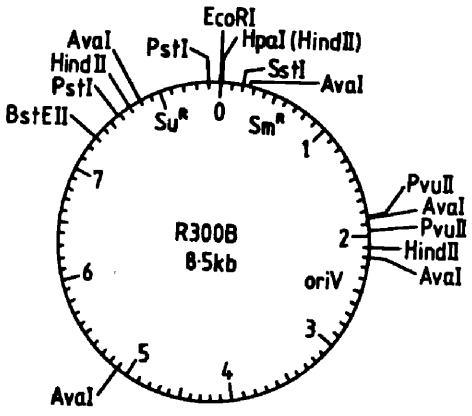



Fig 1. A restriction map of R300B

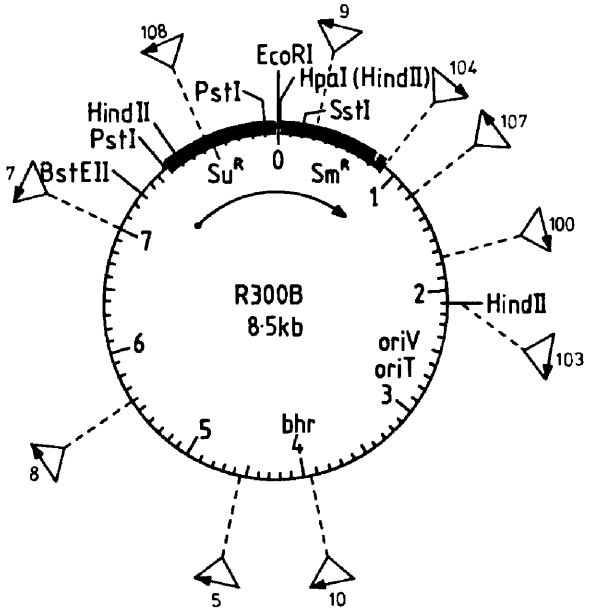



Fig.2. A Tn3 insertion map of R300B. The arrow heads on the Tn3 symbols are at the *bla* end and show the direction of transcription out of the transposon[31] when Tn3 is in its normal repressed (*tnp*R+) state[32]. The arrow within the circle shows the proposed start and direction of transcription of the operon giving drug resistances.

both $Su^R$ and $Sm^R$ whereas pLT9 has lost only $Sm^R$. The promoter is presumably to the left of the *Pst*I site at 7.6 kb since, as noted above, removal of the *Pst*I fragment does not lead to loss of $Sm^R$, although the level of resistance is somewhat reduced.

Insertions of Tn3 at around 4kb affect the broad host-range *(bhr)* properties of R300B. Such plasmids are still mobilizable into *E.coli* but not into *M.methylotrophus*. We do not know whether this is an effect on their transfer into, or maintenance in, the latter species. We also include in Figure 2 the approximate site of the presumed transfer origin (*ori*T). Nordheim and Timmis[33] have shown that the relaxation nick site of RSF1010, presumed to be *ori*T, is close to, but not at, the site of the replication origin (*ori*V).

## Transcription of the *sul aph*C Operon

We have examined the polypeptides expressed by R300B and some of its Tn3 derivatives in minicells. As *sul* and *aph*C are in the same operon, their expressed polypeptides are likely to be present in about equal amounts. In Figure 3 it can be seen that band E is reduced to approximately half intensity by the Tn3 insertion of pLT9 ($Sm^S$), showing that the eliminated polypeptide (the amino-glycoside 3" phosphotransferase) bands at this position and suggesting that the remaining polypeptide in band E is the dihydropteroate synthase (*sul*). Dougan et al[34] have also deduced the identity of the former polypeptide. These polypeptides have a molecular weight of a little under 30,000 which require a coding capacity of just less than 900 base pairs each. The *sul* and *aph*C genes have therefore been drawn as this size on Fig.2. Their positions come from various considerations: (i) the Tn3 insertion in pLT104 does not affect the level of $Sm^R$, (ii) genes in an operon are normally adjacent and (iii) cleavage at the *Eco*RI site of R300B separated but did not inactivate the genes giving $Su^R$ and $Sm^R$ in the construction of pGSS8 and 9 from pGSS6 (next section). The *Eco*RI and probably the *Hpa*I sites therefore appear to be cloning sites within a transcription unit that would not give rise to fused proteins. This *sul aph*C operon appears to be highly expressed from a comparison of band E with the β-lactamase bands F and G.

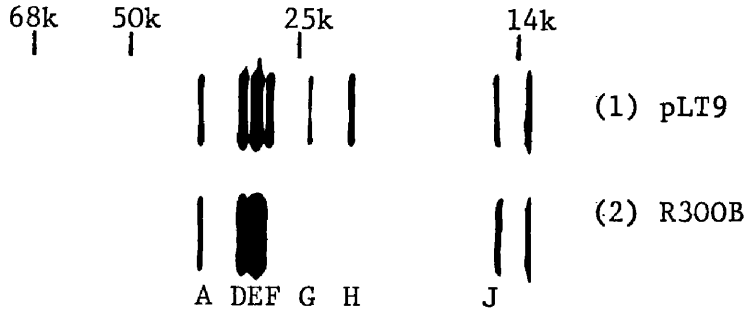



Fig.3. Autoradiogram of $^{35}S$-labelled polypeptides from minicells containing (1) pLT9 (see Fig 2) and (2) R300B. Molecular weight markers are indicated above.

## Construction of new Cloning Vectors from R300B

R300B has few sites for the restriction enzymes normally used for cloning. We therefore generated new derivatives by introducing genes from other plasmids. As pBR322 has been sequenced and is well understood, we used it as a source. By *Hae*II partial cleavage of both R300B and pBR322, followed by ligation, we generated a series of cointegrate plasmids such as pGSS6 (Figure 4). In this figure we have designated the origin of pBR322 as *ori*E and that of R300B as *ori*Q. Replication from *ori*E is dependent on *pol*A$^{+}$[35] whereas from *ori*Q it is not[36]. We have used this difference to distinguish between the two types of origin in the cointegrates and the subsequent cleavage products. We next cleaved pGSS6 and similar cointegrates with *Eco*RI and self-ligated the two fragments formed. From pGSS6, the ApSu$^{R}$ plasmid (pGSS9) produced was found to be *pol*A$^{+}$ dependent and not mobilizable into eg *M.methylotrophus* whereas the TcSm$^{R}$ plasmid (pGSS8) shown in Figure 4 is *pol*A$^{+}$ independent and has the broad host-range of the parental R300B. It also has the *Cla*I, *Hind*III, *Bam*HI and *Sal*I cloning sites in the gene conferring Tc$^{R}$. Transcription of the gene conferring Sm$^{R}$ is however dependent on a backward reading promoter near the beginning of the *tet* gene. Cloning into the *Cla*I or *Hind*III sites can therefore cause loss of both markers.

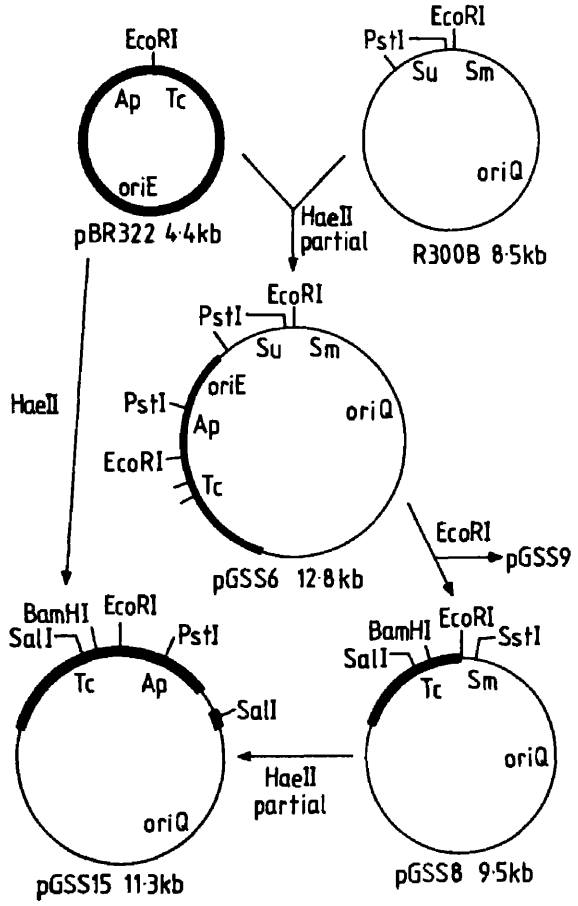



Fig. 4 Construction of R300B derivatives using DNA from pBR322.

Because of this, we decided to introduce the *bla* gene, which has its own promoter, from pBR322. *Hae*II cleavage of pBR322 and pGSS8 followed by ligation and transformation into a *pol*A host led to the recovery of pGSS15. In this plasmid the *bla tet* boundary of pBR322 has been reconstructed but the plasmid has the replication and host-range properties of R300B. (Unfortunately, an extra *Hae*II fragment carrying the *Sal*I site was also introduced. We are at present attempting to remove it).

Another series of vectors was made by addition rather than substitution of genes: these retain the strong *sul aph*C promoter of R300B. We restricted pACYC184[2] and pMK20[38] plasmid DNAs to completion with *Hae*II, ran them on a gel and then electroeluted the 1.3kb band from the former and the 1.5kb band from the latter. These contain the genes conferring $Cm^R$ and $Km^R$ respectively from the two plasmids as shown in Figure 5. These were separately ligated to partially *Hae*II cut R300B DNA eluted from a gel at the whole linear plasmid (8.5kb) position and transformed into *E.coli* selecting for $Cm^R$ or $Km^R$ clones. Few of the $Cm^R$ clones proved to contain R300B::$Cm^R$ plasmids like the example pTB86 in Figure 5: the majority had plasmids consisting of

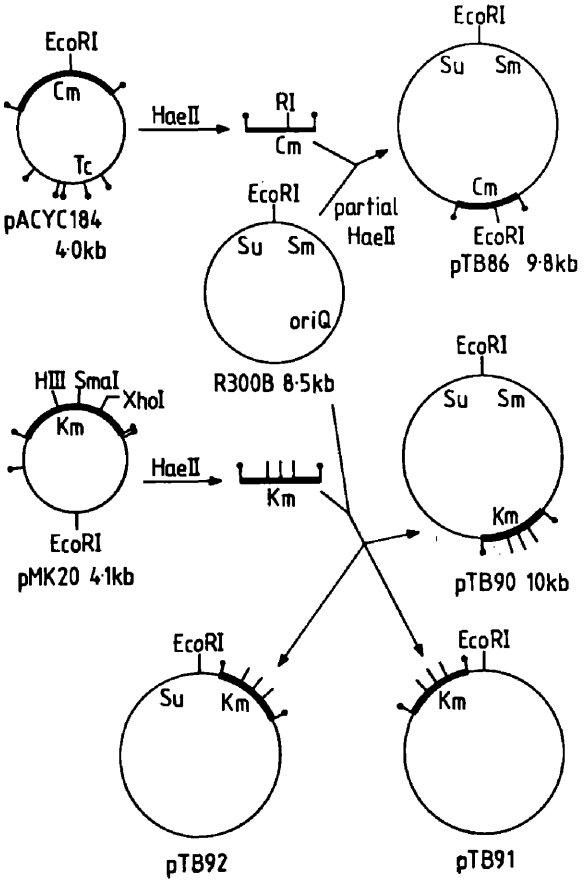



Fig.5. Construction of the $Cm^R$ and $Km^R$ derivatives of R300B. The *Hae*II sites are marked thus ↑.

the two largest, adjacent *Hae*II fragments of pACYC184 which can form a replicon[39]. The original *Eco*RI site needs to be removed from pTB86 and its sisters in order to use the one in the gene conferring $Cm^R$ as an insertional inactivation site.

The $Km^R$ clones were found to contain R300B::$Km^R$ plasmids. Of 78 examined so far, 62 were $SuSmKm^R$ like pTB90, 6 were $Km^R$ only, like pTB91 and 10 were $SuKm^R$ like pTB92. The latter two classes are consistent with the model that R300B has a single *sul aph*C operon with *sul* being proximal to the promoter and suggest that the inserted *Hae*II fragment blocks transcription in either orientation. Each of these plasmids has gained a *Hind*III, *Sma*I and *Xho*I cloning site. R300B has about 20 *Hae*II sites so we would expect insertion of the $Km^R$ fragment at several of these sites causing other changes in phenotype apart from the loss of $SuSm^R$ or $Sm^R$ already noted. We have therefore tested these 62 clones for such changes. Two have a reduced and one an increased, plasmid copy number (as determined by the Km and Sm resistance levels), 13 are non-mobilizable and 10 have a reduced host-range. We are mapping the inserts in these plasmids at present to complement our Tn3 mutagenesis mapping.

## Transposon Derivatives of R300B

Another way of introducing cloning sites into plasmids is by using transposons. Such derivatives can then be further adapted by excising specific segments from them using pre-existing and newly-introduced restriction sites (as we have done for the RP4::Tn7 system[5,7]). This is another reason for our isolating the R300B::Tn3 derivatives described above. A Tn5 derivative of R300B (pTB70) has been used by us recently to clone *gdh*$^+$ from *E.coli* into a glutamate synthase mutant of *M.methylotrophus*[8]. By the consequent switch in the pathway for ammonium assimilation this has led to a significant improvement in the efficiency of conversion of methanol to single-cell protein by this organism. We have also isolated a Tn1771[40] derivative of R300B. The *tet* region of Tn1771 (and the indistinguishable Tn1721) can be amplified to give multiple tandem repeats[41]. Genes cloned into this portion will therefore be similarly amplified. Furthermore genes cloned into suitable sites on transposon derivatives of R300B can be transposed into the chromosomes of a wide range of organisms.

## DISCUSSION

We have described the genetic structure and properties of R300B and the construction from it of some broad host-range cloning vectors which are proving to be very useful in our cloning systems. Bagdasarian et al[29] have also generated IncQ vectors, with *Bgl*II and *Xba*I insertional inactivation cloning sites. Meyer and Shapiro[37] and Bagdasarian et al[29] have reported the construction of cointegrates between a ColE1 and an IncQ plasmid using the *Eco*RI or *Pst*I sites respectively, but these plasmids do not stably inhabit *Pseudomonas*,

or if selected, delete part of the ColE1[37]. (Gautier and Bonewald[30] also made such cointegrates via *Eco*RI but they do not report any reduction in host-range). It seems there may be a region of ColE1 that is inimical to broad host-range maintenance. If so, it must have been disrupted in the construction of pGSS6, as this plasmid (and its derivatives pGSS8 and 15) do not suffer this handicap: they are stable in at least *E.coli*, *P.aeruginosa*, *M.methylotrophus* and *A.eutrophus*.

There is some confusion in the literature about whether or not IncQ plasmids require polymeraseI for replication. Our data confirm the painstaking data of Grindley and Kelley[36] that they do not. Gautier and Bonewald[30] however, have drawn the opposite conclusion. This may be due to the slight instability of these plasmids in some *pol*A mutants[36].

There is clearly plenty of scope for the further development of these broad host-range cloning vectors. We do not know at present whether they can be reduced in size without loss of valuable functions. But we can introduce or select stronger promoters and put in cloning sites suitably down-stream from ribosome binding sites with perhaps a secretion leader sequence between. We also hope that our genetic analysis of R300B will lead to an understanding of how this fascinating plasmid functions.

ACKNOWLEDGEMENTS

We are grateful to T C Hodgman for his analysis of the polypeptides synthesised in minicells and to S A Withe for his help with the restriction analysis of R300B and the photography. We would also like to thank D R Helinski for his hospitality to one of us (PTB) in his laboratory recently, during which time the $Cm^R$ and $Km^R$ derivatives of R300B were isolated with the particular help of D Stalker.

REFERENCES

1 F Bolivar, R L Rodriguez, P J Greene, M S Betlach, H L Heyneker, H W Boyer, J H Crosa and S Falkow, Gene 2:95 (1977).
2 A C Y Chang and S N Cohen, J Bacteriol 134:1141 (1978).
3 S N Cohen and A C Y Chang, J Bacteriol 132:734 (1977).
4 N E Murray, W J Brammar and K Murray, Molec Gen Genet 150:53(1977)
5 P T Barth and N J Grinter, J Mol Biol 113:455 (1977)
6 A DePicker, M Van Montagu and J Schell in "DNA Insertion Elements, Plasmids and Episomes" Bukhari, Shapiro and Adya, eds, 678 (1977), Cold Spring Harbor Laboratory.
7 P T Barth, in "Plasmids of Medical, Environmental and Commercial Importance", K N Timmis and A Pühler eds, 399 (1979), Elsevier/North Holland.
8 J D Windass, M J Worsey, E M Pioli, D Pioli, P T Barth, K T Atherton, E C Dart, D Byrom, K Powell and P J Senior, Nature 287:396 (1980).

9 G Ditta, S Stanfield, D Corbin and D R Helinski, Proc Natl Acad Sci in press.
10 H R Smith, G O Humphreys and E S Anderson, Molec Gen Genet 129:229 (1974).
11 P T Barth, H Richards and N Datta, J Bacteriol 135:760 (1978).
12 N J Grinter and P T Barth, J Bacteriol 128:394 (1976).
13 G J Warren, A J Twigg and D J Sherratt, Nature 274:259 (1978).
14 R P Novick, R C Clowes, S N Cohen, R Curtiss III, N Datta and S Falkow, Bacteriol Rev 40:168 (1976).
15 P T Barth and N J Grinter, J Bacteriol 120:618 (1974).
16 J De Graaff, J H Crosa, F Heffron and S Falkow, J Bacteriol 134:1117 (1978).
17 R Meyer, G Boch and J Shapiro, Molec Gen Genet 171:7 (1979).
18 F Heffron, C Rubens and S Falkow, Proc Natl Acad Sci 72: 3623 (1975).
19 P T Barth, N J Grinter and D E Bradley, J Bacteriol 133:43(1978).
20 G Dougan, M Saul, A Twigg, R Gill and D Sherratt, J Bacteriol 138:48 (1979).
21 P J Kretschmer and S N Cohen, J Bacteriol 139:515 (1979).
22 J A Nowak and A J Clark, Lunteren Lectures on Molecular Genetics (1979).
23 N J Panopoulos, B J Staskawicz and D Sandlin in "Plasmids of Medical, Environmental and Commercial Importance", K N Timmis and A Pühler eds, 365 (1979), Elsevier/North Holland.
24 Gary Ditta, personal communication.
25 JoAnne Williams, personal communication.
26 Tom Schmidthauser, personal communication.
27 Wolfgang Schuch, personal communication.
28 Kim Ellis, personal communication.
29 M Bagdasarian, M M Bagdasarian, S Coleman and K N Timmis, in "Plasmids of Medical, Environmental and Commercial Importance", K N Timmis and A Pühler eds., 411 (1979), Elsevier/North Holland.
30 F Gautier and R Bonewald. Molec Gen Genet 178:375 (1980).
31 C Rubens, F Heffron and S Falkow, J Bacteriol 128:425 (1976).
32 F Heffron, B J McCarthy, H Ohtsubo and E Ohtsubo, Cell 18:1153 (1979).
33 A Nordheim and K N Timmis, Lunteren Lectures on Molecular Genetics (1979).
34 G Dougan, M Saul, A Twigg, R Gill and D J Sherratt, J Bacteriol 138:48 (1979).
35 D T Kingsbury and D R Helinski, Biochem Biophys Res Comm 41:1538 (1970).
36 N D F Grindley and W S Kelley, Molec Gen Genet 143:311 (1974).
37 R Meyer and J Shapiro,EMBO Workshop on Control of Replication and Partitioning of Bacterial Chromosomes and Plasmids,Leicester (1979).
38 M Kahn and D R Helinski, Proc Natl Acad Sci 75:2200 (1978).
39 David Stalker, personal communication.
40 F Schöffl and A Pühler, Genet Res Camb. 33:253 (1979).
41 R Mattes, H J Burkardt and R Schmitt, Molec Gen Genet.168:173 (1979).

# THE DEVELOPMENT AND USE OF A VECTOR SYSTEM FOR *METHYLOPHILUS METHYLOTROPHUS*


## SUMMARY

The obligate methylotroph, *Methylophilus methylotrophus*, used by ICI for its Single Cell Protein production, may represent a valuable alternative host organism to *E.coli* for the commercial production of heterologous gene products. The organism has the advantages of being safe, coupled with an ability to grow well, in large quantities, on a cheap carbon source. This thesis describes the construction, characterisation and analysis of a series of plasmid cloning vectors designed for use in *M.methylotrophus*. The vectors are based on the IncQ plasmid R300B and maintain a broad host range with an increased capacity for easy-to-use cloning sites mainly derived from the *E.coli* plasmids pBR322 and pBR328. All the vectors thus carry antibiotic-resistance genes containing restriction sites which could lead to insertional inactivation as a means of detecting recombinants. One plasmid in particular, pGSS33, has four antibiotic-resistance genes all of which contain at least one such restriction site.

The first demonstration of expression of a eukaryotic coding sequence, murine dihydrofolate reductase (DHFR), in *M.methylotrophus* has been described. This has been followed by expression of *E.coli* B-galactosidase and synthetic human $\alpha$-1 interferon, all making use of pGSS vectors. The pGSS15-DHFR plasmid, pDHFR2.43, may turn out to be a valuable test plasmid for studying the stability of cloned eukaryotic coding sequences in both *E.coli* and *M.methylotrophus*; it is readily detected (trimethoprim-resistant) and can be grown with or without selection.

A start has been made towards providing increased expression from vectors carrying strong *E.coli* promoters (*lac*UV5 and synthetic *trp*) which have been demonstrated to work well in *M.methylotrophus*.

Broad host range cosmid vectors have been constructed which have the potential to be used for the production of gene-banks in Gram-negative organisms other than *E.coli*.

Copy numbers of the vector plasmids have been determined in *E.coli* strains and several different methods of measurement reviewed in an attempt to find one suitable for use with *M.methylotrophus*. An encouraging lead has been identified in the search for a high copy number plasmid, which coupled with a strong promoter, could provide the basis for a very efficient batch production process.

Thus with the availability of easy-to-use cloning vectors, convenient delivery systems and the accumulated evidence of strong *E.coli* promoters working efficiently in *M.methylotrophus*, this organism can seriously be considered as a safe alternative host to *E.coli*.



Geoffrey S. Sharpe



Ph.D. Thesis 1984.