INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

ALTERNATIVE INITIATION OF TRANSLATION IN *ESCHERICHIA COLI*

by

Ashkan Golshani

A thesis submitted in conformity with the requirements
for the Degree of Master of Science
Graduate Department of Botany
University of Toronto

ALTERNATIVE INITIATION OF TRANSLATION IN *ESCHERICHIA COLI*

Ashkan Golshani, Department of Botany, University of Toronto

ABSTRACT In *Escherichia coli*, initiation of translation for mRNAs is known to be promoted by the binding of the 3' end of 16S ribosomal RNA (rRNA) to a purine rich sequence (known as Shine-Dalgarno sequence, SD) 5' to the initiator codon (AUG). Using an *E.coli* expression system we have proposed the presence of alternative sites (loops # 42 and # 43, and nucleotides 458-467) on the 16S rRNA capable of binding functional mRNAs. We have also shown that the 5' untranslated regions of clover yellow mosaic virus and papaya mosaic virus RNA genomes are both capable of initiating translation in *E.coli*.

ACKNOWLEDGMENTS

I would like to thank Dr. Mounir G. AbouHaidar for giving me the chance to work in his laboratory. Aside from the technical skills I have gained in his lab, I feel that his guidance has helped me to mature into an independent researcher and person. Mounir, thanks very much for everything.

I would also like to thank Dr. Ivan Ivanov, the man who introduced me to the wonderful world of molecular biology as well as patient therapy for his guidance and care, but most importantly for his friendship. This research could not have been initiated or completed without his help. Ivan, thanks very much for the good times and unforgettable memories. I hope to see you in Bulgaria soon.

I would like to thank Drs. Rick Collins and Sadhna Joshi for their helpful discussions and commentary on my work as members of my supervisory committee.

Many thanks to Richard Collins, who provided technical knowledge and advice on every aspect of the lab (including Mounir). Also thanks to the "mercenaries" who made this lab a harmonious and joyful place to work: Ali, Alice, Andrei, Duncan, Glen, Huimin, Jianhua, Kathy, Kevin, Richard C., Richard I. and Rosemary. You guys are the greatest.

I also must thank all those wonderful friends who helped me out mentally and physically, on the "last minute" compilation of this "piece of work": Andreas ("Doggie"), Andrea C. ("B.B.C."),

Andrea G. (sorry about the nose!), Carolyn (BC-Vancouver, Vancouver-BC), Cesare ("the best"), Daphne ("Qua Qua"), Jackie ("WOLF"), Jennifer ("Betty-Ann"), Mara (stop kicking my leg in soccer please!), Rejash ("Stud Master"), Rich C. (OK, I'll get you the recipe!), Rich I. (sorry you didn't make it to the Junior Kindergarten Soccer League), Rosemary ("of the higher class"), Valerie ("Jose"), and all those who were wishing me luck. The cheques are in the mail!

Finally, I would like to thank my family for their encouragement and belief in me, and for their wisdom and guidance that has molded me into the person I am. Without their love and support this endeavor would not have been successful.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS

Ap	ampicillin
A-site	aminoacyl-site
ATP	adenosine triphosphate
bp	base pair
BSA	bovine serum albumin
$CaCl_2$	calcium chloride
CAT	chloramphenicol acetyl transferase
Ci	Curie
Cm	chloramphenicol
CoA	acetyl coenzyme A
CYMV	clover yellow mosaic virus
d	dalton
ddNTP	2',2'-dideoxyribonucleotide triphosphate
DEAE	diethylaminoethyl
dH_2O	distilled water
ddH_2O	double distilled water
DB	downstream box
DNA	deoxyribonucleic acid
dNTP	2'-deoxyribonucleotide triphosphate
ds	double-stranded
DTNB	5',5'-dithiobis-2-nitrobenzoic acid
DTT	dithiothreitol
EDTA	ethylenediamenetetraacetic acid
EF	elongation factor
E-site	exit-site
EtBr	ethidium bromide
EtOH	ethanol
fMet-$tRNA_f$	formyl-methionyl-$tRNA_f$
HCl	hydrochloric acid
IF	initiation factor
KAc	potassium acetate
kb	kilobase
kbp	kilobase pair
KCl	potassium chloride
Klenow	*E. coli* DNA polymerase I, Klenow fragment
LB	Luria-Bertani (media)
mA	milliampere
Met	methionine
Met-$tRNA_f$	methionyl-tRNA for start codon
$MgCl_2$	magnesium chloride
M_r	molecular weight
NaAc	sodium acetate
NaCl	sodium chloride
NaOH	sodium hydroxide
NDP	ribonucleotide diphosphate
nm	nanometre
nt	nucleotide
NTP	ribonucleotide triphosphate

OD	optical density
P-site	peptidyl-site
PMV	papaya mosaic virus
RBS	ribosome binding site
RF	releasing factor
RNA	ribonucleic acid
rRNA	ribosomal RNA
S	sedimentation coefficient
SD	Shine-Dalgarno
SDS	sodium dodecyl sulfate
SD_n	SD sequence native to the CAT gene
ss	single-stranded
SSC	1X: 0.15 M NaCl and 0.015 M sodium citrate, pH 7.0
TB	terrific broth (media)
TBE	1X: 0.1 M Tris, 0.1 M boric acid and 2 mM EDTA, pH 8.3
Tc	tetracycline
TE	10 mM Tris-HCl, pH 8.0, 1 mM EDTA
T_h	hybridization temperature
TMV	tobacco mosaic virus
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
TS	50 mM Tris-HCl pH 7.8 and 150 mM NaCl
V	volt
YT	yeast-tryptone (media)
μCi	micromurie
μg	microgram
μl	microlitre
μM	micromolar
ϵ	a sequence derived from RBS of bacteriophage T7 gene *10*
Ω	5' untranslated region of TMV RNA

1. INTRODUCTION

Although the process of protein synthesis occurs continuously from start to finish, it is customary to divide this process into three stages: initiation, elongation, and termination (see Figure 1.1). To examine each of these stages, two features of mRNA directed polypeptide chain assembly must be considered. First, this assembly is a unidirectional process which begins at the amino terminal and proceeds toward the carboxyl end. Second, mRNA translation begins at a start codon (usually AUG) located at a unique position which identifies the 5' end of the mRNA (Watson *et al.*, 1987).

The key organelles in protein synthesis are the ribosomes. Prokaryotic ribosomes (referred to as 70S particles because of their sedimentation characteristic) are composed of one small (30S) and one large (50S) subunits. The 30S ribosomal subunit contains a single ribosomal RNA (rRNA) molecule (16S) of 1542 nucleotides and 21 different proteins (S1-S21). The 50S subunit has two rRNAs associated with 34 different proteins (L1-L34): the large rRNA (23S) consisting of 2904 nucleotides and the smaller rRNA (5S) consisting of 120 nucleotides (Singer and Berg, 1991; Watson *et al.*, 1987).

Initiation of translation in prokaryotes involves the binding of initiator tRNA and mRNA to the 70S ribosome. This process results in the selection of a specific mRNA for translation and determines the initiator codon as well as the exact phasing of the

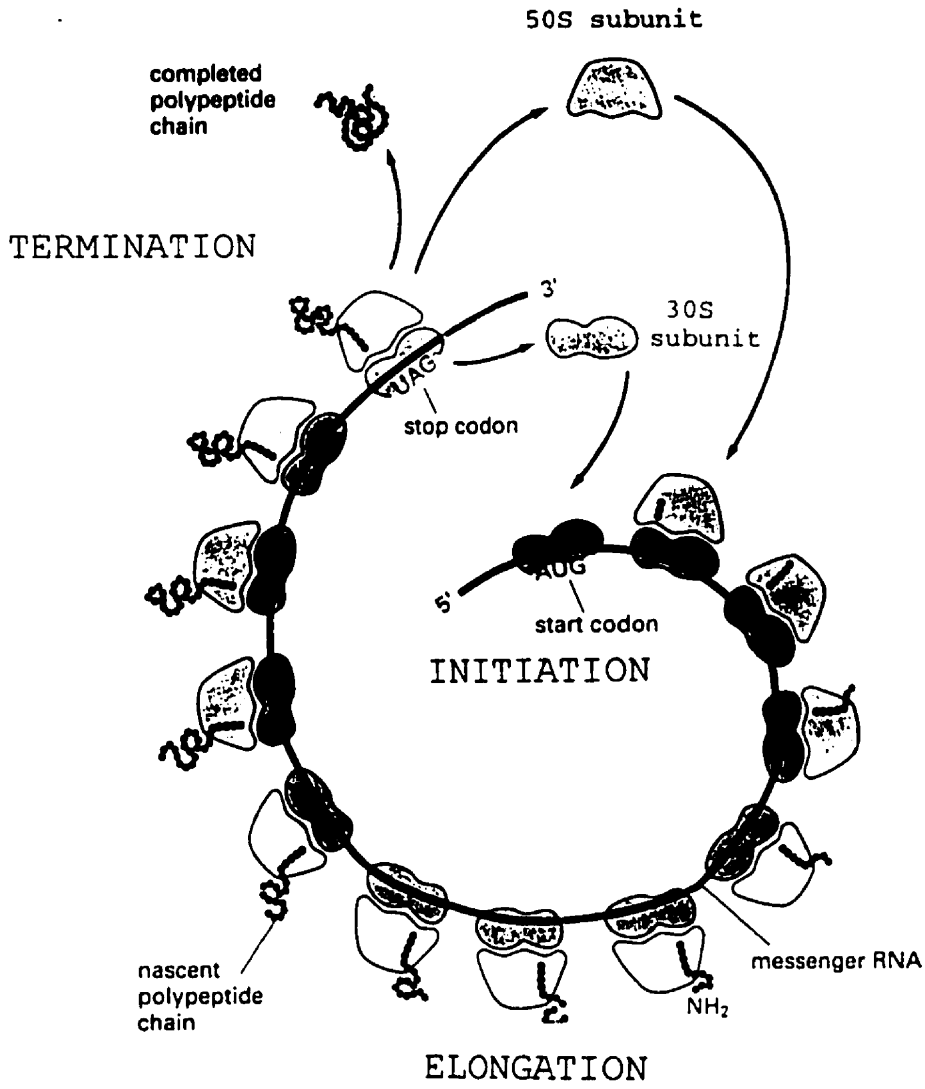


Figure 1.1: Schematic diagram illustrating protein synthesis in prokaryotes. The first step in protein synthesis is the binding of 30S ribosomal subunit to an mRNA molecule. The 50S subunit then binds to the 30S subunit to form a complete initiation complex to which the initiator tRNA binds and translation begins. The ribosomes continue translating the nucleotide sequence into a polypeptide chain (elongation) until they reach a stop codon, at which point the completed polypeptide chain is released (termination)and the ribosomes are dissociate into 50S and 30S subunits. (Adapted from Alberts *et al.*, 1994).

reading frame for the mRNA. It is believed that in most cases the initiation phase is the rate limiting step in the process of protein synthesis; thus, initiation also determines the rate at which the encoded protein is synthesized (Stormo, 1986; van Knippenberg, 1990; Singer and Berg, 1991; Nierhaus *et al.*, 1993).

Initiation of protein synthesis in prokaryotes requires several proteins called initiation factors. Initiation factors IF1 and IF3 cause dissociation of the 70S ribosome by binding to the 30S ribosomal subunit. The 30S subunit carrying IF1 and IF3 further interacts with initiation factor IF2, the initiator tRNA [formyl-methionyl-tRNA$_f$ (fMet-tRNA$_f$)], GTP (guanosine triphosphate), and mRNA through a series of intermediates giving rise to the 30S initiation complex. Formation of a fully functional initiation complex is completed by association of this 30S initiation complex with the 50S ribosomal subunit and subsequent release of the initiation factors and GDP (guanosine diphosphate). The 70S initiation complex is then ready to enter the elongation phase of protein synthesis (Stormo, 1986; van Knippenberg, 1990; Draper, 1993; McCarthy and Brimacombe, 1994).

The association of the 30S and 50S subunits during the initiation process generates the three functional sites required for protein assembly: P (peptidyl) site, A (aminoacyl) site and E (exit) site. The initiator tRNA (fMet-tRNA$_f$) occupies the P-site. Elongation factor EF-Tu forms a complex with an aminoacyl-tRNA and GTP and the resulting complex binds to the A-site releasing GDP bound to EF-Tu (to be recycled). The peptidyl transferase activity

of the 50S subunit catalyses the transfer of the fMet group from its tRNA in the P-site, to the amino group of the aminoacyl-tRNA, in the A-site. Next, translocation of mRNA and the newly formed peptidyl-tRNA from the A-site to the P-site occurs in a reaction involving elongation factor EF-G and the hydrolysis of GTP. The deacylated tRNA is moved to the E-site and then released. The A-site is now ready to accept another aminoacyl-tRNA and the cycle continues (Singer and Berg, 1991; Nierhaus *et al.*, 1993) (see Figure 4.3).

Termination takes place when the ribosome reaches one of the stop codons UAA, UAG, or UGA. Since no tRNA translates these codons, the polypeptidyl-tRNA remains in the P-site. In the A-site, releasing factor RF1 recognizes UAA and UAG while releasing factor RF2 recognizes UAA and UGA. A third termination factor, RF3, catalyzes cleavage of the polypeptide chain from the tRNA. The dissociation of mRNA and tRNAs from the 70S ribosome is promoted by releasing factor RF4 (Singer and Berg, 1991; Tate *et al.*, 1993).

1.1 INITIATION OF TRANSLATION IN *ESCHERICHIA COLI*.

1.1.1 THE 5'-END OF mRNA FORMS A COMPLEX WITH THE 30S RIBOSOMAL SUBUNIT

In 1969, advances in RNA sequencing technology made it possible for different groups to analyze the RNA sequences surrounding initiation codons of *Escherichia coli* (Steitz, 1969).

In these experiments, initiation complexes of mRNA, ribosomes, and fMet-$tRNA_f$ were treated with ribonuclease. The ribosome protected segment of the mRNA was found to contain 30-40 nucleotides with the AUG initiation codon positioned about 12 nucleotides downstream from the 3' end of the protected segment. These regions were called protein synthesis initiation regions (Steitz, 1979) or ribosome binding sites (RBS)(Steitz, 1969). The only well understood component of the RBSs was the start codon. The remaining elements were believed to be involved in binding mRNAs to the ribosomes.

It was known at the time that a specific binding of mRNA to the 30S ribosomal subunit occurred during the translation initiation step (Lodish, 1969; Steitz, 1969). However, it was not known how the ribosome discriminated against so many real and potential AUG start codons and selected the first methionine triplets only. It was obvious that the 5' untranslated region of mRNA possessed more functions than just carrying the initiator (AUG) codon. (Lodish, 1969; Steitz, 1969).

Experimental evidence for the interaction between RBSs and 16S rRNA were obtained by Steitz and Jakes (1975). They studied the initiation complex formed *in vitro* between *E.coli* ribosomes and a ^{32}P-labelled fragment derived from the beginning of the A-protein of coliphage R17. When the mRNA-ribosome complex was treated with colicin E3 (a nuclease that makes a single cut 49 nucleotides from the 3' end of 16S rRNA), the 3' end of 16S rRNA was released as a stable complex with the ^{32}P-labelled mRNA fragment.

Indirect evidence in favour of the involvement of the 3′ end of 16S rRNA in initiation of translation came from various other experiments. Firstly, the ribosomal protein S1, which is required for the binding of fMet-tRNA$_f$ to the mRNA-ribosome initiation complex (van Dieijen *et al.*, 1976), was shown to be located near the 3′ end of 16S rRNA (Czernilofsky *et al.*, 1975). Secondly, initiation factor IF3, (which is also required for initiation), was shown to bind to the 30S ribosomal subunit in a region close to the 3′ end of 16S rRNA, and in association with protein S1 (van Duin *et al.*, 1975). Finally, kasugamycin, a potent inhibitor of initiation, targeted the 3′ end of 16S rRNA (Helser *et al.*, 1971). To better elucidate the molecular mechanism of initiation of translation, Steitz and Steege (1977) studied the interaction of the 3′ end of 16S rRNA using two colicin E fragment-RBS complexes (Steitz and Steege, 1977). One complex, that contained coliphage R17 A protein RBS, was capable of forming 7 base pairs with the colicin E fragment. The other complex, an RBS from bacteriophage lambda, was capable of forming 9 base pairs with the colicin E fragment. Comparative thermal denaturation studies of the two complexes yielded melting temperature values of 32°C and 37°C for the coliphage R17 and lambda RBSs respectively. These values were consistent with the number of hypothesized base pairs formed in the two complexes. This observation further supported evidence of the interaction between the 3′ end of 16S rRNA and mRNA.

In their studies, Steitz and Steege (1977) used a fragment released from the 16S rRNA which might not even be accessible to the mRNA in the intact ribosomes.

To show the accessibility of the 3′ terminal domain of 16S rRNA for base pairing, Taniguchi and Weissmann (1979) inoculated intact ribosomes with a heterogeneous mixture of ^{32}P-labelled oligonucleotides derived from coliphage Qβ mRNA by RNase treatment. They observed selective binding to the ribosomes for oligonucleotides which were complementary to the 3′ end of 16S rRNA. It is important to note, however, that among the bound oligonucleotides none were found in any known RBS (Taniguchi and Weissmann, 1979).

If 16S rRNA interacts with the mRNA RBSs by its 3′ terminus during the initiation step, then the blocking of the 3′ end should either diminish or prevent the mRNA binding to the ribosome. It was shown by Taniguchi and Weissmann (1978) that oligonucleotides complementary to the 3′ end of 16S rRNA inhibited binding of radioactively labelled coliphage Qβ mRNA to *E.coli* ribosomes and the extent of inhibition was found to be proportional to the oligonucleotide used.

1.1.2 SHINE-DALGARNO HYPOTHESIS

In 1974, Shine and Dalgarno observed that the *E.coli* RBSs, whose structures were available at the time, contained a substantial part of the sequence 5′-GGAGGU [later called Shine-

Dalgarno (SD) sequence], 5′ to the AUG start codon (Shine and Dalgarno, 1974). The conservation of this sequences among different mRNAs suggested a specific role for this domain. Through stepwise degradation of the 3′ end of 16S rRNA, Shine and Dalgarno further showed that the 3′ end of 16S rRNA is complementary to the SD sequence (later called "SD binding site" or "anti-SD sequence"). Therefore, they suggested that formation of the mRNA-30S ribosomal subunit complex is likely to be due to the interaction between the 3′ end of 16S rRNA and the domain preceding the start codon in mRNA (Shine and Dalgarno, 1974).

Further evidence supporting the Shine-Dalgarno hypothesis came from mutational studies involving the SD sequence. Translation of *E.coli* mRNAs were greatly reduced by mutations that altered or deleted the SD sequence (Dunn *et al.*, 1978; Singer *et al.*, 1981; Schwartz *et al.*, 1981).

The first such point mutation identified was a transitional mutation (G-to-A) at the 5′ side of the AUG start codon in coliphage T7 gene 0.3 (Dunn *et al.*, 1978). This change eliminated a possible five-base pairing between the SD sequence of the mRNA and the SD binding site of 16S rRNA (Dunn *et al.*, 1978), an interaction previously proposed to be important for the initiation of 0.3 protein synthesis (Steitz and Jakes, 1975). This mutation caused the site of ribosome binding to shift about 15 bases to the 3′ side, centering on an internal AUG, and producing 0.4 protein, which is a distinguishable form of the 0.3 protein. A "pseudorevertant" was further identified that restored the

production of the original 0.3 protein. This suppressor mutation was found to lie adjacent to the original mutation creating a new SD sequence complementary to 16S rRNA (Dunn *et al.*, 1978).

An interesting report further supporting the SD interaction came from the studies by Chapon who was working with mutations in the *malT* gene of *E.coli* (Chapon, 1982). She isolated a mutation which led to an increase in *malT* gene expression. The *malT* gene is preceded by an SD sequence, which is somewhat unusual in that it is mainly complementary to nucleotides 8-12 from the 3' end of 16S rRNA. Sequences preceding other cistrons are usually complementary to nucleotides closer to the 3' end (Gold *et al.*,1981). In the identified mutant, there was an almost complete complementarity of the mRNA with a sequence extending from the fourth to the twelfth nucleotide of the 3' end of 16S rRNA, which was well within the range of distances found in other systems at the time (Gold *et al.*,1981). Thus, it was concluded that the SD sequence located at the beginning of the *malT* mRNA is normally rather weak and the mutation resulted in the formation of a much stronger binding site, allowing more efficient initiation of translation (Chapon,1982).

1.1.3 MUTATIONAL STUDIES RELATED TO THE 3'-END OF 16S rRNA ALSO SUPPORTED THE SHINE-DALGARNO HYPOTHESIS

Using a plasmid-borne specialized ribosome system, Jacob and co-workers constructed a single base mutation at position 1538 in *E.coli* 16S rRNA (Jacob *et al.*, 1987). This position is thought to

be involved in base pairing with mRNA during initiation of translation. The mutation was introduced in a 16S rRNA gene by M13 mutagenesis. The mutant gene was expressed under a temperature sensitive repressor where the shift in temperature resulted in production of mutant ribosomes amounting to up to 50% of the cellular ribosomes. Out of the proteins whose amounts were altered by the induction of these mutant ribosomes, only eight were identified which had genes that had been sequenced. The change in the production of all eight was compatible with the changes in the SD binding site (that is, the 3' end). For example, in the ATPase α-subunit mRNA, the mutation in the SD binding site replaces a G:C base pair with a G.U base pair and reduces its translation (GGAG to GGAG, where underlined bases are complementary to 16S rRNA). Conversely, methionyl-tRNA synthetase mRNA gains a base pair and its translation is increased (UAAGAAG to UAAGAAG) (Jacob *et al.*, 1987).

To further illustrate the interaction between mRNA and the 3' end of 16S rRNA, Hui and de Boer used a plasmid-borne ribosome system carrying a human growth hormone gene (Hui and de Boer, 1987). They constructed the system in such a way that an inducible subpopulation of modified ribosomes were directed to a single mutated mRNA species. This was accomplished by changing the SD sequence preceding the human growth hormone gene from 5'-GGAGG to 5'-CCTCC, which made the translation of the modified mRNAs by wild type ribosomes very inefficient. Further, the SD binding site at the 3' end of 16S rRNA was altered from 5'-CCTCC to 5'-GGAGG,

restoring its potential to base pair with the mutated human growth hormone mRNA. They observed that synthesis of the human growth hormone was dependent on the induction of the mutant ribosomes.

1.1.4 CHARACTERISTICS OF THE SHINE-DALGARNO SEQUENCE IN PROKARYOTIC mRNA

The Shine-Dalgarno sequence is defined as a conserved, purine-rich sequence located upstream from the start codon in prokaryotic mRNAs. It varies in length from about 3 to 8 nucleotides (5 nucleotides on average) which are complementary to the 3′ end of 16S rRNA. The efficiency of the SD sequence is determined by its ability for base pairing with the 3′ end of 16S rRNA. A three base pair interaction is often adequate when the most favored region (CUCC) of the 3′ end is involved. However, SD can also interact with neighbouring nucleotides (left or right) of the 5′-CUCC sequence, and in this less favorable situation, a three base pair interaction usually yields a low level of expression (Kozak, 1983). The extension of complementarity (up to 8 base pairs) between the SD sequence and the 3′ end of 16S rRNA results in an increased expression level. However, the complementarity above 8 base pairs is shown to have an opposite effect in some cases (Gold, 1988).

The consensus SD sequence (5′-AGGAGGU or 5′-AAGGAGGU) is derived from the nucleotide sequences of 5′ terminal untranslated regions of natural *E.coli* and bacteriophage mRNAs (Scherer *et al.*, 1980; Gold *et al.*, 1981). Although the consensus SD sequence has

not yet been found in a natural RBS (Ringquist *et al.*, 1993), it has been chemically synthesized (Jay *et al.*, 1980; Jay *et al.*, 1981) and its efficiency to initiate high levels of translation in *E.coli* cells have been experimentally proven (Rommens *et al.*, 1983; Jay *et al*, 1984a; Jay *et al.*, 1984b; Gigova *et al.*, 1989; Ivanov *et al.*, 1989).

1.1.5 SPACER REGION BETWEEN SD SEQUENCE AND THE START CODON

The original studies on the RBSs indicated that the average distance between the SD sequence and the AUG start codon (SD-AUG spacing) was about 7 nucleotides (Shine and Dalgarno, 1974; Gold *et al.*, 1981; Kozak, 1983). Although some variation from this value was tolerated, mutants in which the SD-AUG spacing was more than nine (Stroynowski *et al.*, 1982) or less than five (Singer *et al.*, 1981) nucleotides showed impaired translation. Such observations suggested that the proper distance between SD sequence and the start codon could be important in the process of initiation (Kozak, 1983).

It has been experimentally shown that variations in the SD-AUG spacing strongly affected the efficiency of translation of a variety of mRNAs (van Wezenbeek *et al.*, 1980; Guarente *et al.*, 1980; Thummel *et al.*, 1981; Gheysen *et al.*, 1982). Guarente *et al.* (1980) found that the optimum SD-AUG spacing for maximum expression of the rabbit β-globin gene in *E.coli* is 6 to 9 nucleotides. Any constructs with SD-AUG spacing of more than 9 or less than 6

nucleotides had a lower gene expression. Similarly, simian virus 40 small t antigen was shown to be expressed with the greatest efficiency when the SD-AUG spacing was about 10 nucleotides (Thummel *et al.*, 1981), while deviations from this value resulted in impaired translation.

1.1.6 OTHER FEATURES OF THE RIBOSOME BINDING SITE CONTAINING THE SD SEQUENCE

Statistical analyses have revealed that AUG triplets preceded by appropriately spaced "SD-like" sequences occur randomly throughout the *E.coli* genome (Gold *et al.*, 1981; Stormo, 1986). Also, the mRNA-rRNA complexes formed on behalf of the SD-anti-SD interaction have different stabilities depending on the length and base composition of the SD sequence (Gold *et al.*, 1981). If the SD-anti-SD is the only interaction accounting for the efficiency of initiation of translation, it is logical to expect differences in gene expression in *E.coli* as the melting temperature (MT) of the complex is varied (Gold *et al.*, 1981). Varying the MT from 20° to 40°C however, has been shown to have no significant difference in the efficiency of translation (Gold *et al.*, 1981). There are also examples of natural mRNAs devoid of an SD or even a 5′ untranslated sequence such as the C1 repressor proteins of phage lambda (Ptashne *et al.*, 1976; Shean and Gottesman, 1992), bacreriophage P2 gene *V* (Christie and Calendar, 1985), tetracycline resistance gene *tetR* (Klock and Hillen, 1986), and some RNAs of the phage 434 (Pirrotta,

1979). Although these genes are poorly expressed *in vivo*, their translation is not initiated by the SD sequence. These arguments suggest that in some cases the SD sequence is dispensable, which means that the RBS should not be identified with the SD sequence. The RBS is a much more complex structure possessing other features and properties contributing to its basic recognition function.

An established feature which may contribute to the recognition of the RBS by the ribosomes is the mRNA secondary (and/or tertiary) structures (Hall *et al.*, 1982). There are examples of poorly expressed mRNAs despite the presence of strong SD sequences in appropriate positions in relation to the initiation codon. In these cases either the SD sequence (Saito and Richardson, 1981), or the start codon (Iserentant and Fiers, 1980), or both (Johnston and Roth, 1981; Gheysen *et al.*, 1982) were found to be involved in stable helical structures which are sometimes capable of eliminating translation completely (Simons and Kleckner, 1983; McPheeters *et al.*, 1986). This concept is supported by mutational studies showing that the translation of mRNA was enhanced when the secondary structure in the translation initiation region was destabilized (Knight *et al.*, 1987).

To assess the relevance of the formation of a secondary structure to the process of translation initiation, Stanssens and co-workers (1985) used a series of structural variants of the RBS which directed the synthesis of human fibroblast interferon as well as the β-galactosidase gene (Stanssens *et al.*, 1985). The structural variants of the RBS were generated by altering the

nucleotide sequences 5′ to the SD sequence. They found that alterations 5′ to the SD sequence affected the rate of mRNA translation considerably. They also observed that the relative efficiency of the various 5′ untranslated nucleotide sequences depended on the downstream coding region, suggesting that secondary (and/or tertiary) structures are of major importance during initiation.

The lack of random distributions of nucleotides at several positions in the RBSs other than those in the start codon and the SD sequence, suggests that other regions in mRNA might also be involved in the recognition process (Gold *et al.*, 1981; Kozak *et al.*, 1983; Gold *et al.*, 1988). A pyrimidine in position -1 (that is, immediately preceding the AUG start codon of mRNA) is shown to promote the highest level of oligonucleotide binding to the ribosomes (Ganoza *et al.*, 1982), while a purine (specifically adenosine) is most effective in position +4 (Schmitt *et al.*, 1982). In some cases, the nature of the four nucleotides following the SD sequence (de Boer and Shepard., 1983) and the three nucleotides preceding the start codon (Hui *et al.*, 1984) has also been shown to have a considerable effect on the translation efficiency. However, no general conclusion has been drawn from such experiments (Hui *et al.*, 1984; Gold, 1988; McCarthy and Gualerzi, 1990).

A specific region, called downstream box or DB, which is located immediately downstream from the start codon, has also been shown to be important in translation efficiency (Peterson *et al.*, 1988; Sprengart *et al.*, 1990). Expression studies using a highly

active RBS of bacteriophage T7 gene 0.3, revealed that the efficiency of this RBS is strongly dependent on nucleotides +15 through +26 of the coding region which are complementary to nucleotides 1471 to 1482 of 16S rRNA. Similar sequences complementary to nucleotides 1469 to 1483 of 16S rRNA are also found in many other *E.coli* and bacteriophage mRNAs downstream of the start codon (Dunn and Studier, 1983; Sprengart *et al.*, 1990), and in some cases their effect as translational enhancers has been illustrated (Faxen *et al.*, 1991; Pohlner *et al.*, 1993).

It has also been demonstrated that certain sequences which do not resemble the SD sequence are capable of enhancing translation when present near the start codon of *E.coli* mRNAs (Sleat *et al.*, 1987; Olins and Rangwala, 1989; Loechel *et al.*, 1991; Helke *et al.*, 1993). These translational enhancers all possess sequences complementary to different parts of 16S rRNA. It is proposed that these regions specify a stimulatory interaction between the mRNA and 16S rRNA besides the SD interaction during the translation initiation step. The 5' untranslated leader sequence ("omega",Ω) of tobacco mosaic virus RNA genome is well documented to be one of the prokaryotic translational enhancers (Sleat *et al.*, 1987; Gallie *et al.*, 1987; Sleat *et al.*, 1988). When present at the 5' end of mRNA, omega significantly enhances translation in both eukaryotes and prokaryotes. A sequence called "epsilon"(ε), derived from the RBS of bacteriophage T7 gene *10*, is another example of translational enhancers when present near the start codon of mRNAs (Olins and Rangwala, 1989). Epsilon, which contains a 10

nucleotide sequence which is complementary to nucleotides 458-467 of 16S rRNA, enhances translation when present either 5′ or 3′ to the start codon (Olins and Rangwala, 1989). It has also been reported that a sequence derived from the RBS of the *tuf* gene in *Mycoplasma genitalium* is capable of enhancing the translation of mRNAs in *E.coli* (Loechel *et al.*, 1991). This sequence has the potential to form a base pair interaction with bases 1082-1093 of 16S rRNA.

1.1.7 NON SHINE-DALGARNO INITIATION OF TRANSLATION IN *ESCHERICHIA COLI*

Recently, it has been demonstrated that the 5′ untranslated leader sequence ("omega",Ω) of the tobacco mosaic virus (TMV) RNA genome is capable of not only enhancing but also mediating initiation of translation of various mRNAs in *E.coli* devoid of SD sequences. The efficiency of Ω to initiate translation was found to be lower than the consensus SD sequence and was variable depending on the reporter gene used (Ivanov *et al.*, 1992).

The lack of sequence similarities between the SD sequence and omega prevents the latter from interacting with the anti-SD site of 16S rRNA (Gallie *et al.*, 1987; Gallie and Kado, 1989; Ivanov *et al.*, 1992) which means that it initiates translation through an alternative non-canonical mechanism of interaction with the ribosome (Ivanov *et al.*, 1992).

Computer studies further revealed that a deletion derivative of omega, which showed the same effect on translation as omega (Gallie *et al.*, 1988; Ivanov *et al.*, 1992), had a sequence complementary to nucleotides 1344 to 1353 of 16S rRNA (Ivanov *et al.*, 1992). To illustrate whether the complementary base pairing between these two regions is responsible for the initiation of translation, three derivatives of omega were constructed (Ivanov *et al.*, 1995). One contained only the natural complementary region between omega and 16S rRNA; the second extended from nine to twelve base pairs complementary region and the third was made non-complementary to either the anti-SD site or to domain 1344-1353 in 16S rRNA. The efficiency of these constructs to initiate translation in *E.coli* cells was studied using two reporter genes, chloramphenicol acetyl transferase (CAT) and the human alpha one interferon ($hIF\alpha_1$). This study showed that the extension of the complementary region between omega and 16S rRNA resulted in an increase in the protein yield, whereas the destruction of the complementary region had an opposite effect. Based on these results it was suggested that omega might mediate initiation through base paring with nucleotides 1344-1355 (omega binding site) on *E.coli* 16S rRNA (Ivanov *et al.*, 1995).

In a recent report, Sprengart's laboratory showed that a downstream box (DB) complementary to nucleotides 1470-1482 of 16S rRNA was an independent translational signal in *E.coli* (Sprengart *et al.*, 1996). A highly active bacteriophage T7 gene *10* RBS which contains a DB complementary to nucleotides 1470-1482 of 16S rRNA

(anti-DB) was placed in front of a dihydrofolate reductase (d*hfr*) reporter gene. The functional activity of the DB was then analyzed by varying its complementarity to the anti-DB in the presence and absence of an SD interaction. It was observed that optimizing this complementarity from 10 to 13 base pairs strongly enhanced the translation regardless of the presence of an SD sequence. Thus, they proposed that the interaction between the DB and the anti-DB mediates initiation.

1.2 TRANSLATION INITIATION AND STRUCTURE OF THE 30S RIBOSOMAL SUBUNIT; HOW IS THE mRNA BINDING SITE ORGANIZED IN THE 30S RIBOSOMAL SUBUNIT?

Three-dimensional structure of the small ribosomal subunit has been extensively studied for more than one decade. The vast amount of data obtained during the years by biochemical, electron microscopy, immunological, X-ray diffraction, etc. studies have been summarized in the first three-dimensional models of the 30S ribosomal subunit published in 1988 (Brimacombe *et al.*, 1988; Stern *et al.*, 1988). Since that time the models of the small ribosomal subunit have been gradually developed and upgraded to include new data obtained by improved methodologies (Malhotra and Harvey, 1994; Brimacombe, 1995; Eeasterwood and Harvey, 1995; Frank, *et al.*, 1995; Noller *et al.*, 1995).

All models are based on the 16S rRNA structure used as a scaffold for attachment of the small ribosomal subunit proteins.

Primary structure of the 16S rRNA is well known and 45 helical (stem-loop) domains are recognized in its secondary structure (Brimacombe, 1991; 1992). Most of these domains are situated in the same way by the different models and all the models agree that the 30S subunit consist of two main parts, head and body (Figure 1.2). They are linked by a string of nucleotides (from the 16S rRNA) and can be easily separated after a mild RNAse treatment (Brimacombe, 1995). Some of the nucleotides of 16S rRNA are expressed on the surface and others are "buried" inside the 30S particle (Oakes and Lake, 1990).

There is a consensus between the models that the nucleotides 1399-1409 and 1492-1507 (belonging to the helical domain number 44) are involved in the organization of the two A and P sites of the *E.coli* ribosome (Eeasterwood and Harvey, 1995; Moazed and Noller, 1990; Doring *et al.*, 1994; Cunningham *et al.*, 1993). These two nucleotide regions (forming the so called 1400-1500 neighbourhood) together with the nucleotides from the helical domain number 11 (forming the 530 neighbourhood) are the main part of the *E.coli* ribosome decoding center.

The anti-SD site, responsible for binding mRNAs containing SD sequences during the initiation of translation, belongs to the last helical domain (number 45) in 16S rRNA. In the three-dimensional model it is found at the upper part of the body close to the cleft between the body and the head of the 30S ribosomal subunit (see Figure 1.2).

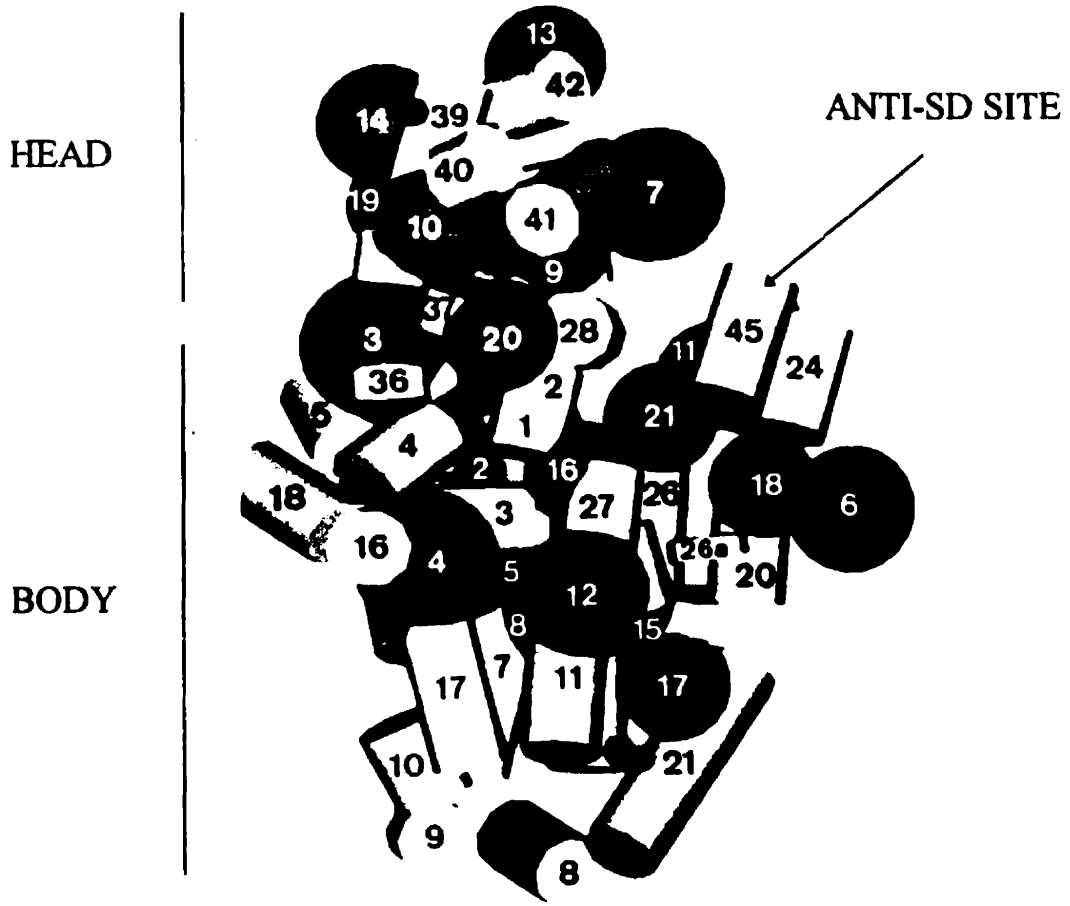


Figure 1.2: Schematic diagram illustrating the tertiary structure model of 30S ribosomal subunit. Ribosomal proteins are represented by dark spheres with white numbers, and 16S rRNA helical domains are represented by cylinders with black numbers. (Adapted from Brimacombe, 1995).

There are two types of nucleotides in 16S rRNA, evolutionarily conserved and variable. The mRNA-rRNA cross- linking studies have shown that most of the conserved nucleotides are located on the track of mRNA which is found on the three-dimensional model as a necklace between the head and the body of the 30S ribosomal subunit (see Figure 1.3 and McCarthy and Brimacombe, 1994; Brimacombe, 1995).

1.3 FOCUS

Literature data related to expression of leaderless mRNAs and mRNAs devoid of SD sequences in *E.coli* clearly indicates that in some cases the SD sequence is dispensable and that the canonical SD-anti-SD interaction might be bypassed during the translation initiation step. These data also show that the anti-SD site is not a unique mRNA binding site on the surface of the 30S ribosomal subunit and that other alternative mRNA binding sites should exist.

The putative "omega binding site" was found partly in the stem and partly in a single-stranded region of the helical domain # 43 in 16S rRNA. It is located in the three-dimensional model of 30S ribosomal subunit at the opposite side of the cleft in front of the anti-SD binding site (carried by the helical domain # 45).

The probable binding sites of the other non-SD translational initiators are so far not related to the model of 30S subunits. Also, for most of these sequences it is not clear whether they only enhance translation or if they initiate translation as well.

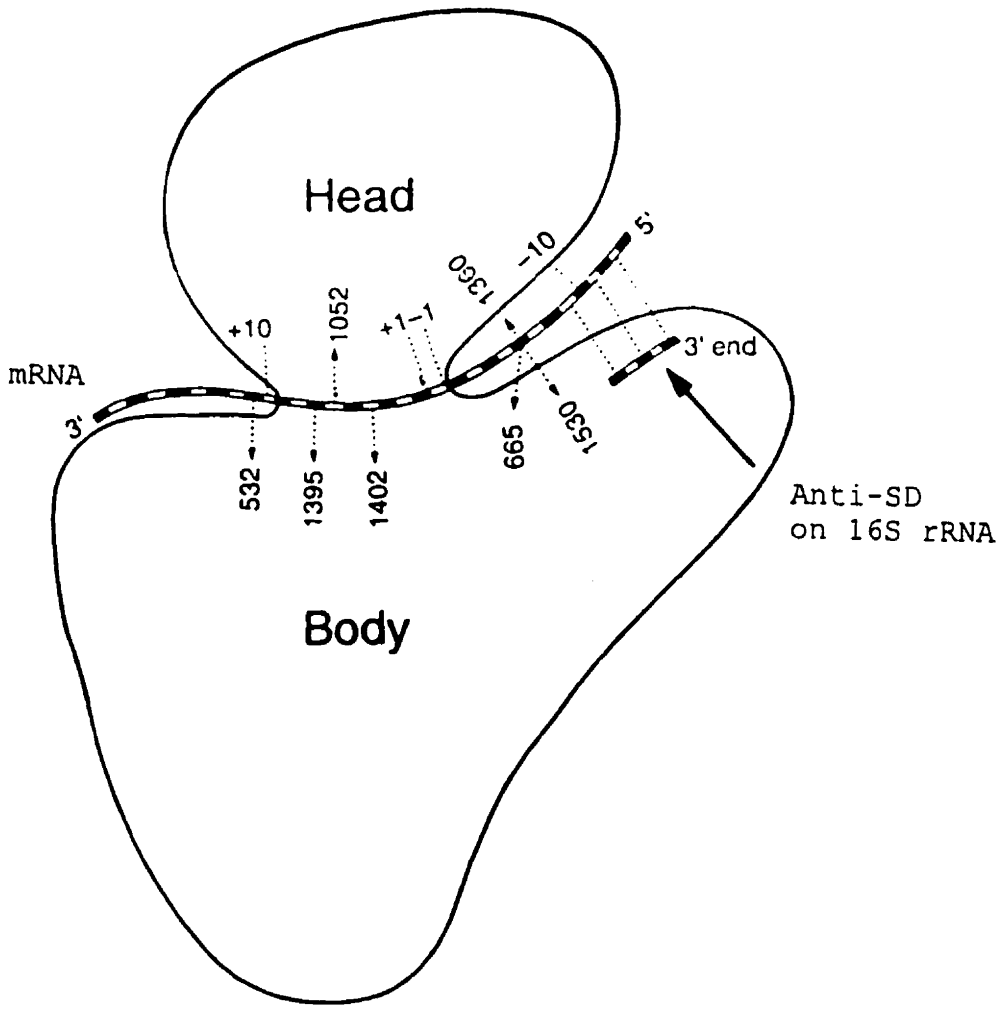


Figure 1.3: Schematic diagram illustrating the positioning of the mRNA track through the 30S ribosomal subunit. An mRNA molecule is shown together with the sites of cross-linking to 16S rRNA. (Adapted from McCarthy and Brimacombe, 1994).

1.4 OBJECTIVES

The general objective of the present thesis is to reveal new non-SD translational initiators and to shed light on the mechanism of their action as well as on the organization of their corresponding binding sites on the surface of the *E.coli* small ribosomal subunit.

To this end the following studies and experiments are designed:

1. Study of the capacity of other areas in domain # 43 of 16S rRNA to bind functional mRNA. The "omega" binding site is partly involved in the stem structure of the helical domain # 43 and it is about as efficient as the natural SD sequence of the chloramphenicol acetyl transferase (CAT) gene (Ivanov *et al.*, 1995). One can assume that if this sequence is located entirely in a single stranded region and on the surface of the small RNA subunit (30S) it would be even more efficient. For this reason, an oligonucleotide complementary to the loop region of domain # 43 is designed to be cloned in front of the ATG codon of the reporter gene chloramphenicol acetyl transferase (CAT) in order to compare its efficiency to initiate translation with that of the SD and omega sequence.

2. Study of the capability of nucleotide sequences from the helical domain # 42 (neighbour to # 43) to bind functional mRNA. An oligonucleotide complementary to the loop region of this domain

is designed to be cloned as above in order to check whether it is also involved in the organization of an alternative mRNA binding site.

3. Studies on the enhancer sequence "epsilon". It is not clear from the original study of Olins and Rangwala (1989) whether the "epsilon" sequence is a translational enhancer only or if it is also an independent initiator of translation. To answer this question three synthetic constructs are designed to be cloned in expression plasmids in place of the SD sequence. One of the constructs contains the natural "epsilon" sequence. In the second construct the "epsilon" is extended with 4 nucleotides complementary to the corresponding binding site in 16S rRNA and the third construct contains the "epsilon" sequence in combination with the consensus SD sequence.

4. Studies on the translation initiation capability of nucleotide sequences from 5' untranslated regions of plant viral RNAs [clover yellow mosaic virus (CYMV) and papaya mosaic virus (PMV)].

2. MATERIALS AND METHODS

2.1 GENERAL MOLECULAR TECHNIQUES

2.1.1 MINI-PREPARATION PROCEDURE FOR PLASMID DNA ISOLATION

Bacterial colonies transformed with various plasmids were inoculated individually into 15 ml culture tubes containing 4 ml of Luria-Bertani (LB) broth (1.0% bacto-tryptone, 0.5% yeast extract and 1.0% NaCl, pH 7.5) supplemented with 50 μg/ml ampicillin (Ap) and incubated at 37°C overnight with shaking.

To extract double-stranded plasmid DNA, a modified version of the Birnboim and Doly (1979) rapid alkaline lysis procedure was used. Cultures grown overnight were aliquoted into 1.5 ml microfuge tubes and pelleted at 16,000 g for 20 sec at 22°C. The supernatant was discarded and pellets were resuspended in 150 μl solution I [50 mM D-glucose, 10 mM ethylenediaminetetraacetic acid (EDTA), 25 mM Tris-HCl pH 8.0] containing 3 mg/ml lysozyme, and incubated on ice for 5 min. To complete cell lysis, 300 μl of solution II [0.2 N sodium hydroxide (NaOH), 1% sodium dodecylsulfate (SDS)] were gently added to each tube and the solutions were mixed by gentle inversion and incubated on ice for 5 min. Finally, 230 μl of solution III [3 M sodium acetate (NaAc) pH 4.8] were added. The tubes were inverted slowly several times and incubated on ice for 10 min. All insoluble material was pelleted by centrifugation at 16,000 g for 15 min at 4°C. The supernatant was transferred into a new tube and plasmid DNA was precipitated by the addition of 1 volume of cold isopropanol followed by incubation at -20°C for 40 min. Insoluble plasmid DNA

was collected by centrifugation at 16,000 g for 10 min at 4°C in an Eppendorf microfuge. Pellets were redissolved in 200 μl TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

Plasmid DNA was further purified through ammonium acetate (NH_4AC) precipitation of genomic DNA and proteins. An equal volume of 5 M NH_4AC (200 μl) was added to the solution and incubated on ice for 30 min. The mixture was centrifuged at 16,000 g for 15 min at 4°C. DNA was precipitated from the supernatant with the addition of 2.5 volumes of 95% ethanol (EtOH) and incubation at -20°C for 40 min, and collected by centrifugation at 16,000 g for 5 min at 4°C. Pellets were washed with 70% EtOH for 5 min, air dried for 30 min, then redissolved in 50 μl of TE^{-1} buffer (1 mM Tris-HCl, 0.1 mM EDTA, pH 8.0).

Plasmid DNA was then analyzed by agarose gel electrophoresis. The agarose gels were made using 1X TBE buffer (0.1 M Tris-HCl, 0.1 M boric acid, 2 mM EDTA, pH 8.3) which was also used as running buffer. Gels were prestained with ethidium bromide (EtBr) (0.5 μg/ml). Two μl of DNA solution were mixed with 3 μl dH_2O and 1 μl 6X loading dye [30% glycerol (v/v), 0.25% xylene cyanol FF (w/v)]. The mixture was electrophoresed through a 1% agarose (w/v) gel at 100 mA constant current until the dye had moved about 2 cm from the wells. DNA bands were visualized under ultraviolet light (300 nm). The quantity of plasmid DNA was estimated visually by the intensity of the bands compared to samples of known concentration.

2.1.2 PHENOL/CHLOROFORM EXTRACTION

The volume of DNA solution to be deproteinized was brought up to at least 50 µl [by addition of double-distilled water (dH_2O)]. DNA was extracted with 1 volume of phenol saturated with 0.1 M Tris-HCl pH 8.0 containing 0.3 M sodium chloride (NaCl) and NaCl was added to the aqueous phase to a final concentration of 0.3 M. The mixture was vortexed and then centrifuged at 16,000 g for 4 min (22°C). The top phase of the mixture (containing DNA) was removed and then one volume of chloroform was added. The mixture was vortexed and centrifuged as before. The top phase was removed and extracted with chloroform twice more. Plasmid DNA was precipitated by addition of 2.5 volumes of 95% EtOH, incubated at -20°C for 40 min, and collected by centrifugation at 16,000 g for 5 min at 4°C. DNA pellets were dissolved in different volumes of TE^{-1} as required.

2.1.3 TRANSFORMATION OF *ESCHERICHIA COLI*

Preparation and transformation of competent cells of *E.coli* LE392 was done according to Maniatis *et al.* (1982). Four ml of LB broth were inoculated with 5 µl from an *E.coli* LE392 glycerol stock (see below) and incubated overnight in a 37°C shaker at 300 RPM. The next day, 30 ml of LB broth were inoculated with 300 µl of this overnight culture and incubated at 37°C with shaking until an optical density of 0.4-0.6 at 590 nm was obtained (about 2 hrs). The cells were cooled on ice for 10 min and centrifuged at 1,500 g for 6 min.

The cells were resuspended by gentle vortexing in half of the initial culture volume (15 ml) of ice-cold 50 mM calcium chloride ($CaCl_2$) and kept on ice for 20 min. This was followed by a second centrifugation as above. The pelleted cells were resuspended by gentle vortexing in 2 ml of 50 mM $CaCl_2$ and kept at 4°C to be used the next day. Cells were now "competent".

Five to ten ng of plasmid DNA (1-2 μl) or 100-200 ng of ligation reaction mixture (10-30 μl) were added to a prechilled 1.5 ml microfuge tube. After the addition of 200 μl of competent cells the mixture was gently vortexed. After incubating the cells on ice for 20 min, they were heat-shocked for 90 sec in a 42°C water bath. One ml of LB broth was then added and the tubes were incubated at 37°C for 1 hr. The cells were pelleted in a microfuge for 10 sec at 12,000 g (supernatant was discarded) and then resuspended in 500 μl of LB broth and incubated at 37°C for 30 min. The cells were then vortexed and spread onto two 2X YT agar plates [1.6% (w/v) tryptone, 1.0% (w/v) yeast extract, 1.0% (w/v) NaCl and 1.5% (w/v) Bacto-agar] containing 50 μg/ml Ap. Cells were divided so that 25% were on one plate and 75% on the other. The plates were incubated overnight at 37°C in an inverted position.

2.1.4 GLYCEROL STOCKS OF TRANSFORMED BACTERIA

Bacteria transformed with desired plasmids were stored at -80°C, as follows: four ml of LB broth containing 50 μg/ml Ap were inoculated with the desired bacteria overnight at 37°C with shaking

at 300 RPM. Two hundred μl of 60% sterile glycerol was added to 400 μl of the overnight culture in a 1.5 ml microfuge tube and vortexed. The mixture was stored at -80°C.

2.1.5 OLIGONUCLEOTIDE SYNTHESIS AND PURIFICATION

Oligonucleotides were made with a Biosearch Cyclone DNA Synthesizer by automated solid phase phosphoramidite based synthesis (Sinha *et al.*, 1984). Ammonium hydroxide (NH_4OH) was used to remove the synthesized oligonucleotides from the column by the following procedure. Two 1 ml disposable syringes, one of which was filled with 30% NH_4OH, were inserted onto each end of the column. The NH_4OH was passed from one syringe to the other through the column at least 15 times. The column, including the syringes, was then covered with Saran wrap and incubated at 37°C for 30 min. The NH_4OH now containing the oligonucleotide was forced into one syringe which was then removed from the column and the NH_4OH was transferred into a 7 ml screw capped glass vial. The bottle was sealed tightly with parafilm and incubated at 60°C overnight. The lid was removed and the bottle was incubated at 24°C for 30 min to allow some NH_4OH to evaporate. One hundred μl aliquots were transferred to 1.5 ml microfuge tubes and spun in a Speedvac concentrator for 90 min on medium heat at -29 mm Hg. The pellets were serially resuspended in a final volume of 400 μl TE buffer and the supernatant was transferred to a new tube where the oligonucleotide was precipitated with 0.2 M potassium acetate (KAc) and 2 volumes of isopropanol at -

20°C for 30 min. The DNA was pelleted by centrifugation at 16,000 g for 5 min, rinsed with 70% EtOH, air-dried and resuspended in 100 µl dH_2O. Spectrophotometric readings were taken at 260 nm to calculate the concentration of the oligomer.

Oligomers designed to be cloned into an expression cassette were purified by thin layer chromatography as described by Gait (1984). The oligonucleotide dissolved in dH_2O was applied to a silica plate (Merck: silica gel 60 F-254) about 2 cm from the bottom using a Pasteur pipette drawn to a fine tip. The plate was then placed in a sealed chamber containing 45% n-propanol and 40% concentrated NH_4OH. The chamber was left at 24°C until the solvent containing the oligomer had migrated to within 1 cm from the top of the plate (about 4 hrs). The plate was then removed from the chamber and allowed to air-dry for 15 min. The intact oligomer was visualized under ultraviolet light (300 nm) and traced with a pencil. Silica containing the oligonucleotide was scraped from the plate and collected in a 1.5 ml Eppendorf tube. To elute the oligomer from the silica, 200 µl of dH_2O was added and the mixture was vortexed. Silica was pelleted by centrifugation at 16,000 g for 5 min, the aqueous layer was transferred to another tube and the silica was re-extracted twice more with 100 µl dH_2O each time. The collected liquid was dried using a Speedvac concentrator as before. The pellet was redissolved in 200 µl of dH_2O and centrifuged at 16,000 g for 3 min. The supernatant was phenol/chloroform extracted and DNA was precipitated with 0.2 M KAc and 2 volumes of isopropanol

at -20°C for 30 min. The oligomer was collected by centrifugation for 5 min, redissolved in 100 µl of dH_2O and its concentration was determined as before.

2.1.6 ISOLATION OF DNA FRAGMENTS FROM AGAROSE GEL

2.1.6.1 TRAPPING

DNAs digested with various restriction enzymes were electrophoresed on a 1% agarose gel until separation of bands was achieved. The level of running buffer was lowered to cover the sides of the gel but not the top surface. A block of agarose was cut out in front of the desired band producing a small well. A small piece of dialysis membrane with a 12,000 Da cut-off was boiled in water for 20 min and placed in the well against the wall closest to the anode electrode. The well was filled with TE buffer and the gel was electrophoresed until the desired fragment was released into the well (about 10 min). The current was reversed for 10 sec and the buffer from the well containing the DNA was transferred into a 1.5 ml Eppendorf tube. DNA was extracted by phenol/chloroform and the DNA was precipitated as outlined earlier.

2.1.6.2 FREEZE AND THAW

DNA digests were electrophoresed on a 1% agarose gel until the separation of desired bands was obtained. A block of the agarose gel carrying the band of interest was cut out and placed in a 0.5 ml

microfuge tube which contained a small quantity of siliconized glass wool, and had a needle size hole at the bottom. The tube was covered with parafilm and placed in a 1.5 ml microfuge tube. The agarose block was frozen and thawed by placing the tubes in -80°C for 10 min, followed by incubation at 37°C for 15 min. The process of freezing and thawing was repeated three times. The buffer held within the agarose block was then forced through the glass wool by centrifugation at 10,000 g for 10 min. A solution containing the DNA was collected in the larger tube at the bottom. The solution was extracted by phenol/chloroform and the DNA precipitated as mentioned earlier.

2.1.7 OLIGOMER HYBRIDIZATION

Five μg of each of the two complementary synthesized oligonucleotides which were purified by thin layer chromatography were mixed with a 50 μl volume of sterile ddH_2O containing 10 mM magnesium chloride ($MgCl_2$) in a 0.5 ml microfuge tube. The tube was sealed with parafilm and heated to 94°C in a water bath for 3 min. The water bath was removed from the heat and allowed to cool to room temperature (about 5 hrs) while stirring and then placed in 4°C for another 2 hrs.

2.1.8 COLONY HYBRIDIZATION

2.1.8.1 DNA

In order to identify plasmids carrying the DNA fragment of interest, large scale screening of colonies was carried out according to a modification of the original colony hybridization procedure described by Grunstein and Hogness (1975), using γ-^{32}P ATP labelled oligonucleotides as probes (Ivanov and Gigova, 1986).

Individual colonies were picked with sterile toothpicks and inoculated onto a nitrocellulose filter disc (BA85; Schleicher and Schuell) which had been previously placed on a 2X YT agar plate containing 50 μg/ml Ap. A separate 2X YT agar plate (with Ap) was also inoculated as a "master" plate.

All plates were incubated, inverted, at 37°C overnight. Filters carrying the bacterial colonies (facing up) were transferred in succession to petri dishes containing Whatmann 3MM filter paper discs soaked in one of four solutions (A-D) for 10 min each; solution A (10% SDS), solution B (0.5 M NaCl), solution C (1 M Tris-HCl pH 7.5, 1.5 M NaCl) and solution D (same as solution C). Next the filters were placed in solution E [6X SSC (0.15 M NaCl and 0.015 M sodium citrate, pH 7.0), 10% formaldehyde], incubated at 60°C for 20 min, air dried on a sheet of Whatmann 3MM filter paper for 20 min at room temperature, and baked for 1 hr at 80°C *in vacuo*.

2.1.8.2 RNA

Preparation of RNA colony hybridization filters was similar to that of DNA with the exception of denaturing steps. RNA filters were placed in solution A (see above), then transferred directly to solution F (3X SSC, 10% formaldehyde) for another 10 min, followed by incubation at 60°C for 20 min. The filters were then air dried and baked as described in section 2.1.8.1.

2.1.8.3 PREPARATION AND PURIFICATION OF ^{32}P-LABELLED OLIGONUCLEOTIDE PROBES

Oligonucleotide probes were end-labelled by the method of Ivanov and Gigova (1986) by adding 10 μg of the oligonucleotide in a 100 μl final reaction volume containing 1X T4-polynucleotide kinase buffer [0.1 M Tris-HCl pH 8.0, 10 mM $MgCl_2$, 5 mM dithiothreitol (DTT)] supplied with 10 μl [γ-^{32}P] ATP (ICN; specific activity of 650 Ci/mmol) and 20 units of T4-polynucleotide kinase (New England Biolabs). The mixture was incubated at 37°C for 1 hr. To purify the oligomer, a DEAE-cellulose column was constructed. The tip of a 1 mL pipette tip (disposable) was cut off leaving a hole 3 mm in diameter. A small quantity of siliconized glass wool was loosely packed into this tip and rinsed with 95% EtOH and then dH_2O. One hundred μl of Sephadex G-50 were added to the column and 150 μl of DEAE-cellulose were layered on top. To pre-equilibrate the contents, 400 μl of 0.2 M NaCl was run through the column. The

prepared probe was mixed with 200 µl 0.2 M NaCl and this mixture was then layered onto the column. Three 200 µl washes with 0.2 M NaCl were used to elute any unincorporated [γ^{-32}P] ATP by gravity. The oligonucleotide probe was recovered from the column with 600 µl 1 M NaCl and its total radioactivity was monitored using a scintillation counter.

2.1.8.4 PREHYBRIDIZATION AND HYBRIDIZATION OF NITROCELLULOSE FILTERS

DNA and RNA filters were prehybridized in SET/SDS solution (100 mM Tris-HCl pH 7.5, 1 M NaCl, 1 mM EDTA, 0.1% SDS) supplemented with 100 µg/ml homo-mix I, for 1 hr at 60°C. To prepare homo-mix I as a blocking agent, 2 g of total yeast RNA [Boehringer-Mannheim] were hydrolysed overnight at 37°C in a 10 ml volume containing 1.7 ml 5 N potassium hydroxide (KOH). The mixture was brought to pH 7 with 1 N hydrochloric acid (HCl) and diluted to 20 ml with dH_2O to give a 100 mg RNA/ml stock.

Radioactively labelled oligonucleotide probes were prepared and purified as described in section 2.1.8.3. The prehybridization solution was replaced with fresh SET/SDS containing 100 µg/mL homo-mix I and 10 µg of labelled probe (10^6-10^7 CPM/ml). Hybridization was carried out overnight in a shaker bath. The hybridization temperature (T_h) was calculated using the formula $T_h = [4(G+C)+2(A+T)]-5°C$ where G, C, A and T are the number of corresponding bases in the probe (Meinkoth and Wahl, 1984). The

filters were washed twice in SET/SDS for 10 min at (T_h-5)°C, then once for 5 min at T_h. Filters were air dried for 1 hr, covered with Saran wrap, then exposed to X-ray film (Kodak XAR5) between intensifying screens at -70°C for 2 hrs.

2.1.9 SEQUENCING OF DOUBLE STRANDED DNA TEMPLATES

Sequencing reactions were performed using the modified T7 polymerase (Sequenase® Version 2.0; USB) (Tabor and Richardson, 1987) and (α-^{35}S) dATP (Biggin *et al.*, 1983). All sequencing reactions used EcoF primer (5'-ATAGGCGTATCACGAGG) which is complementary to a sequence 5' to the promoter in pBRP1TC plasmid. Two μg of plasmid DNA, purified by alkaline lysis mini-prep method and phenol-chloroform extracted, dissolved in 18 μl of dH_2O and to this 2 μl of fresh 2 N sodium hydroxide (NaOH) were added to denature the plasmid and degrade RNA. The mixture was incubated at 37°C for 5 min followed by neutralization with 14 μl 5 M NH_4Ac. DNA was precipitated by addition of 100 μl 95% EtOH and incubation at -80°C for 15 min and centrifugation at 16,000 g for 5 min at 4°C. The pellet was washed twice with 70% EtOH, air dried and resuspended in TE buffer with a total volume of 10 μl containing 2μl 5X Sequenase® reaction buffer (200 mM Tris-HCl pH 7.5, 250 mM NaCl and 100 mM $MgCl_2$) and 50 ng primer. The primer was annealed to the template plasmid DNA by incubating the reaction mixture at 94°C for 3 min followed by quick-cooling on ice for 5 min. The following reagents were added to the annealed DNA mixture in this order; 1 μl 0.1 M

DTT, 1 μl manganese buffer (0.15 M sodium isocitrate and 0.1 M manganese chloride), 2 μl deoxynucleotides (dGTP mix) labelling mixture (prediluted 1:5 with dH_2O), 0.5 μl [α-^{35}S] dATP (1000 Ci/mmol; Amersham) and 2 μl of Sequenase® diluted 1:8 (3.25 units total) in enzyme dilution buffer (10 mM Tris-HCl pH 7.5, 5 mM DTT and 0.5 mg/ml BSA). The reaction mixtures were incubated at 24°C for 6 min. Then 3.5 μl of labelling reaction was transferred to each of 4 tubes containing 2.5 μl termination mixture [dideoxyadenosine triphosphate (ddATP), dideoxyguanosine triphosphate (ddGTP), dideoxycytidine triphosphate (ddCTP) or dideoxythymidine triphosphate (ddTTP)] which had been pre-warmed for 1 min at 37°C. Termination reaction mixtures were incubated at 37°C for 6 min and stopped by the addition of 4 μl of stop solution [95% formamide (v/v), 20 mM EDTA, 0.05% bromophenol blue (w/v) and 0.05% xylene cyanol FF (w/v)] and the tubes were placed on ice. Samples were heated at 94°C for 3 min and quick-cooled on ice immediately prior to loading onto an 8% acrylamide (19:1 acrylamide: bis-acrylamide)/8 M urea wedge gel. The gel was run at 65 watts (constant power) for 2-3 hrs then fixed in 10% EtOH/10% acetic acid for 1 hr. The gel was then transferred to 1 MM Whatman filter paper and dried under vacuum at 80°C for 1 hr, followed by 48 hrs exposure to X-ray film (Kodak XAR5).

2.1.10 NUCLEOTIDE SEQUENCE COMPLEMENTARITY SEARCH

A computer search for sequence complementarity between synthetic oligonucleotides and 16S rRNA of *E.coli* was performed using the programme SEQAID.

2.1.11 PREDICTING SECONDARY STRUCTURES OF THE mRNAs

The programme PCFOLD was used to generate computer predicted secondary structure of CAT mRNA carrying different sequences for initiation of translation.

2.2 CONSTRUCTION OF THE pBRP1'Tc EXPRESSION CASSETTE

The pBRP1Tc plasmid was digested with *Eco*RI and *Hin*dIII, removing the P1 promoter and the SD sequence. Digested DNA was then ligated to double stranded PP1' oligomer (see Appendix for the sequence) and then digested with *Eco*RI prior to transformation in *E.coli* LE392 cells (see Figure 2.1).

2.2.1 *Eco*RI AND *Hin*dIII DIGESTION OF pBRP1Tc CONSTRUCT

One μg of the pBRP1Tc construct was digested in a 50 μl reaction mixture containing 1X NEBuffer 2 (10 mM Tris-HCl pH 7.9, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT), 10 units of *Eco*RI (NEB) and 10 units of *Hin*dIII (NEB) for 4 hrs at 37°C. Digested plasmid DNA was

deproteinized by phenol/chloroform extraction and precipitated with 0.3 M NaCl and 2.5 volumes of EtOH at -20°C for 2 hrs. DNA was recovered by centrifugation at 16,000 g at 4°C for 5 min. The digested DNA pellet was resuspended in 25 μl of TE^{-1} and its concentration estimated by agarose gel electrophoresis.

2.2.2 LIGATION REACTION WITH PP1' OLIGOMER AND DEACTIVATION OF LIGASE

Digested DNA (500 ng) was ligated to 1 μg of double-stranded PP1' oligomer in a 20 μl final reaction volume containing 1X T4 DNA ligase buffer [50 mM Tris-Hcl pH 7.8, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 μg/ml bovine serum albumin (BSA)] and 15 WEISS units of T4 DNA ligase (NEB). The ligation reaction was incubated at 16°C for 3 hrs. T4 DNA ligase was inactivated by incubation at 65°C for 20 min.

2.2.3 DIGESTION OF THE PLASMID PRIOR TO TRANSFORMATION

Prior to transformation, the pBRP1'Tc construct was digested with *Eco*RI through supplementing the ligation reaction mixture with NaCl (50 mM final concentration), and 20 units of *Eco*RI (NEB). The digestion reaction mixture was incubated at 37°C for 1 hr. Half of the reaction mixture (10 μl) was transformed into competent *E.coli* LE392 cells.

2.2.4 SCREENING FOR THE MODIFIED pBRP1'Tc CONSTRUCT

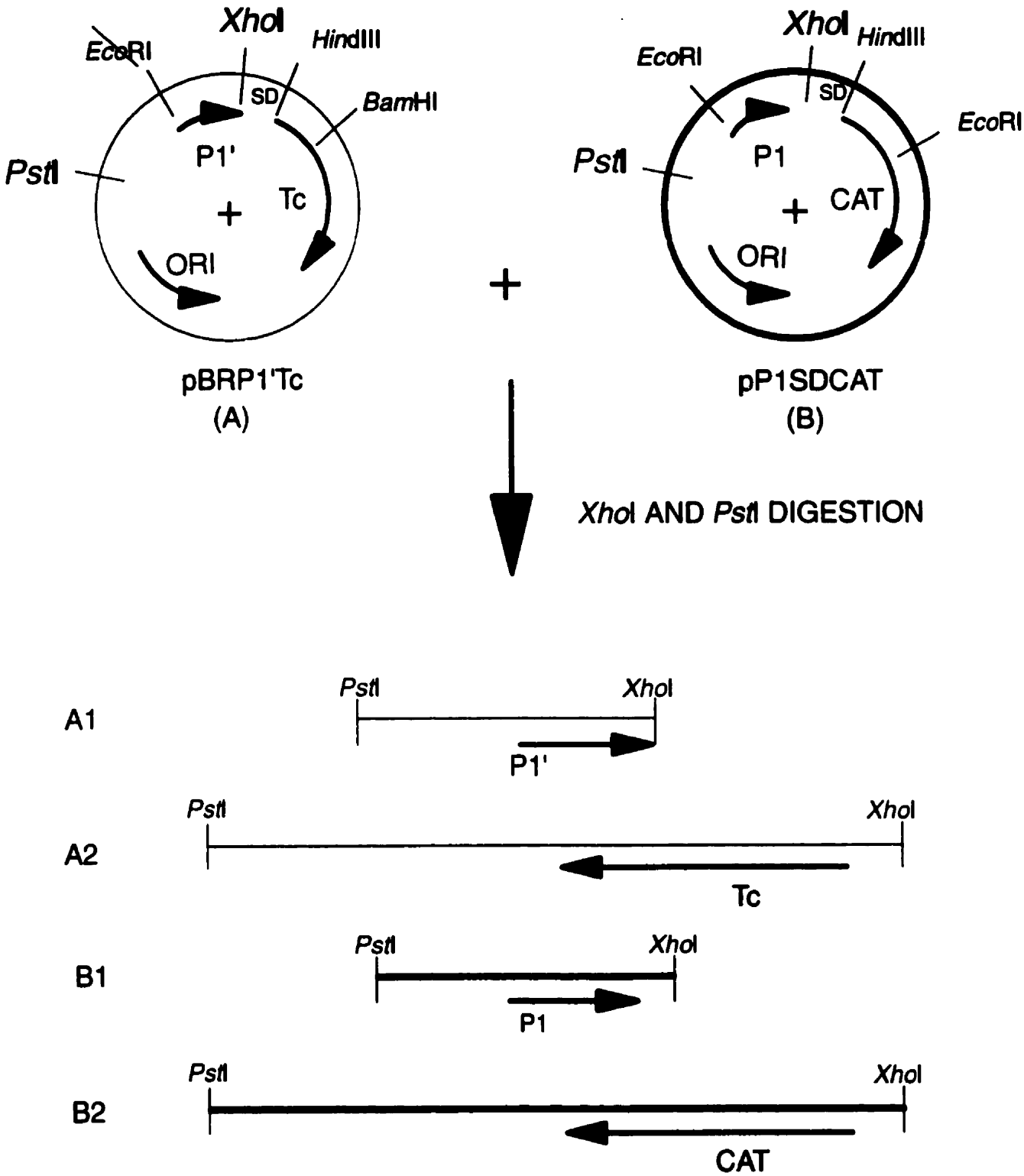
After miniprepping, modified plasmids (designated pBRP1'Tc) were first screened through *Eco*RI digestion. One μg of individual plasmids was digested in a 50 μl reaction volume containing 1X NEBuffer 2 and 10 units of *Eco*RI (NEB) and incubated at 37°C for 1 hr. The plasmids were run individually on a EtBr-prestained 1% agarose gel. Modified plasmids were identified by their resistance to the restriction digestion and then verified through dideoxynucleotide sequencing.

2.2.5 REPLACING THE TETRACYCLINE RESISTANCE GENE BY A CHLORAMPHENICOL ACETYL TRANSFERASE REPORTER GENE

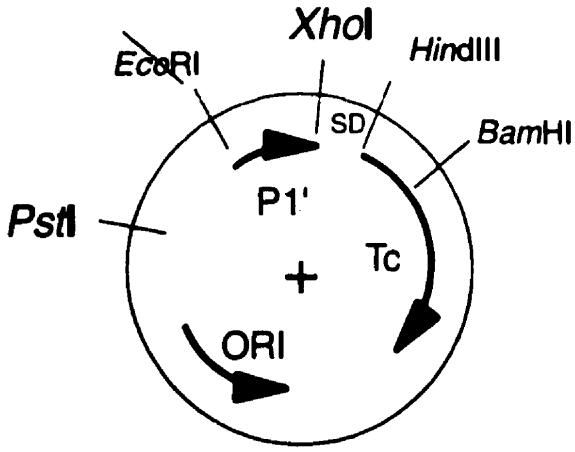
A mixture of pBRP1'Tc and pP1SDCAT plasmids was digested with *Pst*I and *Xho*I restriction enzymes (see Figure 2.2). Digested DNA was then ligated producing four different plasmids, not including the polymers. DNA was digested with *Bam*HI prior to transformation in order to prevent the transformation of the plasmids carrying tetracycline resistance (Tc) gene.

2.2.5.1 DIGESTION REACTION OF pBRP1'Tc

Purified pBRP1'Tc (1 μg) which served as the donor of the P1' fragment, and 5 μg of pP1SDCAT (a modified pBR322 plasmid carrying chloramphenicol acetyl transferase gene) which served as the vector,

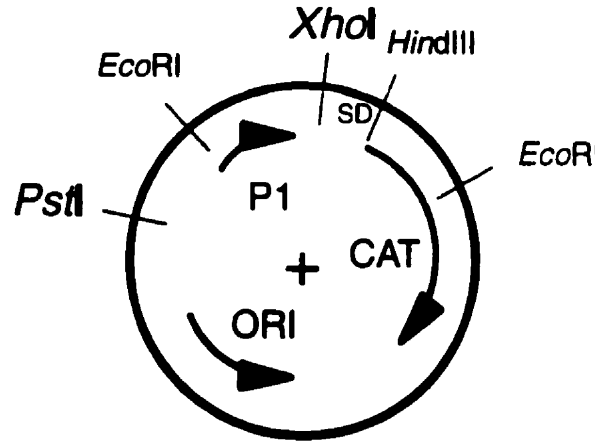

XhoI
EcoRI
HindIII
SD
BamHI
PstI
P1'
Tc
ORI
pBRP1'Tc
(A)
XhoI
EcoRI
HindIII
SD
EcoRI
PstI
P1
CAT
ORI
pP1SDCAT
(B)
XhoI AND PstI DIGESTION
A1
PstI
XhoI
P1'
A2
PstI
XhoI
Tc
B1
PstI
XhoI
P1
B2
PstI
XhoI
CAT

LIGATION OF THE FRAGMENTS



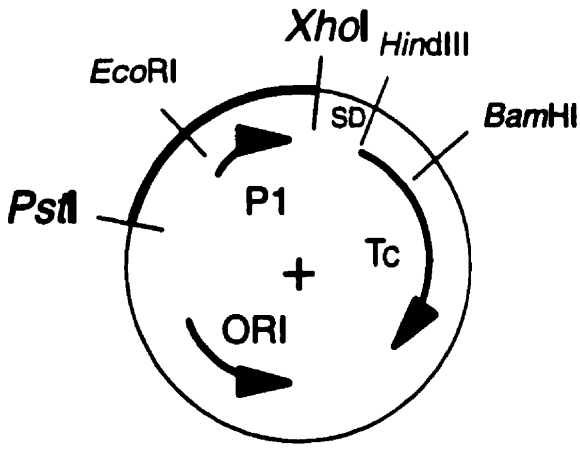
PLASMID 1 (A1+A2)

(pBRP1'Tc)

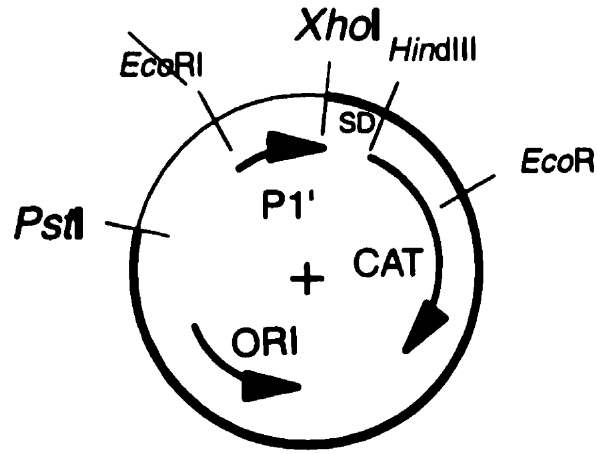


PLASMID 2 (B1+B2)

(pP1SDCAT)



PLASMID 3 (B1+A2)



PLASMID 4 (A1+B2)

(pP1'SDCAT)

*Bam*HI DIGESTION AND
TRANSFORMATION

Figure 2.2: Schematic diagram illustrating the construction procedure of the expression cassette. Plasmids pBRP1'Tc (designated A) and pP1SDCAT (designated B) were digested with *Pst*I and *Xho*I, producing a mixture of four DNA fragments (A1, A2, B1, and B2). The DNA fragments were ligated to generate four types of plasmids (not including the polymers) numbered 1 to 4. Plasmids 1 (A1+A2) and 2 (B1+B2) are the parental plasmids pBRP1'Tc and pP1SDCAT respectively. In plasmid 3 (B1+A2), the P1 promoter controls expression of the Tc resistance gene. Plasmid 4 (A1+B2, pP1'SDCAT) carries the CAT gene under the control of the promoter derived from pBRP1'Tc (designated P1'), and was used as the expression cassette. The ligated plasmids were digested with *Bam*HI prior to transformation in order to prevent the transformation of plasmids 1 and 3.

were digested to completion at 37°C for 4 hrs in a 75 μl reaction volume containing 1X NEBuffer 2 (10 mM Tris-HCl pH 7.9, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT), 100 μg/ml BSA, 60 units of *Pst*I (NEB) and 60 units of *Xho*I (NEB). Digested plasmid DNA was deproteinized by phenol/chloroform extraction, and precipitated with 0.3 M NaCl and 2.5 volumes of EtOH at -20°C for 2 hrs. DNA was recovered by centrifugation at 16,000 g for 5 min at 4°C. The digested DNA pellet was resuspended in 40 μl of TE^{-1} and its concentration was estimated by agarose gel electrophoresis.

2.2.5.2 LIGATION REACTION AND DEACTIVATION OF LIGASE

Digested DNA (2 μg) was ligated in a 40 μl final reaction volume containing 1X T4 DNA ligase buffer and 20 WEISS units of T4 DNA ligase (NEB). The ligation reaction mixture was incubated at 16°C for 4 hrs and then at 24°C for 1 hr. T4 DNA ligase was inactivated by incubation at 65°C for 20 min.

2.2.5.3 DIGESTION OF THE PLASMID PRIOR TO TRANSFORMATION

Prior to transformation, the 4 possible constructs were digested with *Bam*HI by supplementing the ligation reaction mixture with NaCl, to a final concentration of 50 mM, and 20 units of *Bam*HI (NEB). The digestion reaction was incubated at 37°C for 1 hr, and half of this reaction mixture was transformed into *E.coli* LE392 cells.

2.2.5.4 SCREENING FOR THE EXPRESSION VECTOR pP1'SDCAT

Modified plasmids carrying the chloramphenicol acetyl transferase (CAT) reporter gene (designated pP1'SDCAT) were initially screened on the basis of their antibiotic resistance. Individual colonies were picked with sterile toothpicks and each was inoculated onto a 2X YT agar plate containing 50 μg/ml Ap. Duplicates of the colonies were allowed to grow on two 2X YT agar plates; one containing 100 μg/ml chloramphenicol (Cm) and the other supplied with 20 μg/ml tetracycline (Tc). All plates were incubated at 37°C overnight.

The plasmids isolated from the colonies which were resistant to both Ap and Cm but not Tc were characterized through *Eco*RI digestion as before (see 2.2.3). The plasmids carrying an *Eco*RI site were selected and further verified through dideoxynucleotide sequencing.

2.3 PLACING THE CHLORAMPHENICOL ACETYL TRANSFERASE REPORTER GENE UNDER THE CONTROL OF Ω42, Ω43 OR EPSILON OLIGOMERS

The pP1'SDCAT plasmid was digested with *Xho*I and *Hin*dIII, removing the consensus SD sequence. Hybridized double-stranded Ω42, Ω43 or epsilon oligomers (see Appendix for sequences) were then ligated individually into the digested plasmid. Ligated DNA was digested with *Xho*I and *Hin*dIII prior to transformation in order to prevent the transformation of cells with the background plasmid (pP1'SDCAT) (see Figure 2.3).

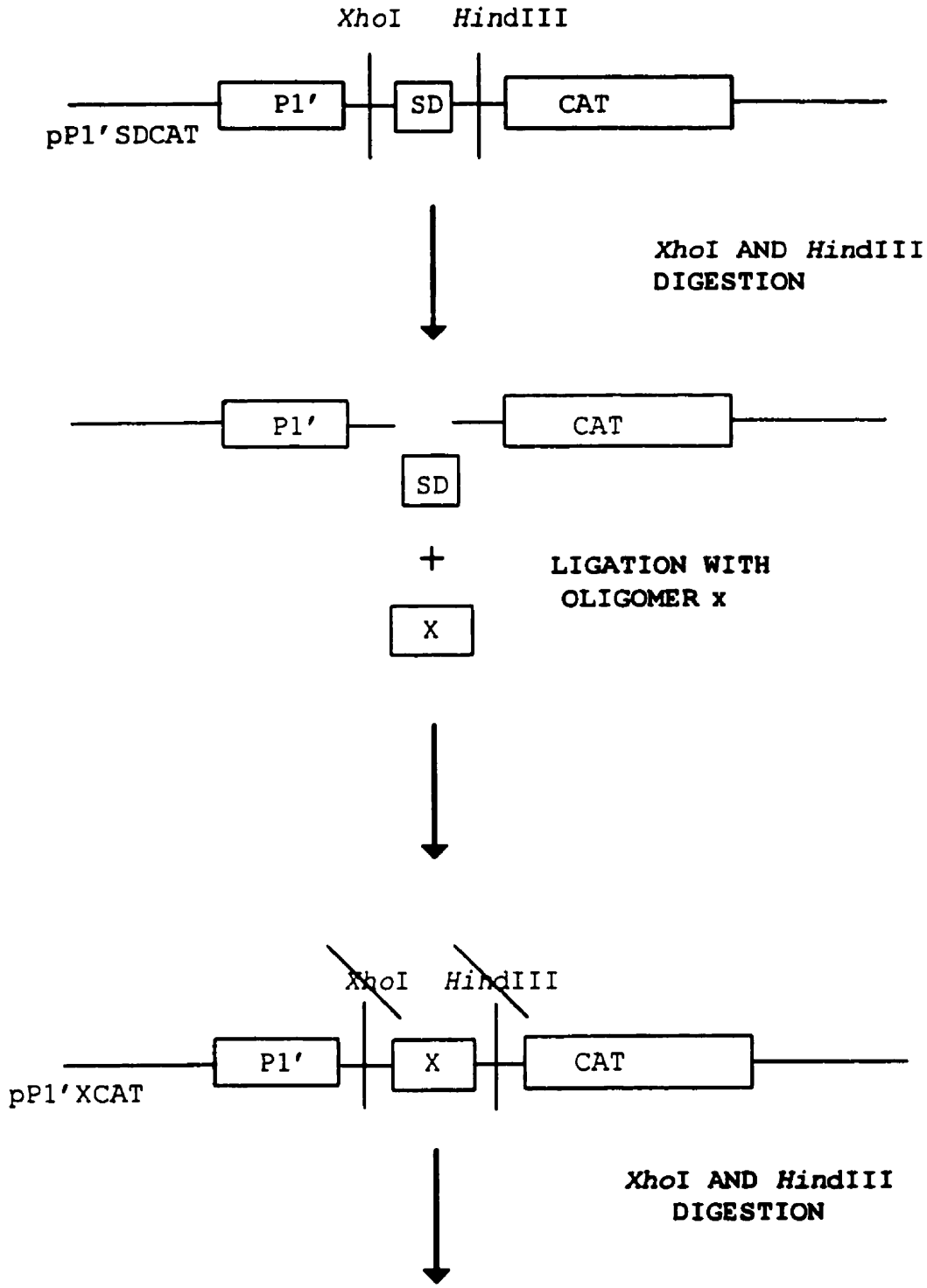

XhoI
HindIII
P1'
SD
CAT
pP1'SDCAT
XhoI AND HindIII
DIGESTION
P1'
CAT
SD
+
LIGATION WITH
OLIGOMER x
X
XhoI
HindIII
P1'
X
CAT
pP1'XCAT
XhoI AND HindIII
DIGESTION
TRANSFORMATION

Figure 2.3: Schematic diagram illustrating the cloning procedure of Ω42, Ω43 and epsilon oligomers. Plasmid pP1'SDCAT was digested with *Xho*I and *Hin*dIII to remove the consensus SD sequence. Hybridized double-stranded oligomers (designated X) were then ligated individually into the digested plasmid. Ligated DNA was digested with *Xho*I and *Hin*dIII prior to transformation.

2.3.1 *Xho*I AND *Hind*III DIGESTION OF pP1'SDCAT

Purified pP1'SDCAT (2 μg) was digested to completion at 37°C for 4 hrs in a 50 μl reaction volume containing 1X NEBuffer 2, 100 μg/ml BSA, 20 units of *Xho*I (NEB) and 20 units of *Hind*III (NEB). Digested plasmid was deproteinized by phenol/chloroform extraction, and precipitated with 0.3 M NaCl and 2.5 volumes of EtOH at -20°C for 2 hrs. DNA was recovered by centrifugation at 16,000 g for 5 min at 4°C. The digested DNA pellet was resuspended in 20 μl TE^{-1} and its concentration was estimated by agarose gel electrophoresis.

To determine the quality of digestion, 100 ng of the digested plasmid was transformed into *E.coli* LE392 cells. The quality of the digestion was deemed acceptable if transformation of 100 ng of digested DNA produced 10 or fewer colonies.

2.3.2 LIGATION REACTION WITH Ω OR ϵ OLIGOMERS AND DEACTIVATION OF LIGASE

Two hundred ng hybridized double stranded oligomer (Ω42, Ω43, ϵI, ϵII or ϵSD; see Appendix for the sequences) were ligated to 1 μg of digested DNA in a 40 μl ligation reaction mixture containing 1X T4 DNA ligase buffer and 20 WEISS units of T4 DNA ligase (NEB). This ligation reaction mixture was incubated at 16°C for 2 hrs and then at 24°C for 1 hr. T4 DNA ligase was inactivated by incubation at 65°C for 20 min.

2.3.3 DIGESTION OF THE PLASMIDS PRIOR TO TRANSFORMATION

The plasmid was digested with *Xho*I and *Hind*III by supplementing the ligation reaction mixture with (50 mM final NaCl), 10 units of *Xho*I and 10 units of *Hind*III. This digestion reaction mixture was incubated at 37°C for 2 hrs, and half of the reaction mixture was transformed into *E.coli* LE392 cells.

2.3.4 SCREENING AND VERIFICATION OF THE CONSTRUCTS

Plasmids carrying the oligonucleotides of interest were initially identified by their resistance to *Xho*I and *Hind*III digestion as before (see 2.2.3). The presence of the oligomers was further verified through dideoxynucleotide sequencing.

2.4 PLACING THE CHLORAMPHENICOL ACETYL TRANSFERASE REPORTER GENE UNDER THE CONTROL OF THE 5' UNTRANSLATED REGIONS OF CYMV AND PMV GENOMES

The pP1'SDCAT plasmid was digested with *Xho*I and *Hind*III thereby removing the consensus SD sequence, which was then released by gel electrophoresis. The plasmid was further ligated with CYMVP or PMVP oligomers followed by transformation of cells (see Figure 2.4).

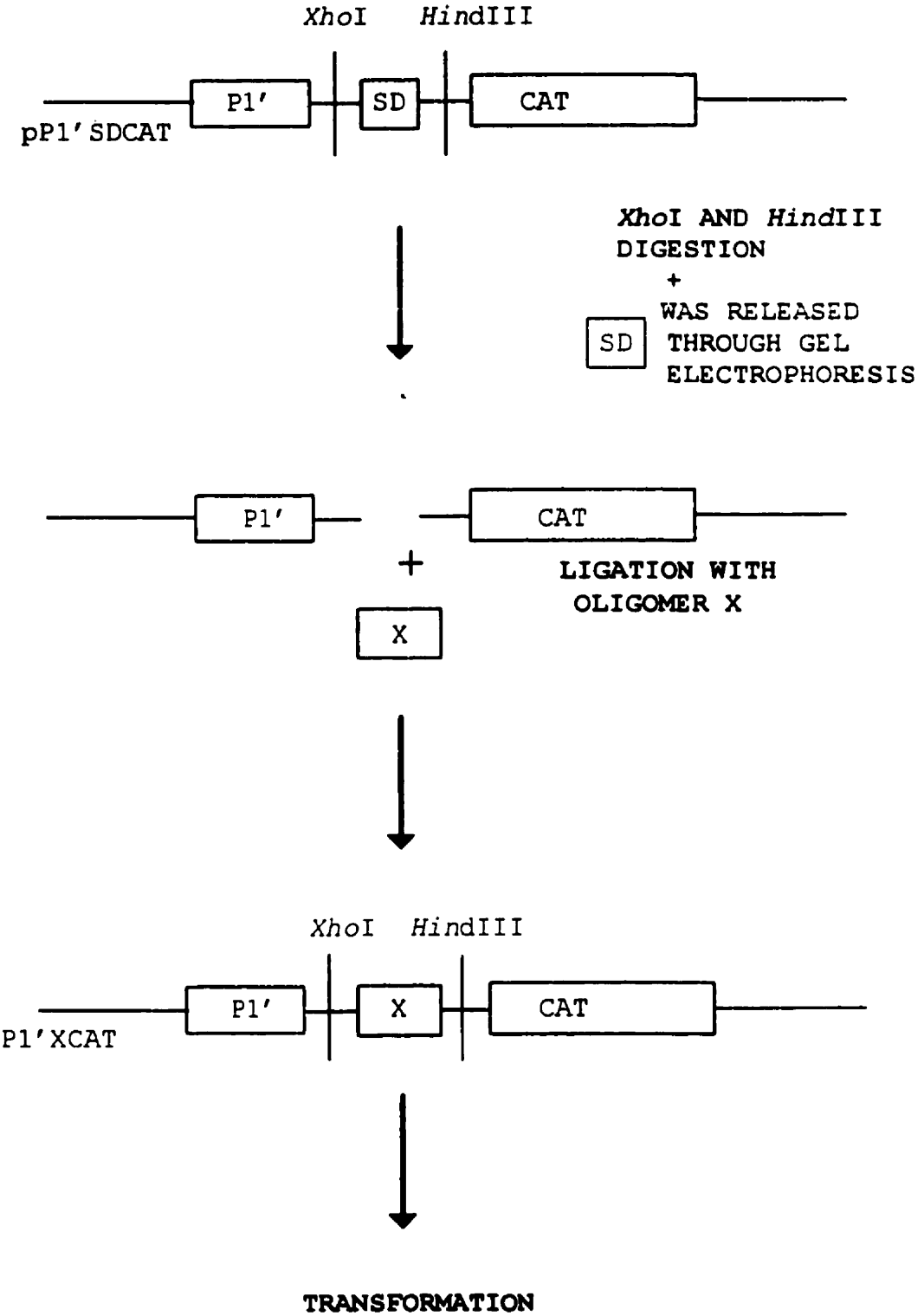

XhoI HindIII
P1'
SD
CAT
pP1'SDCAT
XhoI AND HindIII
DIGESTION
+
SD
WAS RELEASED
THROUGH GEL
ELECTROPHORESIS
P1'
CAT
+
LIGATION WITH
OLIGOMER X
X
XhoI HindIII
P1'
X
CAT
pP1'XCAT
TRANSFORMATION

Figure 2.4: Schematic diagram illustrating the cloning procedure of the CYMVP and PMVP oligomers. Plasmid pP1'SDCAT was digested with *Xho*I and *Hin*dIII. The SD sequence was released through gel electrophoresis and the plasmid was then ligated with hybridized double-stranded oligomers (designated X) individually and transformed.

2.4.1 *Xho*I and *Hin*dIII DIGESTION OF pP1'SDCAT CONSTRUCT

Digestion of 3 μg purified pP1'SDCAT DNA was performed as outlined earlier with *Xho*I and *Hin*dIII (see 2.3.1), followed by phenol/chloroform extraction and precipitation. The concentration of digested DNA was estimated on an agarose gel and its quality examined by transformation.

2.4.2 RELEASE OF SD SEQUENCE FROM pP1'SDCAT THROUGH GEL ELECTROPHORESIS

The digested DNA mixture (see 2.4.1) was run on a 1% agarose gel until the dye migrated about 3 cm from the well. At this stage the fragment containing the SD sequence (16 nucleotides long) was released from the plasmid DNA. Plasmid DNA was then isolated from the gel by trapping or freeze and thaw method followed by phenol/chloroform extraction and precipitation as before.

2.4.3 LIGATION REACTION WITH CYMVP OR PMVP OLIGOMERS AND TRANSFORMATION

Six hundred ng of hybridized double stranded untranslated regions of CYMV or PMV, CYMVP or PMVP oligomers respectively (see Appendix for the sequences), were used in a ligation reaction containing 1 μg of digested plasmid DNA as described earlier (see

3.2.2). Half of the ligation mixture was transformed into *E.coli* LE392 cells.

2.4.4 SCREENING AND VERIFICATION OF THE CONSTRUCTS

The plasmids carrying the DNA fragment of interest were identified through DNA or RNA colony hybridization. Radioactively labelled (see 2.1.8.3) CYMV18 and PMV18 18mers (see Appendix for the sequences), were used as probes to screen for the plasmids carrying CYMVP and PMVP oligomers respectively. The hybridization temperature of 47°C was used for both 18mers. Dideoxynucleotide sequencing was further used to verify the presence of the oligomers.

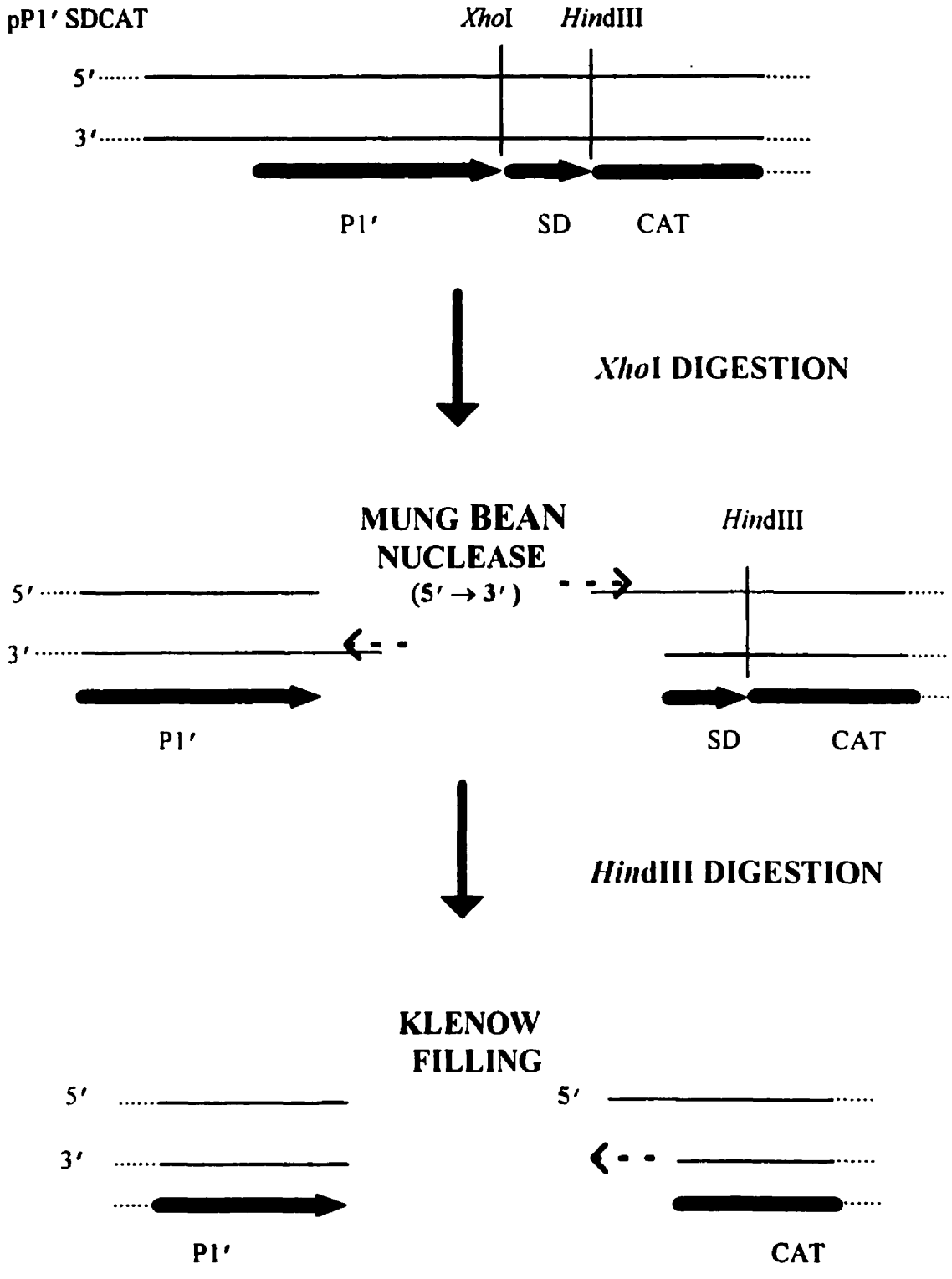
2.5 CONSTRUCTION OF CONTROL CASSETTES

2.5.1 PLACING THE REPORTER GENE UNDER THE CONTROL OF A 5 NUCLEOTIDE-LONG SEQUENCE

A control construct (5NT) carrying a 5 nucleotide sequence (5'-AGCTT) as the mediator of CAT mRNA initiation was made by removing the *Xho*I site and the SD sequence, followed by filling-in the *Hind*III site in pP1'SDCAT plasmid (see Figure 2.5).

2.5.1.1 *Xho*I DIGESTION OF pP1'SDCAT

One μg of purified pP1'SDCAT was digested to completion at 37°C for 4 hrs with *Xho*I as described earlier (see 2.2.2) and then


pP1′ SDCAT
XhoI
HindIII
5′
3′
P1′
SD
CAT
XhoI DIGESTION
MUNG BEAN NUCLEASE
(5′ → 3′)
HindIII
5′
3′
P1′
SD
CAT
HindIII DIGESTION
KLENOW FILLING
5′
3′
P1′
5′
CAT

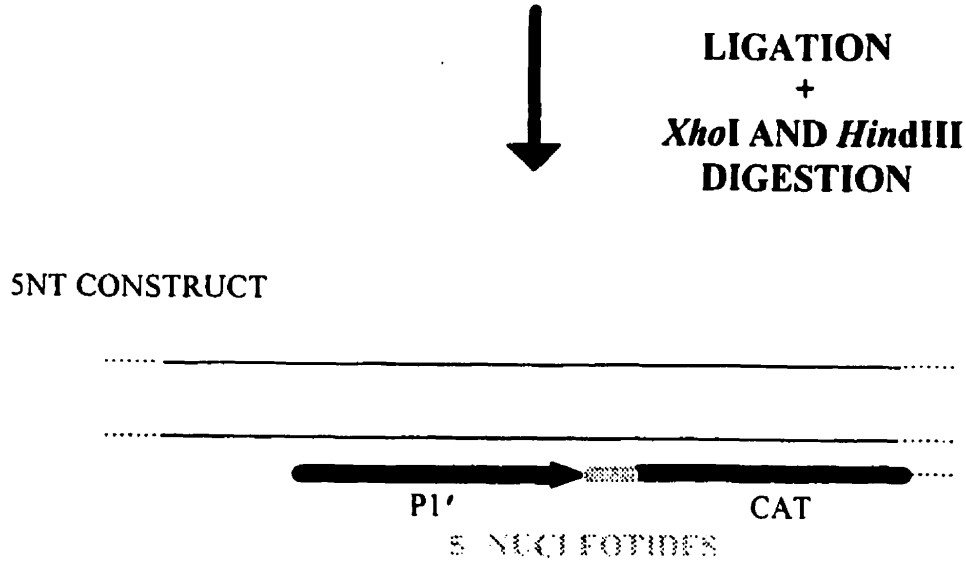


Figure 2.5: Schematic diagram illustrating the construction procedure of 5NT construct. Plasmid pPl'SDCAT was first digested with *Xho*I, producing overhangs which were blunted with mung bean nuclease. The DNA was then digested with *Hind*III and the resulting 3' recessed ends were filled-in with Klenow. The plasmid was blunt-end ligated and then digested with *Xho*I and *Hind*III prior to transformation.

phenol/chloroform extracted and precipitated. The concentration of digested DNA was estimated on an agarose gel and the quality of the digestion was examined by transformation.

2.5.1.2 BLUNT-ENDING OF THE DNA USING MUNG BEAN NUCLEASE

The 5' *Xho*I overhangs were made blunt with mung bean nuclease in a 50 µl reaction volume containing 1 µg of plasmid, 30 mM NaAc pH 4.8, 50 mM NACl, 1 mM zinc chloride ($ZnCl_2$), 5% glycerol (v/v) and 1 unit of mung bean nuclease (diluted to 1 unit/µl with 10^{-3} Tritron X-100 from 169 units/µl; Pharmacia) incubated at 37°C for 20 min. The DNA was phenol/chloroform extracted and precipitated.

2.5.1.3 *Hin*dIII DIGESTION

The blunt DNA was digested to completion with *Hin*dIII as before (see 2.3.1). The reaction was stopped by phenol/chloroform extraction and the DNA was pelleted as before.

2.5.1.4 BLUNT ENDING OF THE DNA USING KLENOW

The 3' recessed *Hin*dIII ends were filled in with Klenow in a 50 µl reaction volume containing 1X Klenow buffer (7 mM Tris-HCl pH 7.5, 50 mM NaCl, 7 mM $MgCl_2$), 4 mM DTT, 250 µM of each deoxynucleotide triphosphate (dNTP) and 5 units of Klenow, and then

incubated at 37°C for 20 min. The DNA was phenol/chloroform extracted and precipitated as before.

2.5.1.5 BLUNT-END LIGATION

The DNA was then blunt-end ligated in a 40 µl reaction volume as outlined earlier (see 2.3.3) with the exception of using 100 WEISS units of T4 DNA ligase (NEB). The ligase was then heat-inactivated.

2.5.1.6 DIGESTION OF THE PLASMID PRIOR TO TRANSFORMATION

The ligated DNA was digested with *Xho*I and *Hind*III prior to transformation as before (see 2.3.3). Half of the reaction mixture was transformed into *E.coli* LE392 cells.

2.5.1.7 SCREENING AND VERIFICATION OF THE CONSTRUCT

The plasmids were initially screened through antibiotic resistance phenotyping as previously explained. The plasmids which were resistant to low concentrations (50 µg/ml) of Cm, but not to higher concentrations (100-400 µg/ml), were further characterized by their resistance to *Xho*I and *Hind*III digestion. The sequence modifications of the plasmid were finally verified through dideoxynucleotide sequencing.

2.5.2 PLACING THE REPORTER GENE UNDER THE CONTROL OF A 36 NUCLEOTIDE LONG SEQUENCE

A control clone (L36), which carries a 36 nucleotide long oligomer, (see Appendix for the sequence) as the mediator of initiation for CAT mRNA, was constructed in the same manner as outlined earlier for CYMVP and PMVP oligomers (see 2.4).
The constructs were screened by RNA colony hybridization at 58°C using ^{32}P end-labelled L36B (see Appendix for the sequence).

2.6 QUANTIFICATION OF CHLORAMPHENICOL ACETYL TRANSFERASE ACTIVITY

Chloramphenicol acetyl transferase activity was quantified for the constructs using a spectrophotometric assay described by Shaw (1975), with some modifications. Transformed colonies of bacteria were inoculated individually into 4 ml of LB broth supplemented with 100 μg/ml chloramphenicol (Cm) and incubated at 37°C overnight with shaking at 300 RPM. The next day, 4 ml of LB broth supplemented with 50 μg/ml Ap were inoculated with 80 μl of this overnight culture and incubated at 37°C with shaking until an optical density of 0.8-1.2 units at 590 nm was obtained (about 4 hrs). Two optical density units of the cells were then removed, cooled on ice for 10 min and pelleted at 1,500 g for 6 min. The pelleted cells were resuspended in 1 ml of ice-cold ST buffer (50 mM Tris-HCl pH 7.8 and 150 mM NaCl) and centrifuged as before. One ml of ice-cold TM buffer (50 mM Tris-HCl pH 7.8 and 50 μM β-mercaptoethanol) was used

to resuspend the pellet. The mixture was transferred into a glass tube and sonicated for 2 min on ice using a power setting of 30 Watts. This sonicated suspension was then transferred into a microfuge tube and centrifuged at 16,000 g for 10 min. The supernatant was removed and kept on ice until assayed for the CAT activity or used to determine the total protein concentration.

The reaction mixture was freshly prepared from the individual reagents. The final concentration of each component was as follows: 100 mM Tris-HCl pH 7.8, 0.1 mM acetyl coenzyme A (CoA) and 0.4 mg/ml 5,5'-Dithiobis-2-nitrobenzoic acid (DTNB). After the cuvette (1 cm light path) containing 10 μl of sonicated cell suspension and 1 ml of reaction mixture was equilibrated in a 37°C water bath, the absorption at 412 nm was set to zero. The reaction was started by addition of Cm to a final concentration of 0.1 mM. The optical density was recorded every minute for five minutes. The final optical density of 0.3-1.0 units was used for further calculations; in case of deviation from these values, the assay was repeated with a proper concentration of cell lysis suspension. The change in optical density per minute was divided by 13.6 (extinction coefficient) to give the rate of Cm acetylation in micromoles per minute. Since 1 unit of CAT is defined as 1 micromole of Cm acetylated per minute at 37°C, the calculation also yields the number of units of enzyme in the cuvette which was subsequently divided by the total protein (mg) (see below) to give the units of enzyme per mg total protein for a CAT assay.

2.7 DETERMINATION OF TOTAL PROTEIN CONCENTRATION IN CAT ASSAY

A modified procedure of Bradford (1976) was employed to determine the total protein concentration for the constructs. A standard curve of known protein concentration was obtained using bovine serum albumin (BSA). The Standard curve samples as well as the samples of unknown protein concentrations (prepared as in CAT assay) were brought to a final volume of 100 μl with dH_2O. One ml coomassie Brilliant Blue solution was added to each sample and vortexed. After inoculation for 2 min at 24°C the optical density of each tube at 592 nm was obtained.

3. RESULTS

3.1 CONSTRUCTION OF EXPRESSION PLASMIDS

3.1.1 PLASMID pPBRP1'Tc

The expression plasmids used in this study carried a modified version of the strong synthetic constitutive promoter P1 (variant of the T5 phage early promoter; Rommens *et al.*, 1983). In the starting plasmid pBRP1Tc (kindly provided by Dr. Ivan Ivanov), a 14 nucleotide spacer region separated the P1 promoter from a synthetic SD consensus sequence which mediates initiation of translation for the tetracycline (Tc) resistance gene. This plasmid was modified for the purposes of this study by substituting a synthetic double stranded oligonucleotide (PP1') for the original P1 promoter, the 14 nucleotide untranslated region and the SD sequence (see below). The PP1' fragment carried the core part of the P1 promoter and the same SD consensus sequence separated by a *Xho*I restriction site. For convenient cloning the PP1' was flanked by *Eco*RI (at the 5'-end) and *Hin*dIII (at the 3'-end) overhangs. As seen in Figure 3.1 the fifth nucleotide of the *Eco*RI site was modified (G instead of C). This led to loss of the *Eco*RI site after ligation in pPBRP1TC and facilitated the screening for positive clones with the PP1' insert.

Modified plasmid (designated pBRP1'Tc) was selected first on the basis of resistance to *Eco*RI. Figure 3.2 shows that while the plasmids in lanes 4 and 5 were digestible, the plasmids in lanes 1,2,3, and 6 remained resistant to *Eco*RI and were used for further verification of their primary structure.

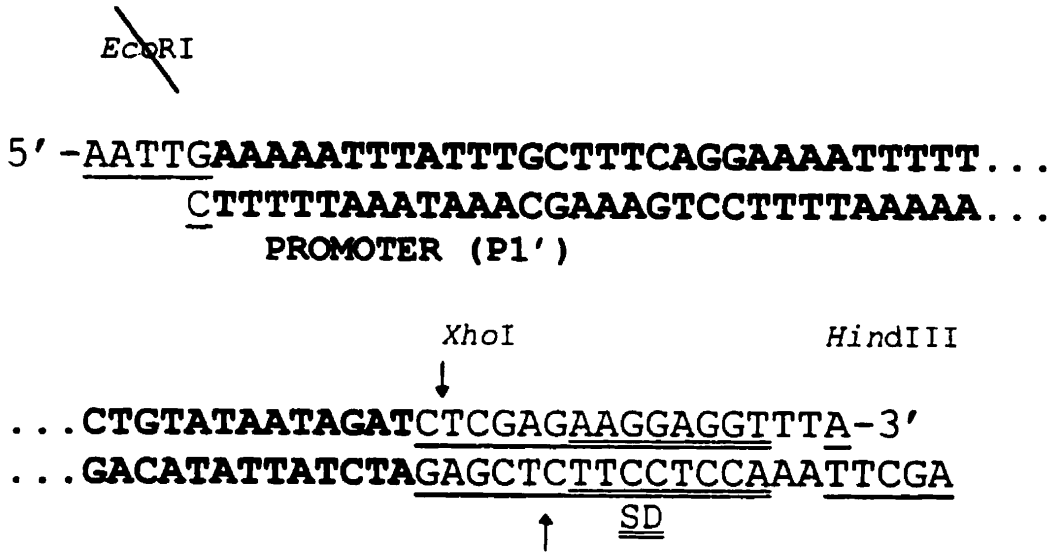


Figure 3.1: Nucleotide sequence of PP1′ oligomer. The promoter region is in bold letters. Restriction sites are underlined and the SD sequence is doubly underlined. The *Eco*RI site in pBRP1Tc is destroyed after ligation.

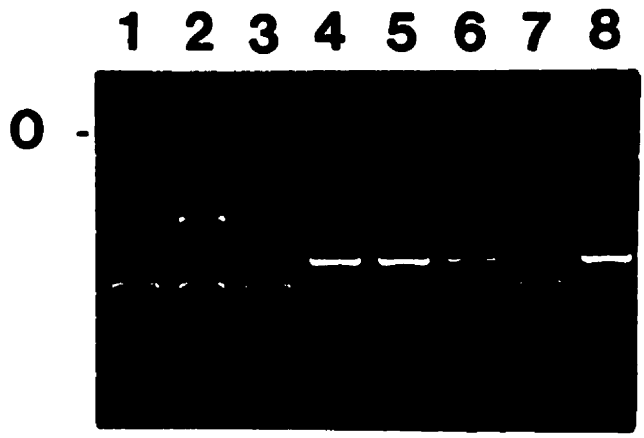


Figure 3.2: Screening of pBRP1'Tc constructs by *Eco*RI. Plasmids in lanes 1, 2, 3 and 6 were not cut by *Eco*RI enzyme and thus they may carry the P1' promoter. Plasmids in lanes 4 and 5 were digested by *Eco*RI, suggesting that they are of the parental (pBRP1Tc) type. Lanes 7 and 8 (controls) contain undigested and digested pBRP1Tc plasmids, respectively. Plasmid DNAs were run on a 1% agarose gel. O=origin.

The deletion of the 14 nucleotide long untranslated region from the original plasmid was verified by dideoxynucleotide sequencing (Figure 3.3).

3.1.2 pP1'SDCAT CONSTRUCT

The P1' promoter was cloned into pP1SDCAT by swapping the *PstI/XhoI* fragment between the pBRP1'Tc and pP1SDCAT. The latter was constructed on the basis of pBR322 (kindly provided by Dr. Ivan Ivanov) and carried a chloramphenicol acetyl transferase (CAT) reporter gene. The newly made expression plasmid was designated pP1'SDCAT.

A three-step screening procedure was employed for the selection of positive clones. Clones were initially screened for their antibiotic resistance. Only clones which survived in media containing 50 μg/ml Ap and 100 μg/ml Cm but were sensitive to Tc (20 μg/ml) were selected for further analysis. Plasmids isolated from the positive clones were then analyzed by *EcoRI* digestion (Figure 3.4). The *EcoRI* site in the parental plasmid is abolished (see above) so the *EcoRI* site generated by the CAT gene is unique for the new plasmid pP1'SDCAT. Thus the linearization of the plasmid (single cut) by *EcoRI* is an indication for a correct construct. Plasmid DNAs in lanes 2,3,4, and 6 each appeared as a single band and therefore carried only one *EcoRI* restriction site. Plasmid DNA in lane 5 released a fragment (about 850 base pairs) indicating that it belongs to the parental pP1SDCAT plasmid type.

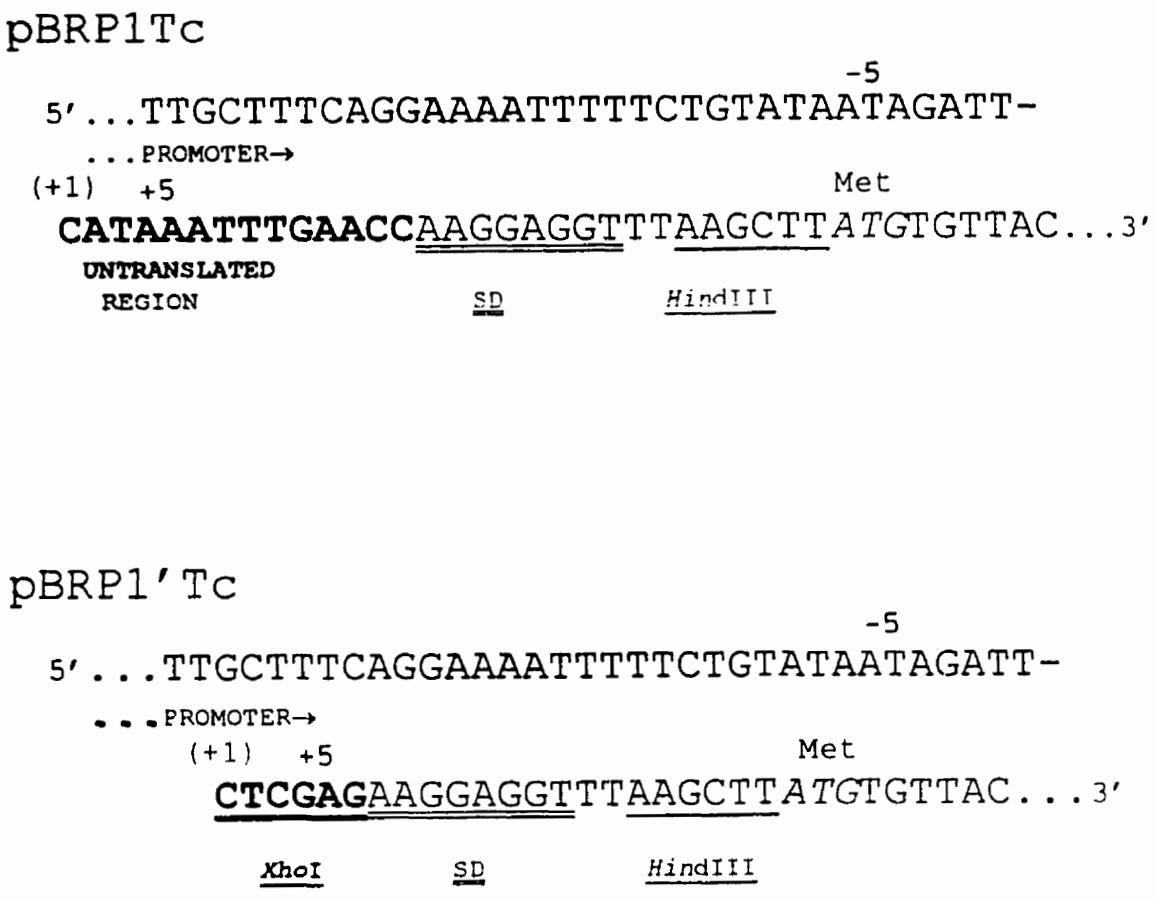


Figure 3.3: Nucleotide sequence of the promoter and the SD regions of pBRP1Tc and pBRP1'Tc. The spacer between the promoter and the SD sequence is in bold letters. The restriction sites are underlined and the SD sequence is doubly underlined. (+1) indicates the start of transcription.

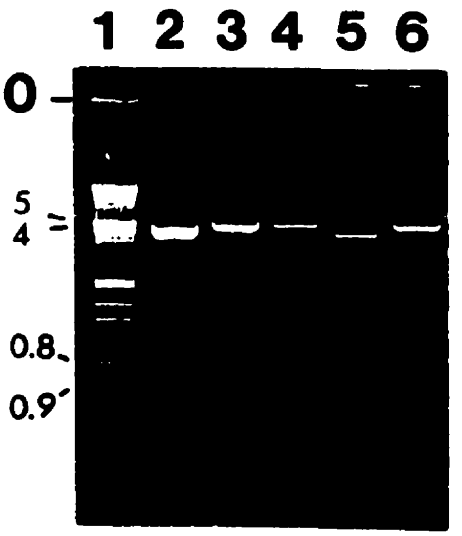


Figure 3.4: Screening of pP1'SDCAT DNA using *Eco*RI. Digestion of plasmid DNAs in lanes 2,3,4 and 6 produced a single DNA fragment of about 4800 base pairs in size, suggesting that the plasmids may carry the CAT gene and hence they may be the expression cassette of interest (pP1'SDCAT). Plasmid DNA in lane 5 produced 2 fragments (about 4000 and 850 base pairs long) suggesting that it Is probably of the background (pP1SDCAT) type. Lane 1 contains λ DNA digested with *Hin*dIII, used as a size marker. Plasmid DNAs were run on a 1% agarose gel. Fragment sizes are in kilo base pairs. 0=origin.

Finally, the correct structure of the newly constructed gene expression cassette was verified by dideoxynucleotide sequencing analysis (Figure 3.5).

3.2 ARE LOOPS # 42 AND # 43 OF THE 16S rRNA INVOLVED IN THE ORGANIZATION OF ALTERNATIVE mRNA BINDING SITES?

To answer this question synthetic oligonucleotides carrying nucleotide sequences complementary to the loop regions of either domain # 43 or # 42 (see Figure 3.6) in 16S rRNA are designed to be cloned in place of the SD sequence from the expression plasmid pP1'SDCAT. The two synthetic oligonucleotides are designated Ω43 and Ω42, respectively. As shown in Figures 3.7 and 3.8, the complementary region for oligomer Ω43 and loop 43 in 16S rRNA (nucleotides 1358-1366), is 9 base pairs, whereas that for Ω42 and loop 42 of 16S rRNA (nucleotides 1316-1326) is 11 base pairs. The distance between the complementary region and the AUG start codon of the CAT gene was set to be 21 for both oligomers [the same as in the original construct (Ivanov *et al.*, 1992; 1995) carrying the TMV Ω sequence as translational initiator].

Both oligomers had two sticky ends; one compatible with the *Xho*I and the other with the *Hind*III sticky ends. Since the last nucleotides adjacent to the double stranded regions of the oligomers are changed, both *Xho*I and *Hind*III sites were abolished after ligation to the vector pP1'SDCAT.

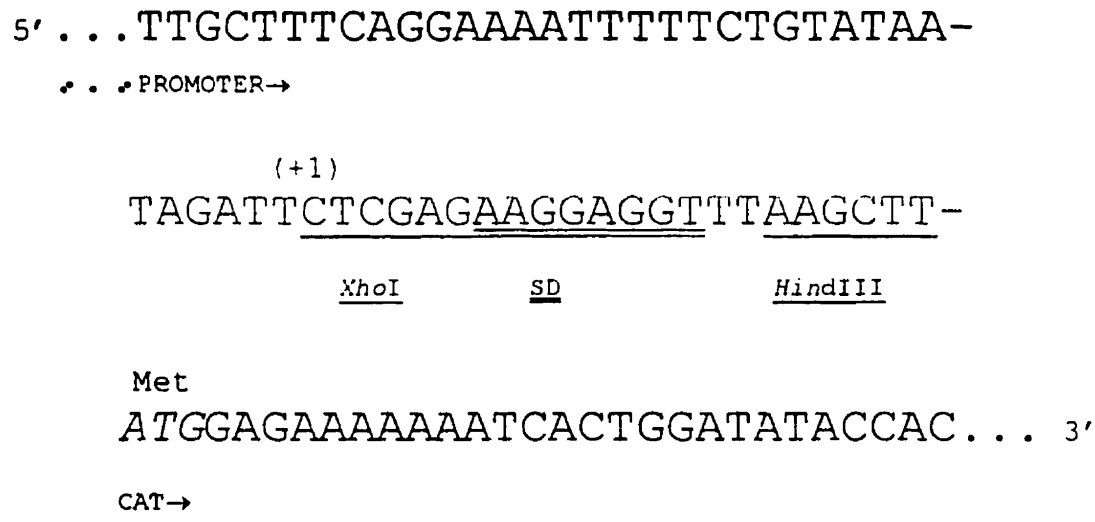
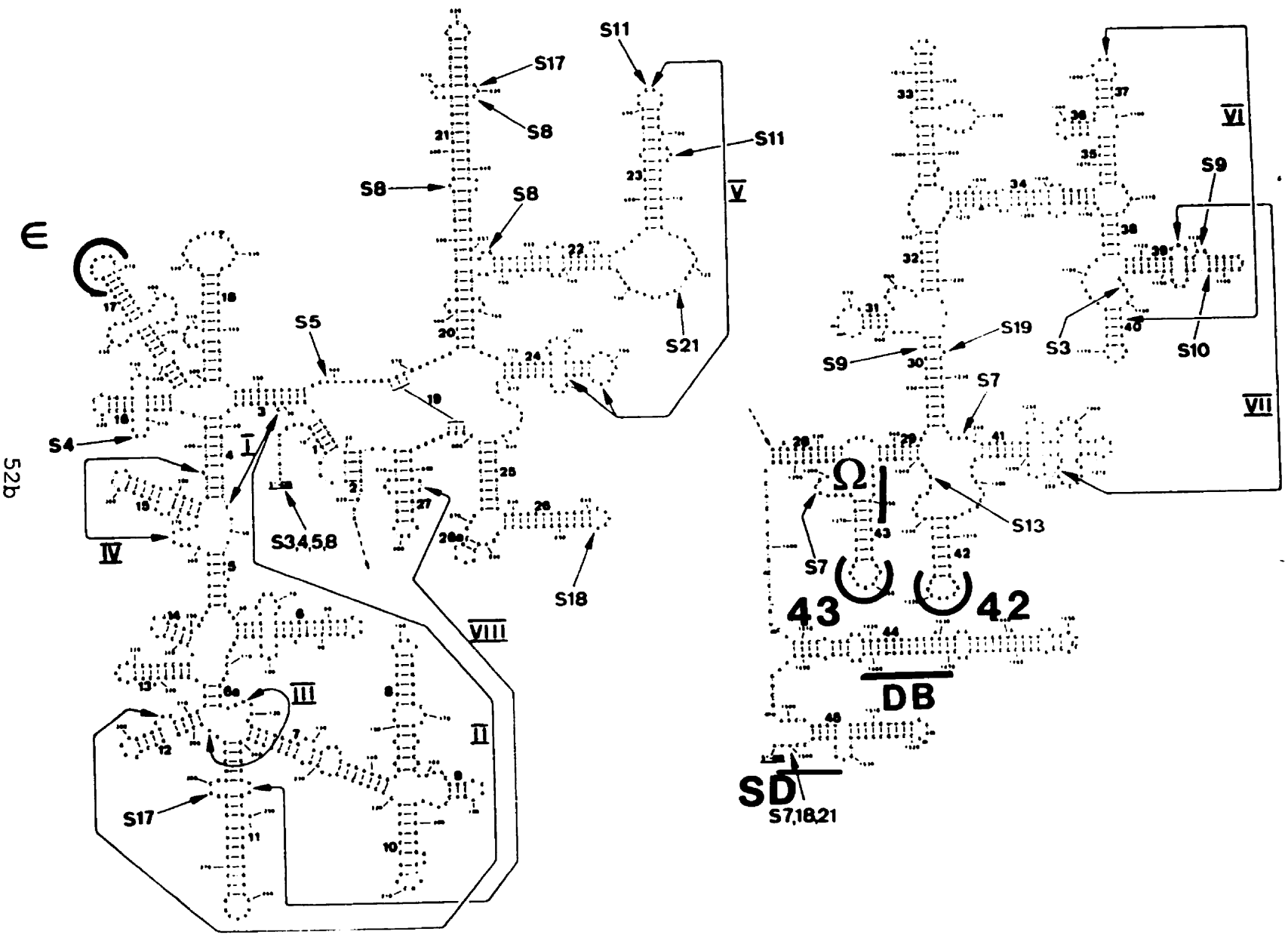


Figure 3.5: Nucleotide sequence of 5′ non-coding region of CAT gene in pP1′SDCAT. Restriction sites are underlined and the SD sequence is doubly underlined. (+1) indicates the start of transcription.


E
S17
S11
S8
S11
S8
S8
V
S5
S21
S4
I
S3,4,5,8
IV
S18
VIII
III
II
S17
VI
S9
S3
S10
S19
S9
S7
VII
Ω
S13
S7
43
42
DB
SD
S7,18,21

52b

Figure 3.6: Schematic diagram illustrating the proposed secondary structure of *E.coli* 16S rRNA. The locations of the intra-RNA cross-links (Roman numerals) and RNA-protein cross-link sites which are used to build the three dimensional model of 30S ribosomal subunit (see Figure 1.2) are indicated by arrows. Sequences complementary to different regions of rRNA are bolded (SD is Shine-Dalgarno sequence, DB is downstream box, Ω is 5' untranslated region of TMV RNA, ϵ is a sequence derived from the RBS of bacteriophage T7 gene *10*, 42 and 43 are the oligomers complementary to loops # 42 and 43 respectively). (Adapted from Brimacombe, 1992).

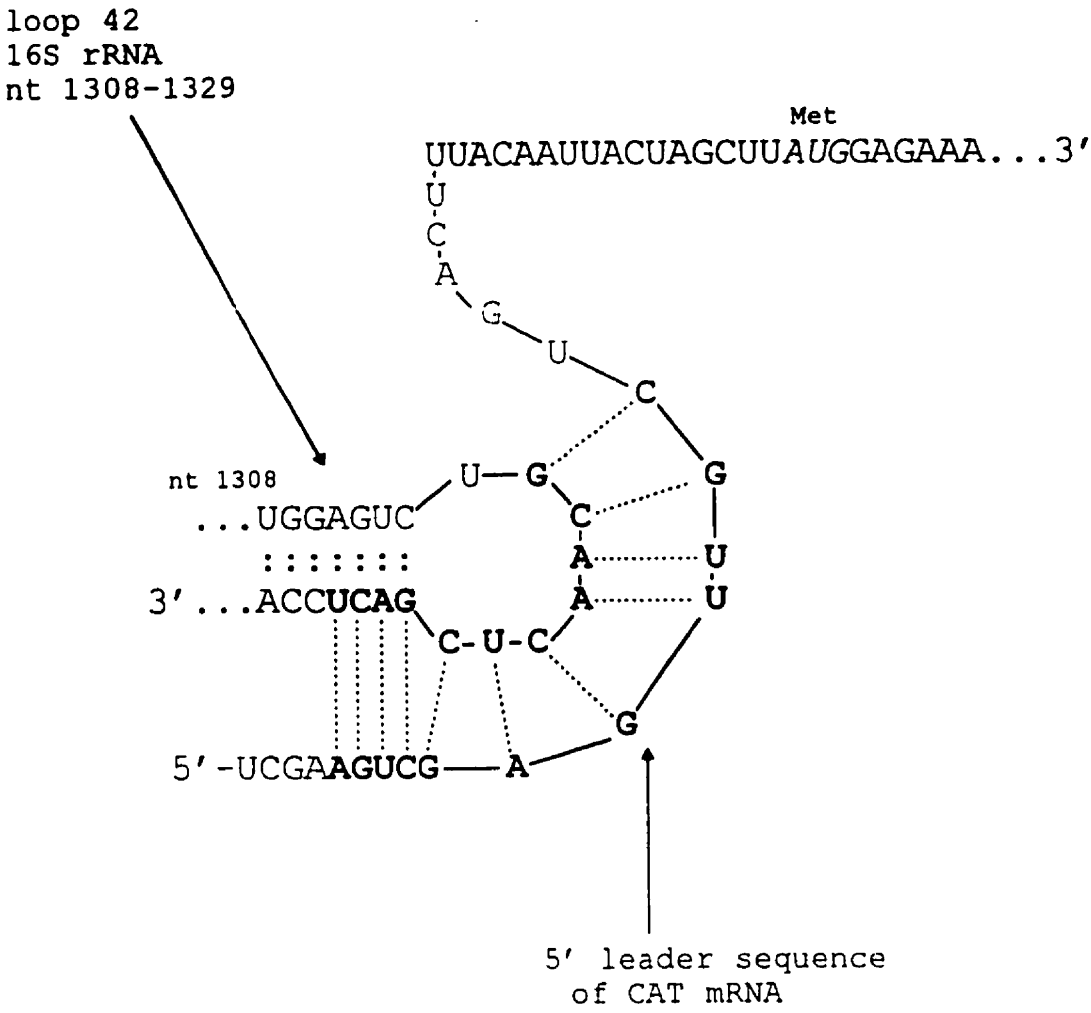


Figure 3.7: Potential base pairing of Ω42 and loop # 42 of 16S rRNA. The complementary base pairing between the oligomer and 16S rRNA is indicated by dashed lines and bold letters.

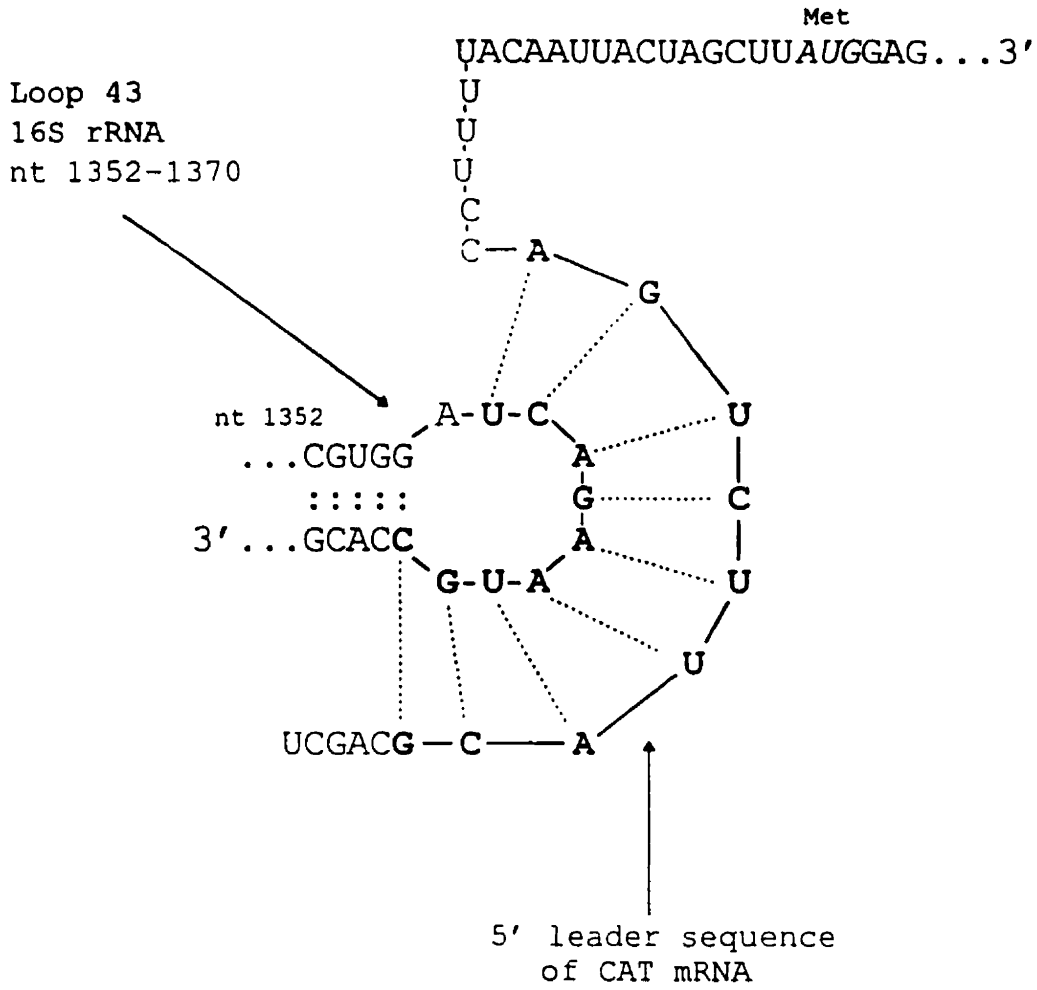


Figure 3.8: Potential base pairing of Ω43 and loop # 43 of 16S rRNA. The complementary base pairing between the oligomer and 16S rRNA is indicated by dashed lines and bold letters.

Constructs carrying the two oligomers Ω42 or Ω43 were identified firstly by their resistance to digestion with *Xho*I and *Hind*III, and secondly by dideoxynucleotide sequencing (Figure 3.9). Unexpectedly, an abnormal construct was also discovered carrying an altered Ω42 oligomer (Ω42b) in which the distance between the complementary region to the 16S rRNA and the start codon of the CAT gene was extended from 21 to 36 nucleotides (see Figure 3.9).

3.2.1 EFFICIENCY OF TRANSLATION INITIATION BY Ω42 and Ω43

Table 3.1 shows the relative efficiency of some nucleotide sequences to initiate translation of CAT mRNA in *E.coli* LE392 cells. The values of CAT activity for each construct is normalized to that of the consensus SD sequence. This SD sequence produced about 1500 units of CAT protein/mg total *E.coli* protein (this value is the basis for calculating the expression of CAT produced by other constructs).

The relative activities of constructs carrying Ω42, Ω43 and SD_n sequence (native to the CAT) as translational initiators were 6%, 22% and 23% respectively. When the distance between the complementary region of Ω42 oligomer to 16S rRNA was increased from 21 to 36 nucleotides (as in construct Ω42b) the efficiency of translation initiation was increased to 72% (Table 3.1).

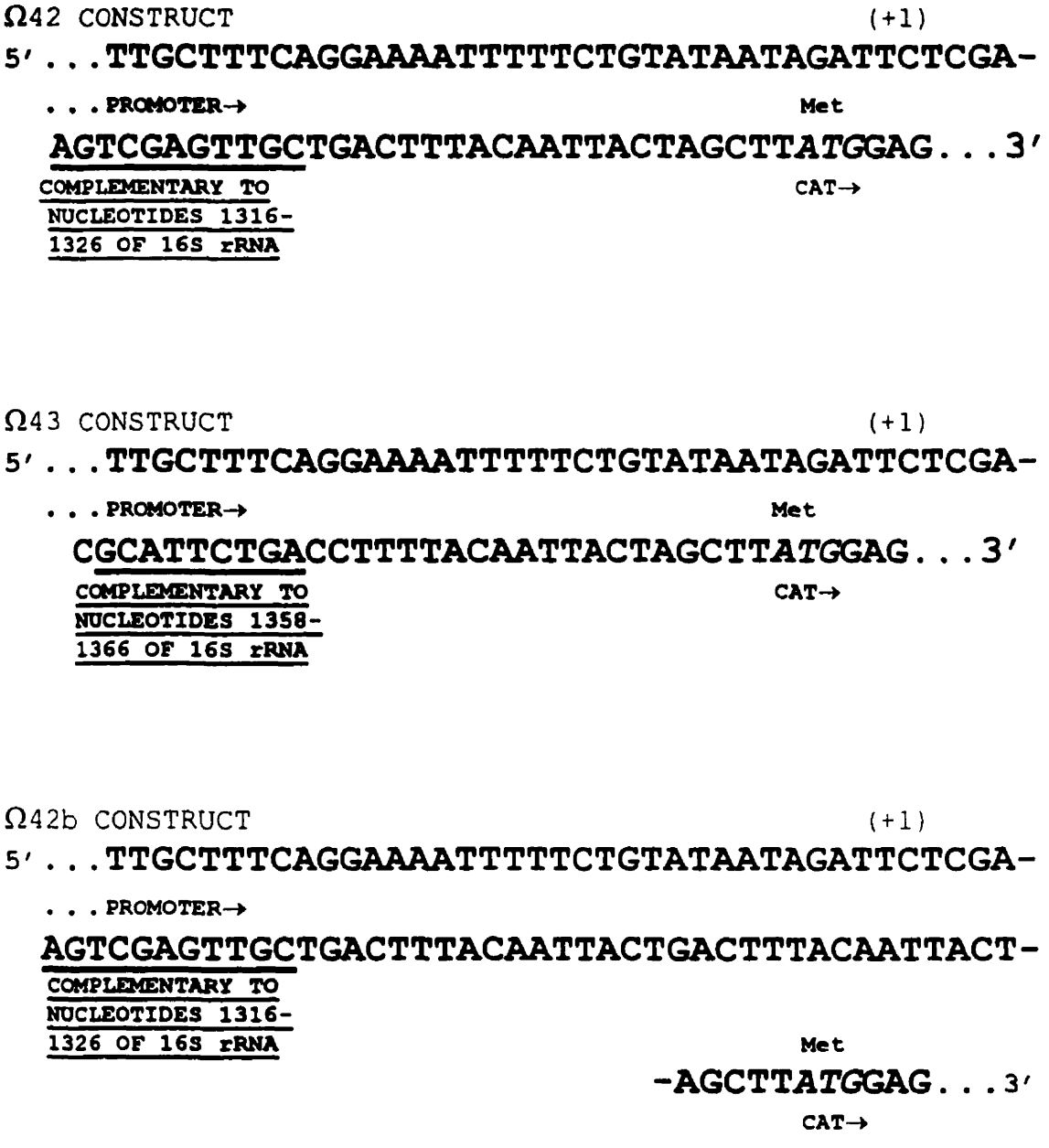


Figure 3.9: Nucleotide sequence of 5′ non-coding region of CAT gene for the constructs carrying Ω42 and Ω43 oligomers. The spacer between the complementary region of the oligomer to 16S rRNA (underlined) and the CAT start codon is 21 nucleotides for Ω42 and Ω43, and 36 nucleotides for Ω42b. (+1) indicates the start of transcription.

CONSTRUCT	RELATIVE ACTIVITY (CONSTRUCT/SD)
SD	$1.00 \pm 0.1 \times 10^{-1}$
SD_n	$2.3 \times 10^{-1} \pm 0.1 \times 10^{-1}$
Ω42	$0.6 \times 10^{-1} \pm 0.1 \times 10^{-1}$
Ω42b	$7.2 \times 10^{-1} \pm 0.1 \times 10^{-1}$
Ω43	$2.2 \times 10^{-1} \pm 0.1 \times 10^{-1}$

Table 3.1: Relative efficiency of oligomers complementary to different loops on 16S rRNA in mediating the initiation of translation of CAT mRNA in *E.coli*. Values represent the average ratio (three independent experiments) of translational activity for each construct over that of the consensus SD sequence. SD is the consensus SD sequence and SD_n is the SD sequence native to the CAT gene. ± represents the standard deviation.

3.2.2 SECONDARY STRUCTURES OF mRNAs CARRYING THE NUCLEOTIDE SEQUENCES Ω42 and Ω42b

Computer predicted secondary structure of the 5' end of the CAT mRNA for Ω42 (the first 150 nucleotides) and Ω42b (the first 165 nucleotides) constructs, are shown in Figures 3.10 and 3.11, respectively. The complementary regions to 16S rRNA are not involved in secondary structures which could have an inhibitory effect on the initiation of translation.

3.3 EFFECT OF EPSILON (ϵ) SEQUENCE ON THE INITIATION OF TRANSLATION OF CAT mRNA IN *E.COLI*

Three oligomers, ϵSD, ϵI, and ϵII were designed to be cloned in place of the SD sequence in the expression plasmid pP1'SDCAT. The oligomer ϵI was designed to be complementary to nucleotides 458-467 in 16S rRNA (Figure 3.12). The ϵSD containing the same complementary sequence to 16S rRNA also contained a consensus SD sequence (3.13). The difference between the ϵII oligomer and ϵI was that the complementary region to 16S rRNA was increased from 10 nucleotides to 14 nucleotides (within nucleotides 451-467) (Figure 3.14). The distance between the complementary regions to 16S rRNA and the AUG start codon of the reporter gene was set to be 9 nucleotides for all the oligomers.

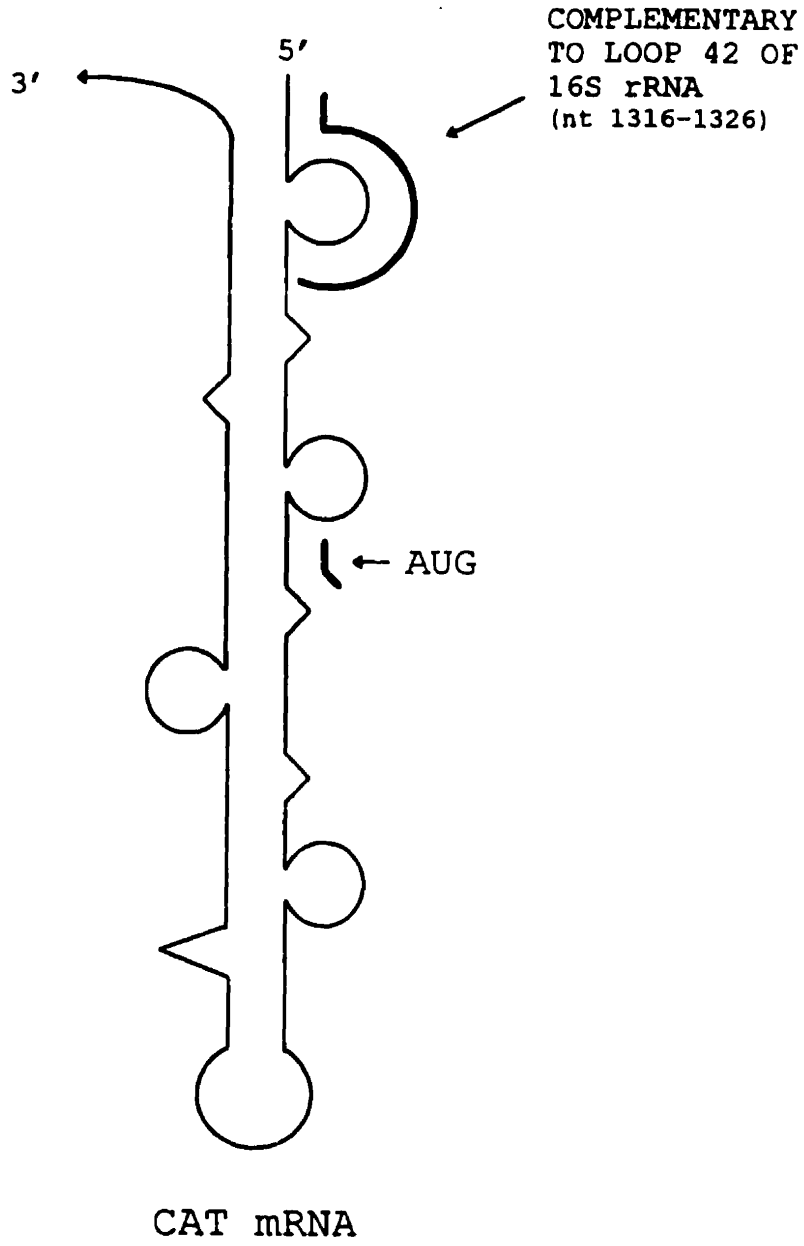


Figure 3.10: Schematic diagram illustrating the predicted secondary structure of the CAT mRNA for Ω42 construct. The programme PCFOLD was used to predict the secondary structure of the first 150 nucleotides of the mRNA.

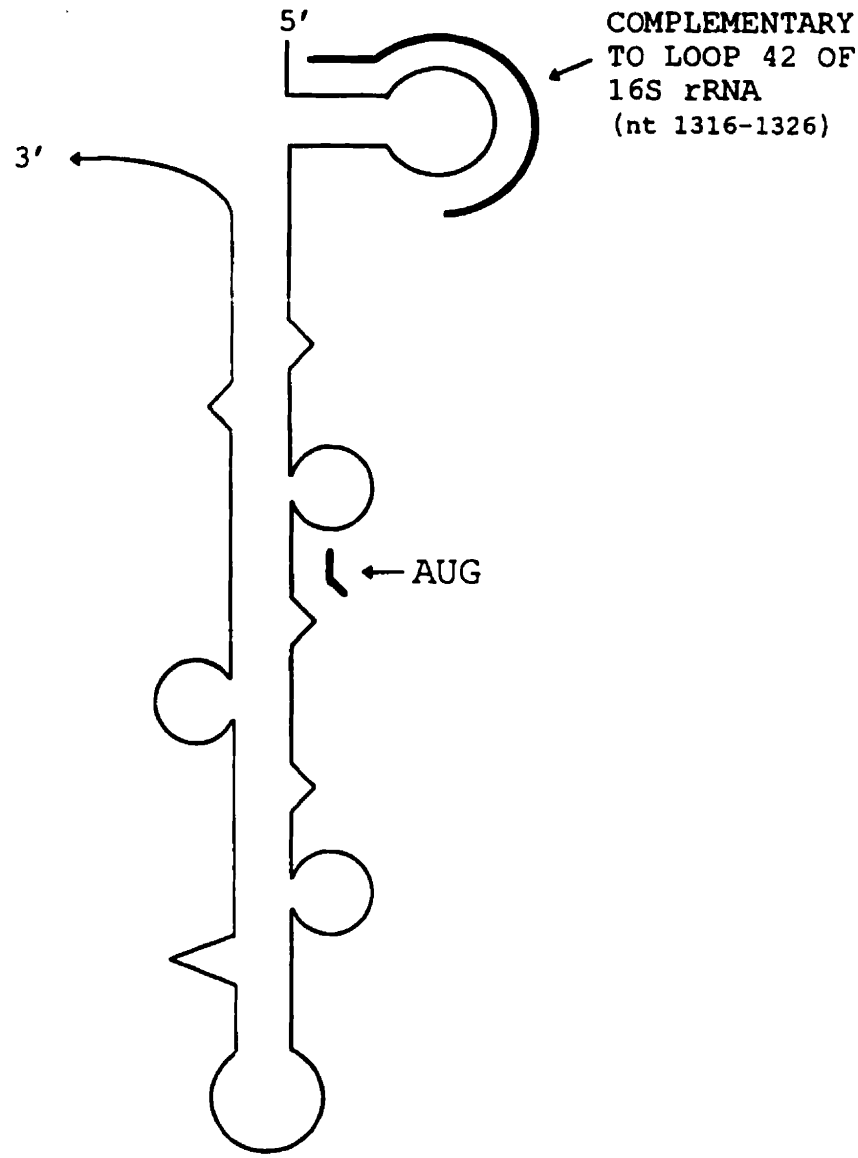


Figure 3.11: Schematic diagram illustrating the predicted secondary structure of the CAT mRNA for Ω42b construct. The programme PCFOLD was used to predict the secondary structure of the first 165 nucleotides of the mRNA.

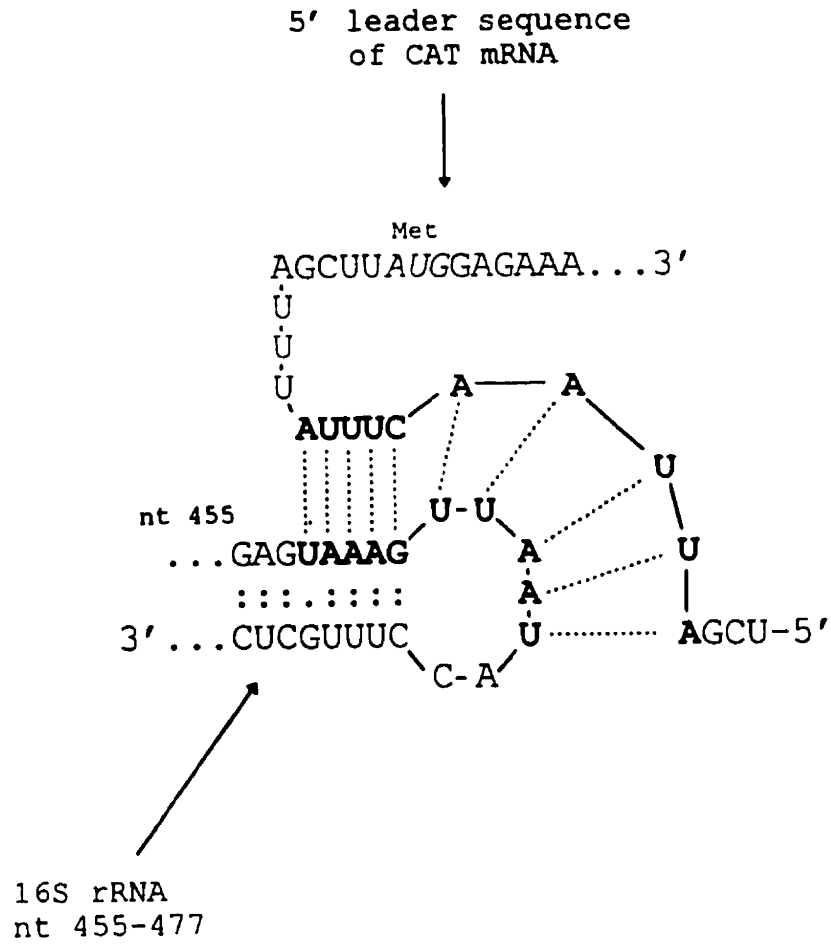


Figure 3.12: Potential base pairing of ∈I and nucleotides 458-467 of 16S rRNA. The complementary base pairing between the oligomer and 16S rRNA is indicated by dashed lines and bold letters.

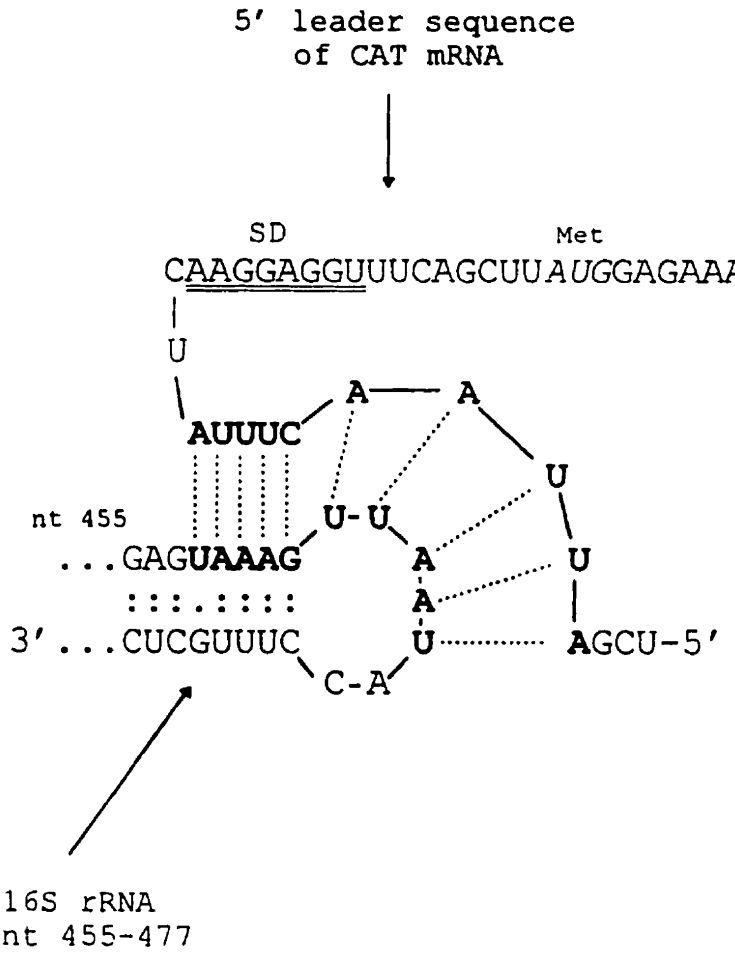


Figure 3.13: Potential base pairing of ϵSD and nucleotides 458-467 of 16S rRNA. The complementary base pairing between the oligomer and 16S rRNA is indicated by dashed lines and bold letters. The SD sequence is doubly underlined.

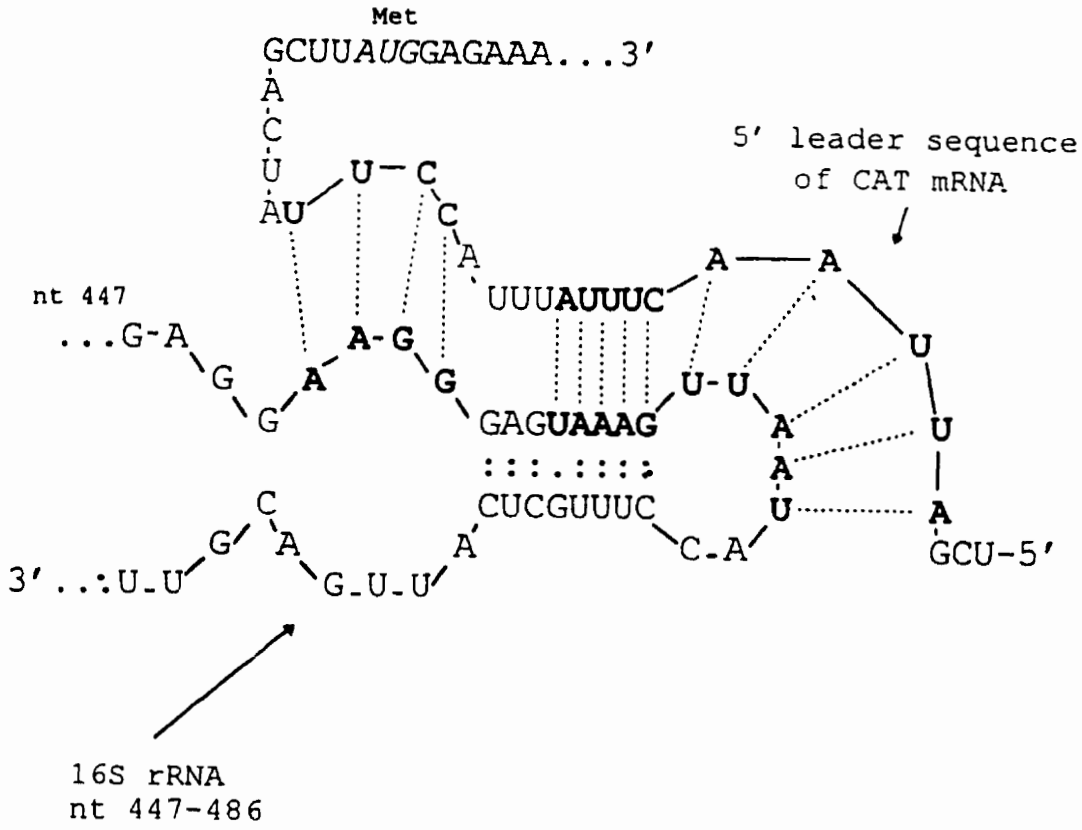


Figure 3.14: Potential base pairing of εII and nucleotides 451-467 of 16S rRNA. The complementary base pairing between the oligomer and 16S rRNA is indicated by dashed lines and bold letters.

The ϵ oligomers contained two "sticky ends"; one compatible with the *Xho*I and the other with the *Hind*III overhangs. Both sites were designed to be destroyed following ligation of the oligomer to their respective compatible ends.

The primary structures of all new constructs are shown in Figure 3.15.

3.3.1 EFFICIENCY OF TRANSLATION INITIATION BY THE EPSILON SEQUENCE

The efficiency of translation initiation was measured by the yield of CAT protein in *E.coli* LE392 cells and the results are presented in Table 3.2. The values show the ratio of translational activity of each construct over that of the consensus SD sequence. The constructs carrying ϵSD oligomer, ϵI oligomer, and ϵII sequences as translational initiators showed relative activities of 250%, 1%, and 85% respectively. The relative activity of the construct carrying the SD_n sequence (native to the CAT gene) was 23%.

3.4 EFFECT OF THE 5′ UNTRANSLATED REGIONS OF CYMV AND PMV RNAs ON THE INITIATION OF TRANSLATION OF CAT mRNA IN *E.COLI* CELLS

Taking into consideration that the 5′ untranslated terminal regions of clover yellow mosaic virus (CYMV) and papaya mosaic virus (PMV) RNA genomes are strong translational enhancers in eukaryotes (as is the TMV Ω sequence) their capability to initiate translation in *E.coli* cells was tested. Two oligomers CYMVP and PMVP were

Figure 3.15: Nucleotide sequence of 5′ non-coding region of the CAT gene for the constructs carrying epsilon oligomers. The spacer between the complementary region of the oligomers to 16S rRNA (underlined) and the CAT start codon is 8 nucleotides. The SD sequence is doubly underlined. The complementary region of the oligomer to the epsilon binding site of 16S rRNA is 10 nucleotides for ∈I and ∈SD, and 14 nucleotides for ∈II. (+1) indicates the start of transcription.

CONSTRUCT	RELATIVE ACTIVITY (CONSTRUCT/SD)
SD	$1.00 \pm 0.1 \times 10^{-1}$
SD_n	$2.3 \times 10^{-1} \pm 0.1 \times 10^{-1}$
ϵI	$0.1 \times 10^{-1} \pm 0.1 \times 10^{-1}$
ϵSD	$2.50 \pm 0.1 \times 10^{-1}$
ϵII	$8.5 \times 10^{-1} \pm 0.1 \times 10^{-1}$

Table 3.2: Relative efficiency of epsilon oligomers in mediating initiation of CAT mRNA translation in *E.coli*. Values represent the average ratio (n=3) of translational activity for each construct over that of the consensus SD sequence. SD is the consensus SD sequence and SD_n is the SD sequence native to the CAT gene. ± represents the standard deviation.

designed to be cloned into the expression vector pP1'SDCAT in place of the SD sequence. Both oligomers contain a 5' *Xho*I and a 3' *Hind*III overhangs and their length was 97 and 102 nucleotides for the CYMVP and PMVP, respectively (Figure 3.16).

Since the two restriction sites (*Xho*I and *Hind*III) were not designed to be destroyed after cloning, an alternative approach was applied for screening and selection of positive clones.

The constructs were screened by colony hybridization using a ^{32}P end-labelled synthetic 18mer DNA fragment, complementary to either the CYMVP or PMVP oligomers. The results from a DNA colony hybridization experiment with the oligomer CYMVP is shown in Figure 3.17. Two out of 52 harvested clones produced strong hybridization signals and were used for further verification. Figure 3.18 shows the result of an RNA colony hybridization experiment with the PMVP oligomer. Three out of 52 harvested clones were found to produce strong hybridization signals and were selected for further verification by dideoxynucleotide sequencing analysis (Figure 3.19).

3.4.1 EFFICIENCY OF CYMVP AND PMVP TO INITIATE TRANSLATION OF CAT mRNA IN *E.COLI* CELLS

The relative translation initiation activity of the constructs carrying CYMVP and PMVP as the initiators of translation in *E.coli* LE392 over that of the consensus SD sequence is shown in Table 3.3. Relative activities of the constructs carrying CYMVP and PMVP oligomers were determined to be 7% and 20%, respectively. For

CYMVP OLIGOMER

XhoI

5′ -**TCGAG**AAAACAAAACGAAAACAAACAAAATCTTC-
CTTTTGTTTTGCTTTTGTTTGTTTTAGAAG-

-GTAAACACCCTCCAACACACCATTCAATCCCGAATCG-
-CATTTGTGGGAGGTTGTGTGGTAAGTTAGGGCTTAGC-

-CCCGCAAACCAGCAAAGATTCGCAT**A**-3′
-GGGCGTTTGGTCGTTTCTAAGCGTA**TTCGA**
HindIII

PMVP OLIGOMER

XhoI

5′ -**TCGAG**AAAAGAAACACAAAGCAAAGCAAAGCAAA-
CTTTTCTTTGTGTTTCGTTTCGTTTCGTTT-

-GCAACTCAAATAAACCATATTTGGCCAAGGCACTTGG-
-CGTTGAGTTTATTTGGTATAAACCGGTTCCGTGAACC-

-TAATCAAACGGGCACAACCCTAGATTAACG**A**-3′
-ATTAGTTTGCCCGTGTTGGGATCTAATTGC**TTCGA**
HindIII

Figure 3.16: Nucleotide sequences of CYMVP and PMVP oligomers.
Restriction sites are in bold letters.

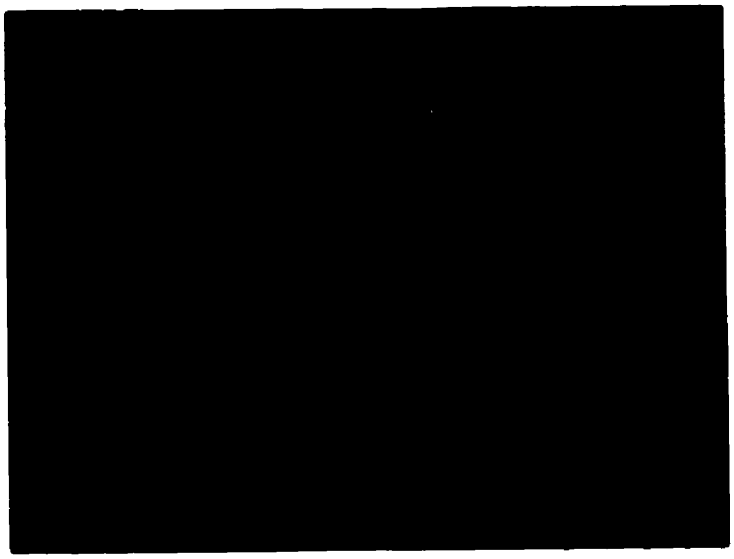


Figure 3.17: Screening of CYMVP constructs using DNA colony hybridization. The constructs were screened using ^{32}P labelled CYMV18 18mer DNA fragments. Two of the constructs showed strong hybridization signals.

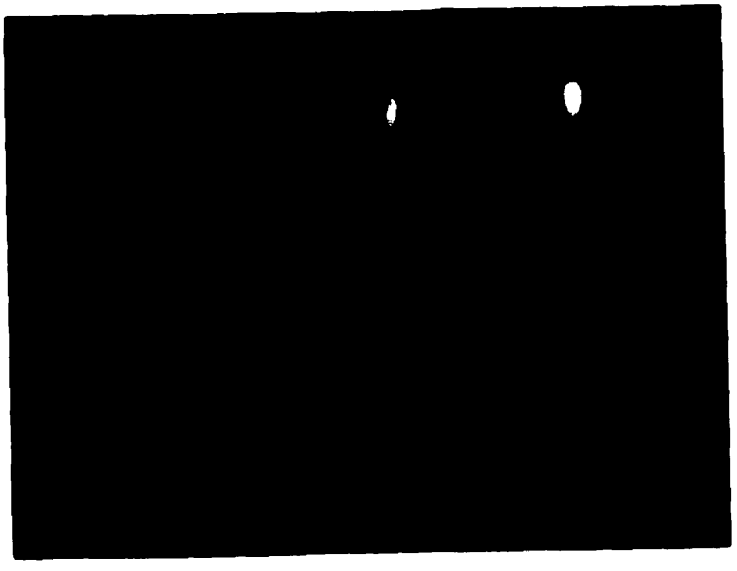
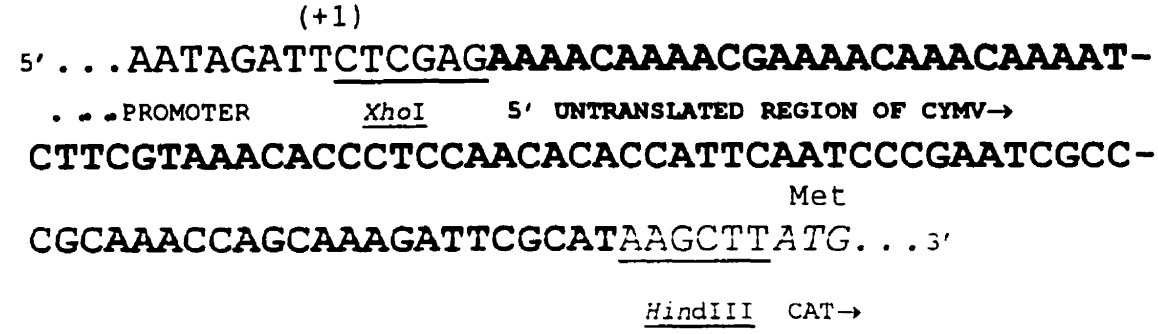


Figure 3.18: Screening of PMVP constructs using RNA colony hybridization. The constructs were screened using ^{32}P labelled PMV18 18mer DNA fragments. Three of the constructs showed strong hybridization signals.

CYMVP CONSTRUCT



PMVP CONSTRUCT

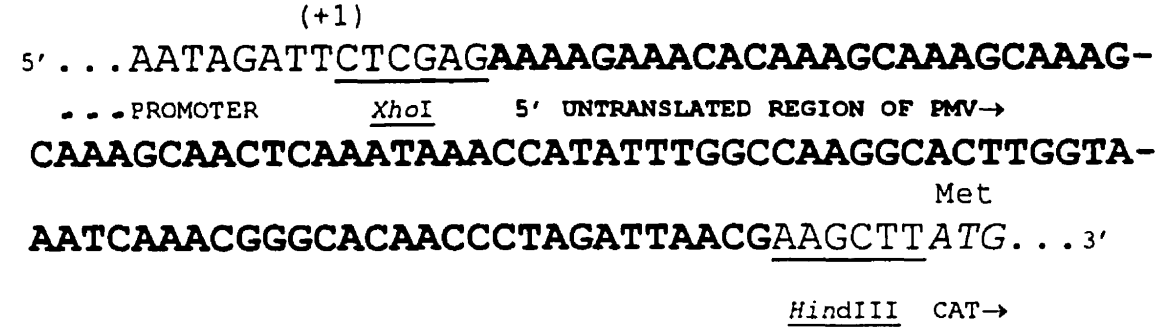


Figure 3.19: Nucleotide sequence of the 5′ end non-coding region of the CAT gene for the constructs carrying CYMVP and PMVP oligomers. Restriction sites are underlined and the 5′ untranslated viral leader sequences are in bold letters. (+1) indicates the start of transcription.

CONSTRUCT	RELATIVE ACTIVITY (CONSTRUCT/SD)
SD	$1.00 \pm 0.1 \times 10^{-1}$
SD_n	$2.3 \times 10^{-1} \pm 0.1 \times 10^{-1}$
Ω	$1.0 \times 10^{-1} \pm 0.1 \times 10^{-1}$
CYMVP	$0.7 \times 10^{-1} \pm 0.1 \times 10^{-1}$
PMVP	$2.0 \times 10^{-1} \pm 0.1 \times 10^{-1}$

Table 3.3: Relative efficiency of CYMVP and PMVP oligomers in mediating initiation of CAT mRNA translation in *E.coli*. Values represent the average ratio (n=3) of translational activity for each construct over that of the consensus SD sequence. SD is the consensus SD sequence and SD_n is the SD sequence native to the CAT gene. Ω is the 5′ untranslated sequence of TMV. ± represents the standard deviation.

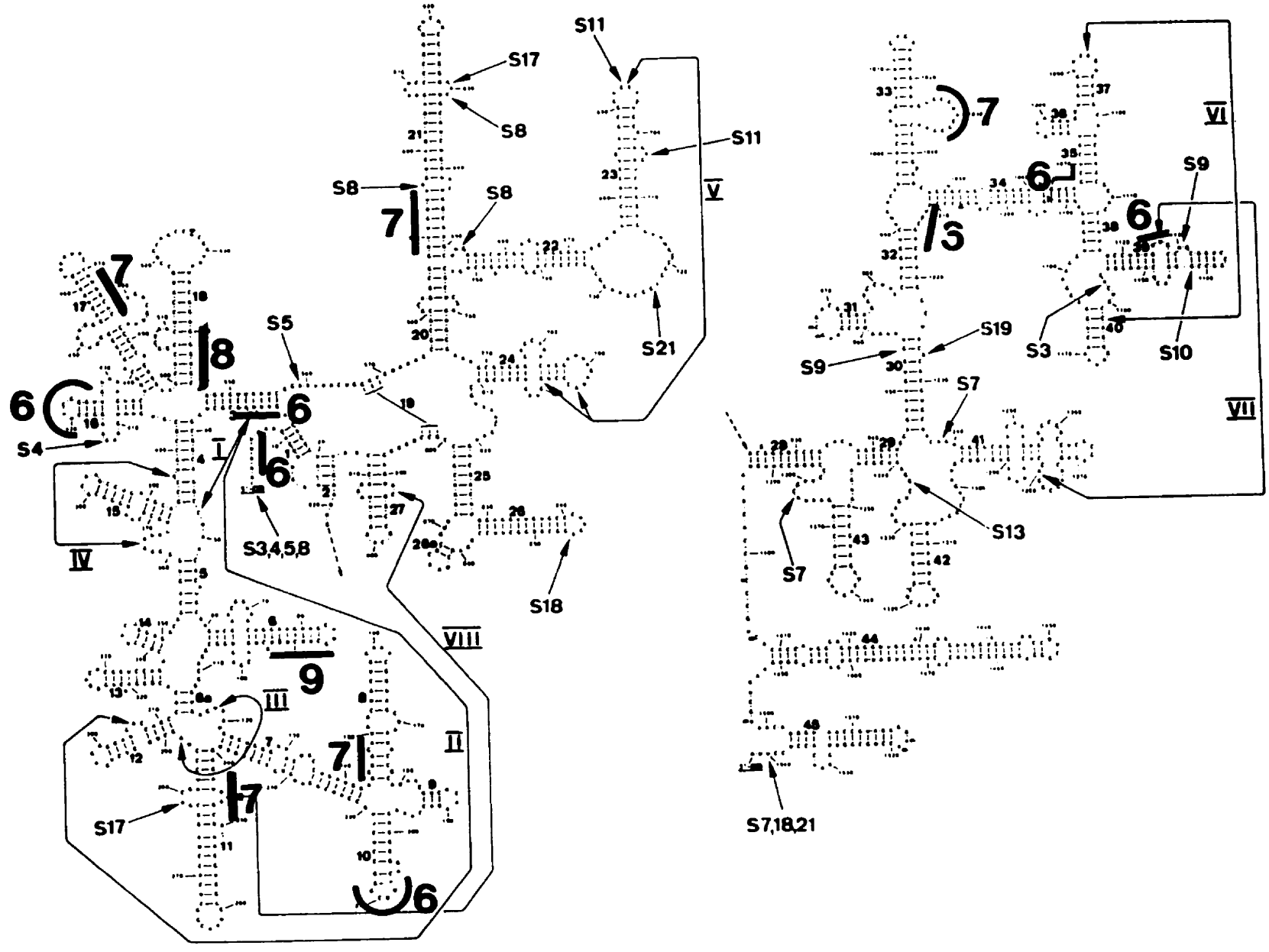
comparison, it should be mentioned that the relative CAT activity for the construct carrying TMV Ω sequence was only 9%.

3.4.2 SEQUENCE COMPLEMENTARITY BETWEEN 16S rRNA AND THE CYMVP AND PMVP OLIGOMERS

The result of sequence complementarity search for CYMVP and PMVP against 16S rRNA, is shown in Figures 3.20 and 3.21 respectively. Regions of 16S rRNA which have 6 or more base pairs complementary to the oligomer are indicated with solid lines.

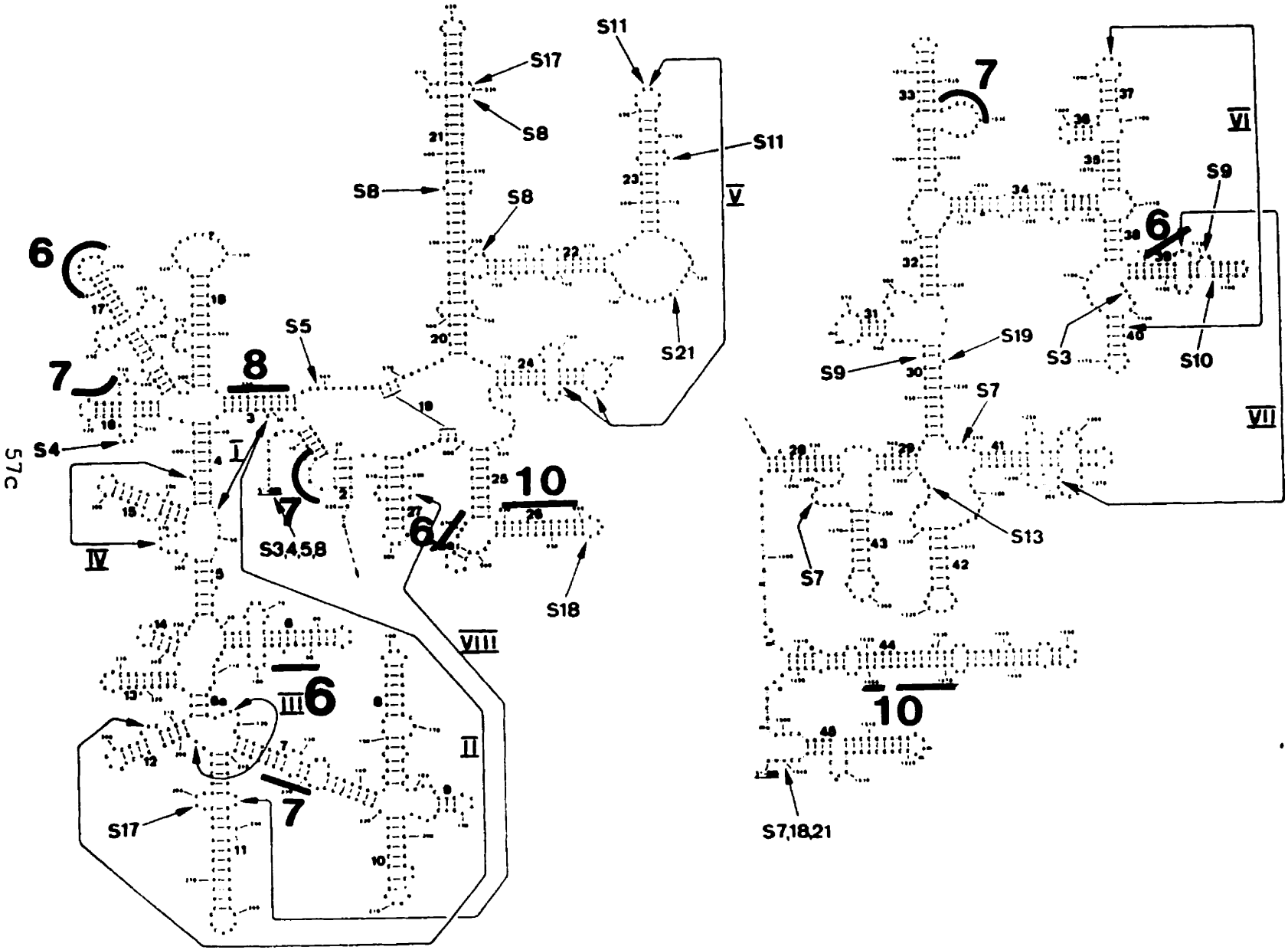
3.5 BACKGROUND LEVEL OF CAT ACTIVITY

To measure the background level of CAT activity for plasmid pP1'SDCAT two control constructs were prepared. In the first construct (5NT) the SD sequence was replaced by a 5 nucleotides long sequence (5'-AGCTT), by first making the *Xho*I site blunt-ended in plasmid pP1'SDCAT using mung bean nuclease and then filling the *Hin*dIII site with Klenow followed by blunt-end ligation of the plasmid. This allows the determination of CAT activity in the absence of an SD-like sequence in plasmid pP1'SDCAT. To measure the effect of the length of the untranslated sequence on the CAT activity, a second construct (L36) was prepared by cloning a 36 nucleotide-long sequence (L36 oligomer), in place of the SD sequence in plasmid pP1'SDCAT. The L36 oligomer (Figure 3.22) was designed to be non-complementary to the known mRNA binding sites in the 16S


S11
S17
S8
S11
S8
S8
V
7
S21
7
S5
8
6
6
6
S4
I
6
S3,4,5,8
IV
S18
VIII
9
III
II
7
7
S17
6
7
6
S19
S9
S3
S10
S9
S7
VII
S13
S7
S7,18,21

57a

Figure 3.20: Sequence complementarity between CYMVP oligomer and 16S rRNA. Complementary regions between the oligomer and 16S rRNA are indicated by thick solid lines; the number of complementary nucleotides is indicated next to each line. The secondary structure of *E.coli* 16S rRNA is taken from Brimacombe (1992; see Figure 3.6).


S11
S17
S8
S11
S8
S8
V
6
S5
8
7
S4
I
7
S3,4,5,8
6
10
IV
S18
VIII
III 6
II
7
S17
7
33
37
VI
S9
6
34
38
39
32
S19
S3
S10
S9
30
S7
VII
29
29
41
S13
S7
43
42
44
10
45
S7,18,21

57c

Figure 3.21: Sequence complementarity between PMVP oligomer and 16S rRNA. Complementary regions between the oligomer and 16S rRNA are indicated by thick solid lines; the number of complementary nucleotides is indicated next to each line. Secondary structure of *E.coli* 16S rRNA is taken from Brimacombe (1992; see Figure 3.6).

L36 OLIGOMER

XhoI

5′-TCGAGCAATATCACCCTCTTTTACAATTACA
CGTTATAGTGGGAGAAAATGTTAATGTTCGA

HindIII

Figure 3.22: Nucleotide sequence of L36 oligomer. Restriction sites are underlined.

rRNA (SD, Ω, and DB binding sites) and contained two sticky ends; one compatible with the *Xho*I and the other with *Hin*dIII overhangs.

The screenings of the 5NT and L36 constructs were carried out by *Xho*I/*Hin*dIII digests and RNA colony hybridization (Figure 3.23), respectively, as outlined earlier. Selected clones were verified by dideoxynucleotide sequencing (Figure 3.24).

3.5.1 MEASURE OF CAT ACTIVITY FOR THE CONTROL CONSTRUCTS

Relative translation initiation activities of control constructs in *E.coli* LE392 cells are shown in Table 3.4. Values are presented as the ratio of translational activity of each construct over that of the consensus SD. Relative activities of the constructs carrying 5NT and L36 were 1% and 2%, respectively.

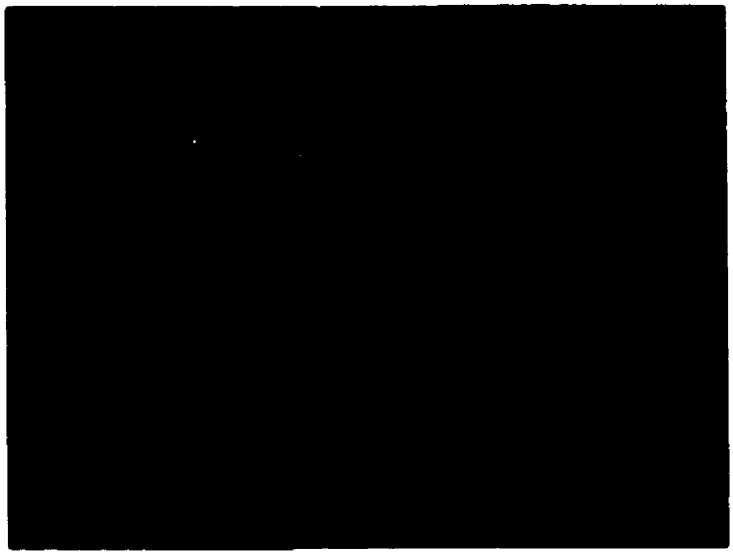


Figure 3.23: Screening of L36 constructs using RNA colony hybridization. The constructs were screened using ^{32}P labelled L36B oligomers. One of the constructs showed a strong hybridization signal.

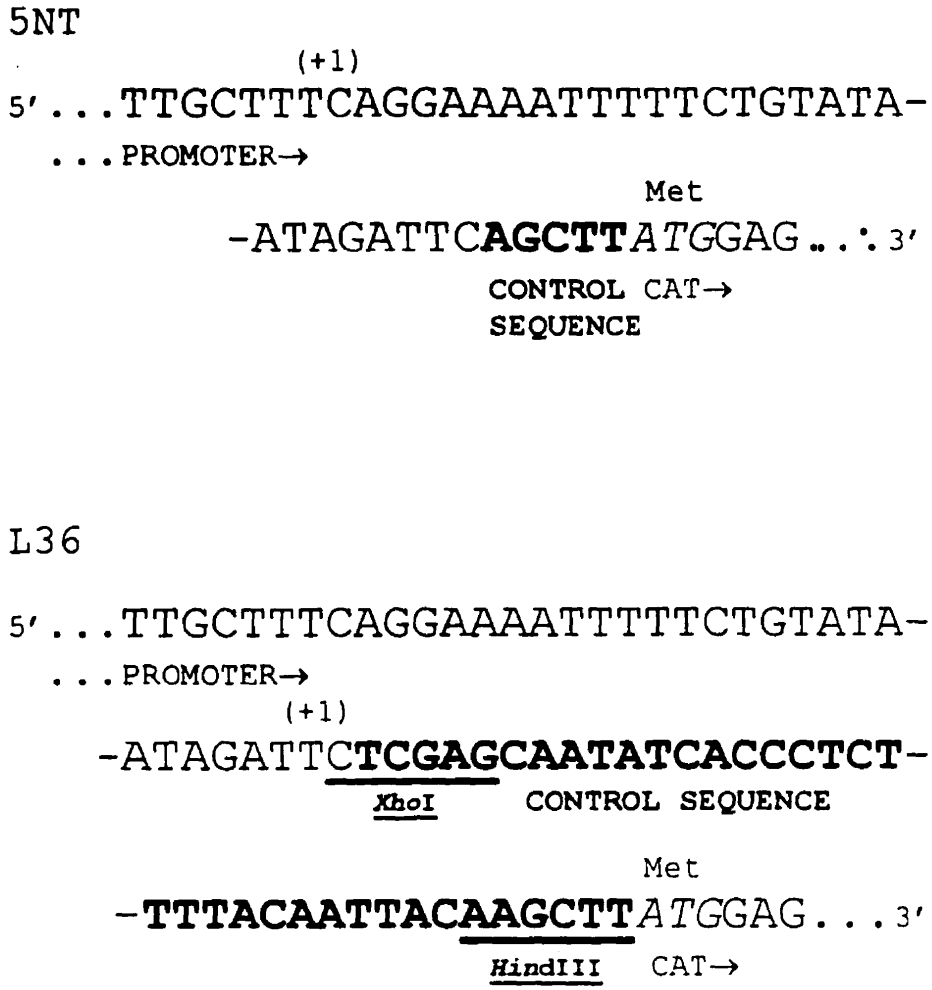


Figure 3.24: Nucleotide sequence of 5′ non-coding region of the CAT gene for 5NT and L36 constructs. Restriction sites are underlined and the control sequences (used to measure the background level of CAT activity) are in bold letters. (+1) is the start of transcription.

CONSTRUCT	RELATIVE ACTIVITY (CONSTRUCT/SD)
SD	$1.00 \pm 0.1 \times 10^{-1}$
SD_n	$2.3 \times 10^{-1} \pm 0.1 \times 10^{-1}$
5NT	$0.1 \times 10^{-1} \pm 0.1 \times 10^{-1}$
L36	$0.2 \times 10^{-1} \pm 0.1 \times 10^{-1}$

Table 3.4: Relative efficiencies of control oligomers in mediating the initiation of CAT mRNA translation in *E.coli*. Values above represent the average ratio (n=3) of translational activity for each construct over that of the consensus SD sequence. SD is the consensus SD sequence and SD_n is the SD sequence native to the CAT gene. ± represents the standard deviation.

4. DISCUSSION

4.1 LOOPS # 42 AND # 43 OF 16S rRNA AS ALTERNATIVE SITES FOR COMPLEMENTARY INTERACTION WITH mRNAs

In the tertiary structure model of *E.coli* 30S ribosomal subunit the SD and omega binding sites and nucleotides 1470-1482 (all proposed to be responsible for complementary interaction with the mRNAs in mediating initiation of translation) are described to be in close proximity to one another (see Figure 4.1 and Brimacombe, 1995; Ofengand *et al.*, 1993; Ivanov *et al.*, 1995). These results raised the possibility of the presence of other mRNA interaction sites on the 30S ribosomal subunit within the vicinity of the omega binding site and the 3' end of 16S rRNA (SD binding site). To investigate this possibility, two oligomers (Ω42 and Ω43) were designed to be complementary to two adjacent loops on each side of the omega binding site on the 16S rRNA. The effect of these primers on the initiation of translation was then studied by placing them in front of the AUG start codon of the CAT mRNA.

Oligomer Ω42 showed a relative translational efficiency of 6% (with respect to the consensus SD sequence) in mediating initiation of CAT mRNA in *E.coli*, which is 3 times higher than the background level (1 and 2% for constructs designated 5NT and L36, respectively). The effect of this oligomer on translation initiation may be explained by its complementarity to loop # 42 (nucleotides 1316-1326) of 16S rRNA which could form a possible 11 base pair interaction.

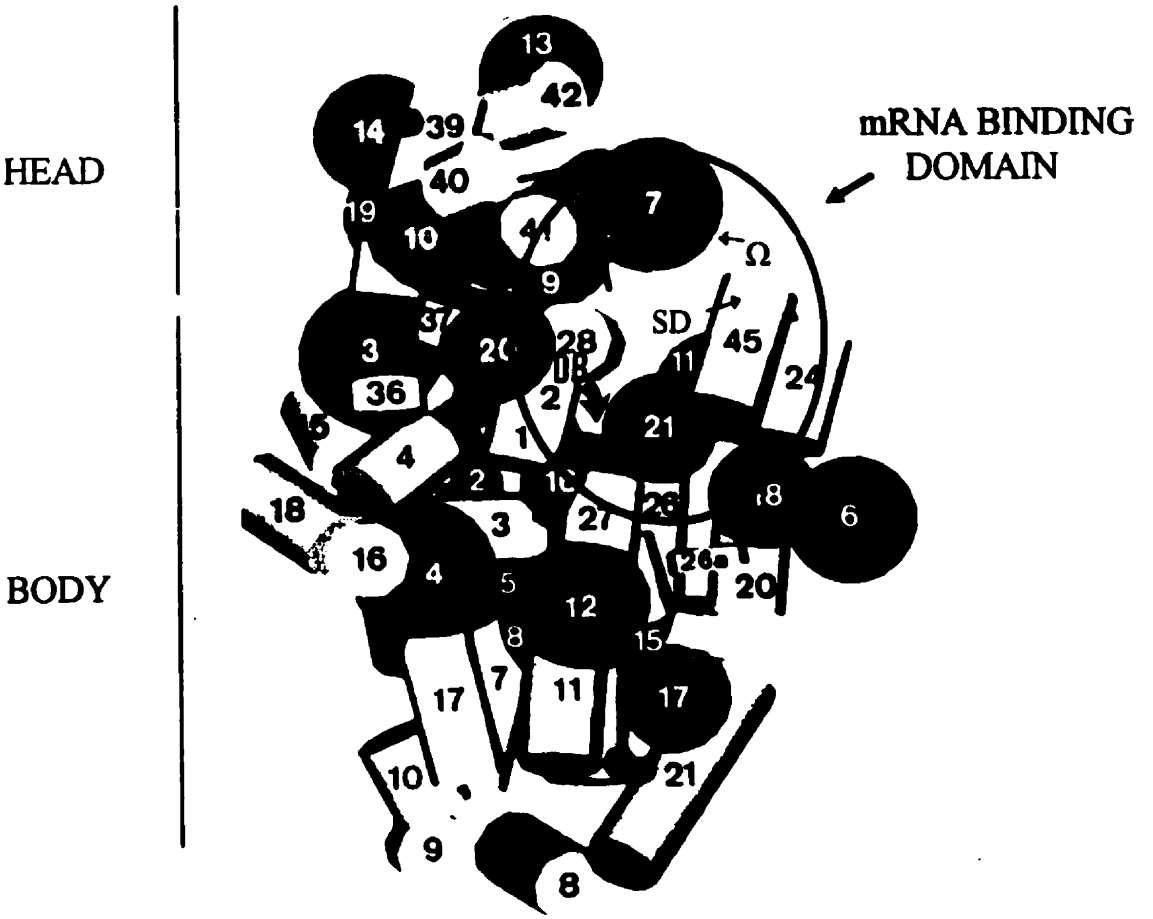


Figure 4.1: Schematic diagram illustrating the position of mRNA binding sites in the 30S ribosomal subunit. The SD, Ω, and DB binding sites are enclosed in an ellipse. Ribosomal proteins are represented by dark spheres with white numbers, and 16S rRNA helical domains are represented by cylinders with black numbers. (Adapted from Brimacombe, 1995).

When the distance between the complementary region of Ω42 and the AUG start codon increased from 21 to 36 bases (Ω42b), the relative efficiency of translation of the oligomer increased to 72% which is more than 3 times higher than the relative efficiency of the SD sequence for the native CAT gene (23%, for SD_n).

At present, it is difficult to elaborate on the difference between the ability of Ω42 and Ω42b to mediate initiation. The simplest explanation however, is that the difference in spacing between the complementary regions of the oligomer and the AUG start codon makes the start codon more accessible for the ribosome. This explanation is supported by the fact that slight variations in SD-AUG spacing were shown to dramatically influence gene expression (Stroynowski *et al.*, 1982; Singer *et al.*, 1981; Kozak, 1983; Gheysen *et al.*, 1982).

In contrast, initiation of translation is shown to be independent of the spacing between AUG and the previously published non-SD sequences. Changing the spacing between omega and the start codon from 42 to 10 nucleotides yielded no significant difference in the level of expression (Gallie *et al.*, 1988; Gallie and Kado, 1989). Also, the effect of epsilon on initiation did not differ when the sequence was placed 5′ or 3′ from the start codon with a distance of 17 and 5 nucleotides, respectively (Olins and Rangwala, 1989). Furthermore, increasing the distance between the start codon and the downstream box (DB) derived from bacteriophage T7 gene *10* RBS, from 1 to 10 nucleotides showed no change in the level of expression (Sprengart *et al.*, 1996).

It could be assumed that the formation of a strong secondary (and/or tertiary) structure is responsible for the reduced activity of Ω42 in the initiation of translation. Similarly, it may be that formation of a secondary and/or tertiary structure may place the AUG codon in a position more accessible for the initiation process in the Ω42b construct. It is known that the presence of secondary and/or tertiary structures involving either the SD sequence (Saito and Richardson, 1981) or the start codon (Iserentant and Fiers, 1980) or both (Gheysen *et al.*, 1982; Johnston and Roth, 1982) could even eliminate translation completely (Simons and Kleckner, 1983; McPheeters *et al.*, 1986). It has been proposed that the secondary structures which do not involve the SD sequence or the start codon of an RBS, could in some cases increase the expression by making the AUG codon more accessible to the ribosomes (Helke *et al.*, 1993).

The proposed secondary structure of the RBS for CAT mRNA in the Ω42 construct is similar to that of Ω42b, with no detectable formation of inhibitory structures. This however, does not rule out the possibility of a secondary and/or tertiary structure which is not detected by the computer or the involvement of Ω42 in RNA-protein interactions that influence the difference in expression between these two constructs.

The difference in the sequence of the spacers between the complementary region and the start codon in Ω42 and Ω42b constructs may also be responsible for the difference in expression between them. It has previously been shown that the sequence used as the spacer in Ω42 construct does not influence the expression of CAT

mRNA (I. Ivanov, unpublished data). The effect of the duplicate form of this sequence (the spacer in Ω42b) however, has not been examined.

Gene expression studies with the sequence Ω43 cloned in front of the CAT gene showed that this sequence mediates initiation with a relative efficiency of 22% which is almost the same as with SD_n (23%). The capability of Ω43 to initiate translation may be explained by its complementarity to loop # 43 (nucleotides 1358-1366).

Through the use of an RNA-protein binding site protection approach and protein foot-printing technique, loops # 42 and # 43 of 16S rRNA are shown to be protected by protein S19 (Stern *et al.*, 1988; Brimacombe, 1991; Brimacombe, 1992). Such protection against specific chemical modification or nuclease digestion however, does not necessarily imply that the loops are unavailable for mRNA binding. Recent understanding of the secondary and tertiary structures of 16S rRNA has made it clear that RNA-protein binding site protection studies mostly reflect the stability of the RNA regions concerned (Brimacombe, 1991; Brimacombe, 1992; Verschoor *et al.*, 1993). The foot-printing method also has limitations. It is a) very sensitive to allosteric effects, as explained by the fact that in most cases, the binding of a particular protein gives rise to sites on the RNA showing enhanced reactivities relative to the control samples, as well as b) very sensitive to sites showing protection effects with various modifying reagents or nucleases (Brimacombe, 1991; Brimacombe, 1992; Verschoor *et al.*, 1993).

In contrast to observations from protein protection studies, loops # 42 and # 43 are not observed to be closely associated with any proteins in RNA-protein cross-link studies (Brimacombe, 1992). The cross-linking method provides direct information concerning the RNA-protein vicinity without necessarily implying that a physical binding exists between these components.

The ability of Ω42, Ω42b and Ω43 primers to mediate the initiation of CAT mRNA suggests that loops # 42 and # 43, adjacent to omega binding site on 16S rRNA, may be used as alternative mRNA binding sites for initiation of translation. Since these domains are located at the bottom part of the head of 30S ribosomal subunit on the opposite site of the cleft in front of the anti-SD site, one could assume that a broad area on the surface of the small ribosomal subunit in that vicinity is involved in the interaction with mRNA and could initiate translation of mRNAs which are devoid of SD sequence.

4.2 EFFECT OF EPSILON SEQUENCE ON INITIATION OF TRANSLATION

The presence of the epsilon sequence together with the consensus SD sequence in front of the CAT mRNA in *E.coli* (ϵSD construct) had an enhancing effect on the translation as compared to the initiation mediated by the consensus SD sequence alone. It was observed that the relative activity of the ϵSD construct was 2.5 times higher than that of the consensus SD. This observation is in agreement with previous studies by Olins and Rangwala (1989), who

reported an enhancing effect for epsilon sequence over the initiation mediated by consensus SD sequence in *E.coli*. They proposed that this stimulatory effect might be due to the complementary region between epsilon and nucleotides 458-467 of 16S rRNA.

To examine the ability of epsilon to mediate initiation, a CAT mRNA was placed under the control of epsilon in the absence of an SD sequence (ϵI construct). It was found that the relative efficiency of epsilon in mediating translation was similar to the background level (1 and 2% for 5NT and L36, respectively). However, when the complementarity between epsilon and 16S rRNA was increased from 10 to 14 nucleotides (ϵII construct), the relative efficiency of the primer to mediate initiation was observed to be 85% which is more than 3 times higher than that of SD_n (Table 3.2). This suggests that epsilon alone may have the potential to mediate initiation when it is provided with better binding abilities to the 16S rRNA. Subsequently, the region between nucleotides 451-467 of the 16S rRNA may have the capability of binding mRNA and hence be used to mediate the initiation of translation in *E.coli*.

In the tertiary structure, the epsilon binding site (nucleotides 451-467) is located in domain # 17 which is close to neither the omega, nor the SD binding sites (see Figure 4.2 and Brimacombe, 1992; Ofengand *et al.*, 1993). The RNA-protein binding site protection approach, the protein foot-printing technique and RNA-protein cross-link analysis have all suggested that the epsilon binding region is free of ribosomal proteins and therefore, it is

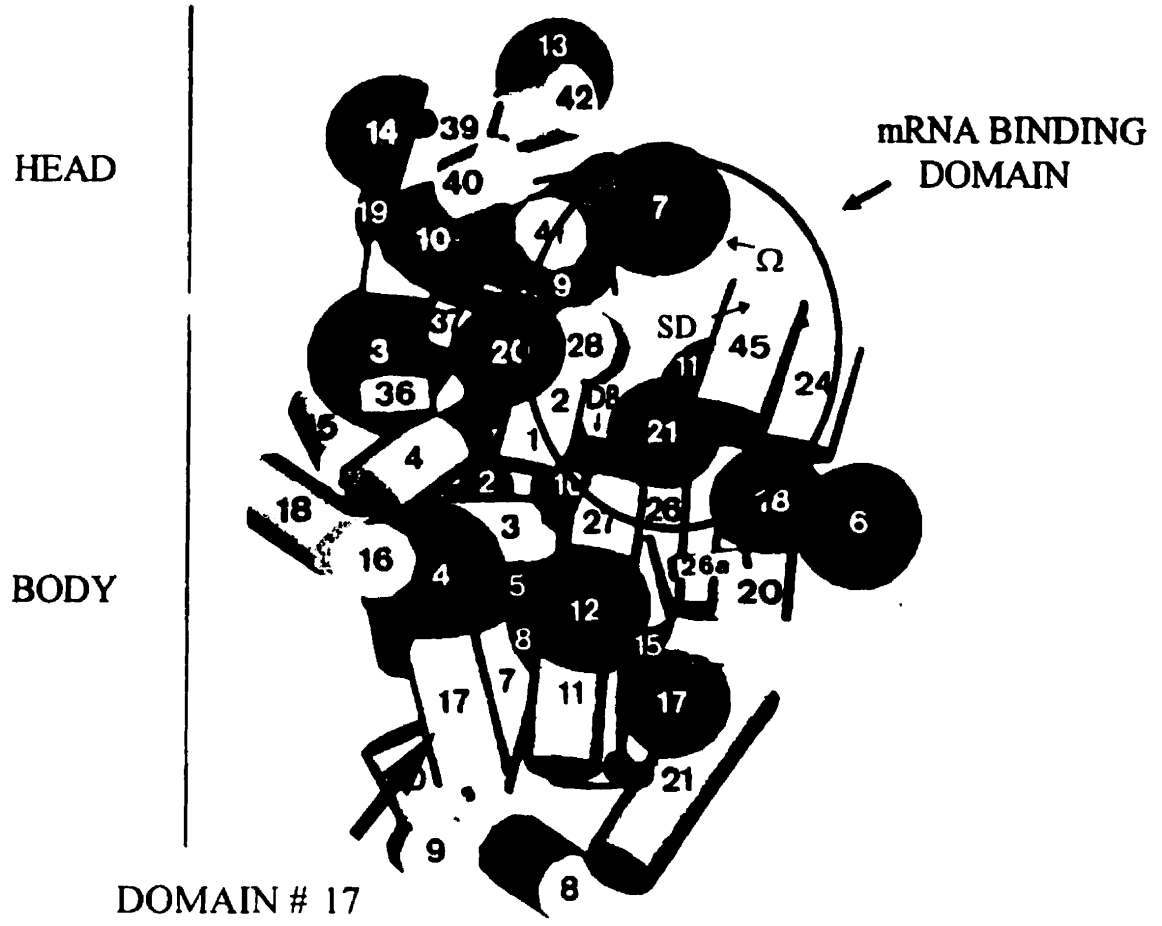


Figure 4.2: Schematic diagram illustrating the position of domain # 17 in the 30S ribosomal subunit. The SD, Ω, and DB binding sites are enclosed in an ellipse. Ribosomal proteins are represented by dark spheres with white numbers, and 16S rRNA helical domains are represented by cylinders with black numbers. (Adapted from Brimacombe, 1995).

not involved in RNA-protein interactions (Brimacombe, 1991; Brimacombe, 1992; Verschoor *et al.*, 1993). Although the helical domain # 17 of 16S rRNA is localized in the lower part of the body in the three-dimensional model of 30S ribosomal subunit, conditionally it could be considered as a part (or in the vicinity) of the "530 nucleotide" neighbourhood. As already discussed, this neighbourhood together with the nucleotides from the 1400-1500 neighbourhood are involved in the organization of the decoding centre of *E.coli* ribosomes. The lack of ribosomal proteins associated tightly with the nucleotides belonging to the helical domain # 17 suggests that they are available for RNA-RNA interaction with complementary nucleotides in mRNA. The close proximity of the two helical domains # 17 and # 18 (the latter carrying the 530 neighbourhood nucleotides) and the positive effect of the extension of complementarity of ϵ sequence on gene expression is a reason to assume that this domain is involved in the organization of a very efficient (alternative to SD) mRNA binding site on the *E.coli* ribosome. To exploit the efficiency of ϵ as an alternative translational initiator however, the complementary sequence to the ϵ binding site must be extended in comparison with the complementarity to the canonical anti-SD sequence.

If the nucleotides from the helical domain # 17 might participate in an interaction with mRNA leading to initiation of translation, then the following question arises: How can nucleotides in a domain located at the bottom part of the 30S

ribosomal subunit (as in domain # 17) bring the AUG codon to the ribosomal P-site which is known to be located between the head and the body of the 30S subunit (see Figure 4.3).

This question cannot be answered now in the light of the current three-dimensional models of the 30S ribosomal subunit. Until recently a similar controversy existed in relation to the localization of the "530" neighbourhood nucleotides. According to the RNA cross-linking data (Shatsky *et al.*, 1991) these nucleotides are related to the decoding centre (A, P and E sites) whereas according to others (see Brimacombe, 1995) they were found far away from the decoding centre (close to domain # 17). Only recently the nucleotide domain # 18 was relocated (Heilek and Noller, 1996) and the 530 neighbourhood nucleotides were placed closer to the decoding centre.

The nucleotide domain # 17 has not attracted very much attention of the investigators so far and it is still at the same place as in the first model of the 30S subunit (Brimacombe *et al.*, 1988). However, we believe that it will be relocated soon, by the same method as was done with the neighbour helical domain # 18.

4.3 THE 5'-TERMINAL UNTRANSLATED REGIONS OF CYMV AND PMV RNAs AS TRANSLATIONAL INITIATORS

The capability of the 5' untranslated regions of CYMV and PMV RNA to initiate translation in *E.coli* was studied by placing synthetic variants of these sequences in front of a CAT mRNA in

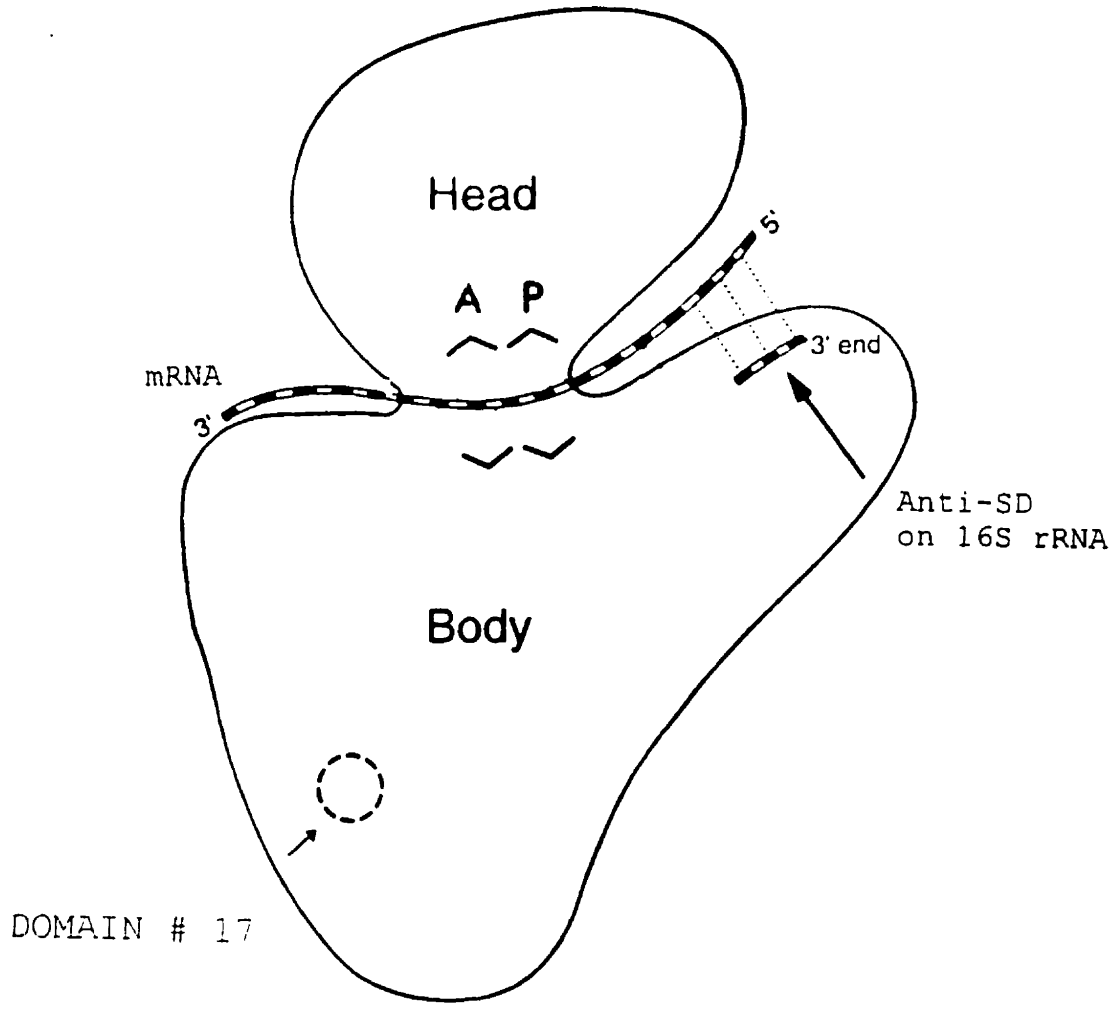


Figure 4.3: Schematic diagram illustrating the positions of the A and P sites in the 30S ribosomal subunit. The A (aminoacyl) and P (peptidyl) sites are located towards the reader on the three dimensional structure of the ribosome. (Adapted from McCarthy and Brimacombe, 1994).

E.coli. Our results showed that the CYMV-mediated initiation of translation was 7% of that of the consensus SD sequence and about 3 times the background level (1 and 2% for 5NT and L36, respectively). The PMV sequence was much more efficient than CYMV reaching a relative efficiency of 20% which is more than twice that of omega (9%) and almost as efficient as the natural SD_n sequence of the CAT gene (23%).

To understand the mechanism by which these two sequences might interact with the *E.coli* ribosomes, a computer homology analysis was carried out. Taking into consideration the low level of expression of CAT mRNA containing CYMV sequence, we can conclude that its interaction with the possible complementary sites shown in Figure 3.20 is not very efficient.

More interesting are the data obtained with the PMV sequence. Among the several possible sites of interaction with 16S rRNA (see Figure 3.21) two of are worthy of discussion. As seen in Figure 4.4 the PMV sequence contains 6 nucleotides complementary to the epsilon and 10 nucleotides complementary to the DB binding sites.

The ϵ-like sequence is located 7 nucleotides before the initiation codon and the DB like sequence is found 74 nucleotides upstream of the ϵ sequence.

In a recent study Sprengart *et al.* (1996) came to the conclusion that "the DB is an efficient and independent translational initiation signal in *Escherichia coli*". Since the role of the spacer region between the DB (as a translational initiator) and the initiation codon is not studied, we cannot determine whether

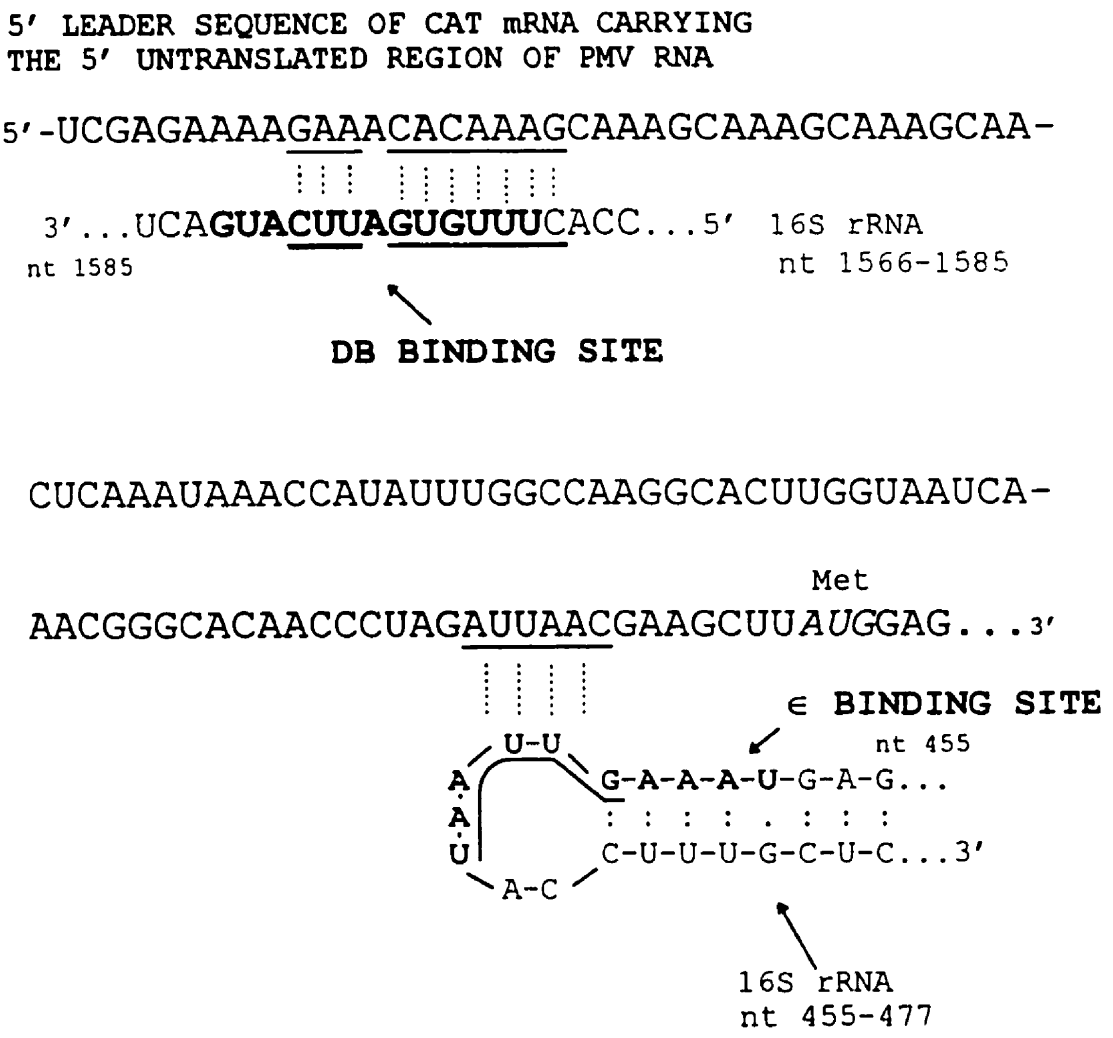


Figure 4.4: Potential base pairing of PMVP to DB and epsilon binding sites of 16S rRNA. The DB and ∈ binding sites are in bold letters. The complementary regions between PMVP and the binding sites are underlined and indicated by dashed lines.

or not it is possible to have an efficient translation initiation of CAT mRNA carrying the DB-like sequence found here. However, the results of others (Olins and Rangwala, 1989) as well as our results presented above show that the ϵ sequence alone is not an efficient translational initiator with its natural complementary track. Taking all these considerations into account we conclude that the high level of translation of CAT mRNA carrying the PMV sequence is due to the interaction of both ϵ-like and DB-like sequences with the ribosome, thus producing a mutual enhancing effect which results in a proper initiation of translation.

4.4 CONCLUSIONS AND HYPOTHESES

The results presented in this thesis clearly show that the *E.coli* ribosome contains more than one mRNA binding site. Besides the conventional (anti-SD) binding site used by the vast majority of *E.coli* mRNAs, other domains in 16S rRNA, rather than the 3'-terminal nucleotides, are also capable of binding mRNA in a way leading to a functional assembly of the translation initiation complex and to a real initiation of translation. Surprisingly, some of these non-SD binding sites are even more efficient than the canonical anti-SD site.

Due to the scarce information about the structure and organization of these new mRNA binding sites on the *E.coli* ribosome, it is impossible, at the present time, to conclude whether they are naturally in use. A more comprehensive computer analysis of the 5'

leader sequences of natural prokaryotic mRNAs is required.

Most of the alternative, non-SD translation initiator sequences were found in highly translated plant viral or phage mRNAs. The question might be asked: Is it coincidental or not?

The non-SD translational initiators might be evolutionarily old sequences preserved in some recent organisms and still active in bacteria. However, they could also be a product of a recent evolution. If this is the case, then the plant viral RNAs could utilize not only the cytoplasmic protein synthesis apparatus, but also the chloroplast and mitochondrial machinery to translate their genomes. This is consistent with some observations from studies on TMV. It has previously been shown that TMV coat protein can be synthesized efficiently using a chloroplast system (Camerino *et al.*, 1982). TMV RNA has been found within chloroplasts isolated from infected tobacco cells (Schoel and Zaittin, 1989) and was also shown to be capable of entering chloroplasts of tobacco cells (Schoel and Zaittin, 1989). The presence of TMV coat protein in chloroplasts of infected tobacco cells was further confirmed by electron microscopy (Osbourn *et al.*, 1989).

4.5 FUTURE EXPERIMENTS

We plan to investigate other mRNA binding sites on 16S rRNA by using random primers. These primers lack guanine residues and consequently only sites capable of mediating initiation in a non-SD

manner will be identified.

We are also planning to examine the ability of omega, loop (# 42 and 43), and epsilon primers to form complementary base pair interactions with sequences of 16S rRNA using site-specific mutation studies. A plasmid-borne specialized ribosome system (kindly provided by Dr. Lea Braker Gingras) which can supply up to 75% of the total ribosome population of *E.coli* cells, will be used to study the effect of an increase or decrease in complementarity between 16S rRNA and the mediators of initiation.

5. LITERATURE CITED

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J. D. (1994). *Molecular biology of the cell*. 3rd ed. Garland Publisher, New York. p. 134.

Biggin, M. D., Gibson, T. J. and Hong, G. F. (1983). Buffer gradient gels and [^{35}S] label as an aid to rapid DNA sequence determination. *Proceedings of the National Academy of Sciences, U.S.A.* 80: 3963-3965.

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

Bradford, M.M. (1976). A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein dye binding. *Anal of Biochemistry* 72: 248-254.

Brimacombe, R. (1991). RNA-protein interaction in the *Escherichia coli* ribosome. *Biochimie* 73: 927-936.

Brimacombe, R. (1992). Structure-function correlations (and discrepancies) in the 16S ribosomal RNA from *Escherichia coli*. *Biochimie* 74: 319-326.

Brimacombe, R. (1995). The structure of ribosomal RNA: a three-dimensional jigsaw puzzle. *European Journal of Biochemistry* 230: 365-383.

Brimacombe, R., Atmadja, J., Stiege, W. and Schuler, D. (1988). A detailed model on the tree-dimensional structure of *E.coli* 16S ribosomal RNA *in situ* in the 30S subunit. *Journal of Molecular Biology* 199: 115-136.

Camerino, G., Savi, A. and Ciferri, O. (1982). A chloroplast system capable of translating heterologous mRNAs. *FEBS Letters* 150: 94-98.

Chapon, C. (1982). Expression of *malT*, the regulator gene of the maltose regulon in *Escherichia coli*, is limited both at transcription and translation. *EMBO Journal* 1: 369-374.

Christie, G. E. and Callendar, R. (1985). Bacteriophage P2 late promoters. II. Comparison of the four late promoter sequences. *Journal of Molecular Biology* 181: 373-382.

Cunningham, P. R., Nurse, K., Weitzman, C. J. and Ofengand, J. (1993). Functional effects of base changes which further define the decoding centre of *Escherichia coli* 16S ribosomal RNA: mutation of C1404, G1405, C1496, G1497 and U1498. *Biochemistry* 32: 7172-7180.

Czernilofsky, A.P., Kurland, C.G. and Stoffler, G. (1975). Localization of protein S1 in 30S ribosomal subunit. *FEBS Letters* 58: 281-284.

De Boer, H.A. and Shepard, H.M. (1983). Strategies for optimizing foreign gene expression in *Escherichia coli*. in *Horizones in Biochemistry, vol.7. Genes, Structure and Expression*. Kroon, A.M. ed. John Wiley ans Sons, Inc., New York. pp. 205-248.

Doring, T., Mitchell, P., Osswal, M. and Bochkariov D. and Brimacombe, R. (1994). The decoding region of 16S rRNA: a cross-linking study of the ribosomal A, P and E sites using tRNA derivatized at position 32 on the anticodon loop. *EMBO Journal* 13: 2677-2685.

Draper, D.E. (1993). Mechanism of Translational Initiation and Repression in Prokaryotes. in *The Translational Apparatus*. Nierhause, K. H., Franceschi, F., Subramanian, A. R., Erdmann, V. A. and Wittmann-Liebold, B. eds. New York: Plenum Press, pp. 197-207.

Dunn, J. J., Buzash-Pollert, E. and Studier, F. W. (1978). Mutations of bacteriophage T7 that affect initiation of synthesis of the gene *0.3* protein. *Proceedings of the National Academy of Sciences, U.S.A.* 75: 2741-2745.

Dunn, J.J. and Studier, F.W. (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *Journal of Molecular Biology* 166: 477-535.

Eeasterwood, T. R. and Harvey, S. C. (1995). Modelling the structure of the ribosome. *Biochemistry and Cell Biology* 73: 751-756.

Faxen, M., Plumbridge, J. and Isaksson, L. A. (1991). Codon choice and potential complementarity between mRNA downstream of the initiation codon and bases 1471 - 1480 in 16S ribosomal RNA affects expression of *glnS*. *Nucleic Acids Research* 19: 5247-5251.

Frank, J., Vershoor, A., Li, Y., Zhu, J., Lota, R. K., Radermacher, M., Penczek, P., Grassucci, R., Agrawi, R. K. and Srivastava, S. (1995). A model of the translational apparitor based on a three-dimensional reconstruction of the *Escherichia coli* ribosome. *Biochemistry and Cell Biology* 73: 757-765.

Gait, M. J. (1984). *Oligonucleotide synthesis*. IRL Press, Oxford.

Gallie, D. R. and Kado, C. I. (1989). A translational enhancer derived from tobacco mosaic virus is functionally equivalent to a Shine-Dalgarno sequence. *Proceedings of the National Academy of Sciences, U.S.A.* 86: 129-132.

Gallie, D. R., Sleat, D.E., Watts, J. W., Turner, P. C. and Wilson, T. M. A. (1987). The 5'-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts *in vitro* and *in vivo*. *Nucleic Acids Research* 15: 3257-3272.

Gallie, D. R., Sleat, D.E., Watts, J. W., Turner, P. C. and Wilson, T. M. A. (1988). Mutational analysis of the tobacco mosaic virus 5'-leader for altered ability to enhance translation. *Nucleic Acids Research* 16: 883-893.

Ganoza, M. C., Sullivan, P., Cunningham, C., Hader, P., Kofoid, E. and Neilson, T. (1982). Effect of bases contiguous to AUG on translation initiation. *Journal Biological Chemistry* 257: 8228-8232.

Gheysen, D., Iserentant, D., Derom, C. and Fiers, W. (1982). Systematic alteration of the nucleotide sequence preceding the translation initiation codon and the effects on bacterial expression of the cloned SV40 small-t antigen gene. *Gene* 17: 55-63.

Gigova, L., Wishart, P., Ushera, A., Ivanova, M., Bardarov, S., Jay, E. and Ivanov, I. (1989). Expression of repetitive human calcitonin genes in *Escherichia coli*. *Biotechnology and Applied Biochemistry* 11: 401-412.

Gold, L. (1988). Posttranscriptional regulatory mechanisms in *Escherichia coli*. *Annual Review of Biochemistry* 57: 199-233.

Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B. S. and Stormo, G. (1981). Translational initiation in prokaryotes. *Annual Review of Microbiology* 35: 365-403.

Grunstein, M. and Hogness, D. S. (1975). Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proceedings of the National Academy of Sciences, U.S.A.* 72: 3961-3965.

Guarente, L., Lauer, G., Roberts, T. M. and Ptashne, M (1980). Improved methods for maximizing expression of a cloned gene: a bacterium that synthesizes rabbit ß-globin. *Cell* 20: 543-553.

Hall, M. N., Gabay, J., Debarbouille, M. and Schwartz, M. (1982). A role for mRNA secondary structure in the control of translational initiation. *Nature* 295: 616-618.

Heilek, G. M. and Noller, H. F. (1996). Site-directed hydroxyl radical probing of the rRNA neighborhood of ribosomal protein S5. *Science* 272: 1659-1662.

Helke, A., Geisen, R. M., Vollmer, M., Sprengart, M. L. and Fuchs, E. (1993). An unstructured mRNA region and a 5' hairpin represent important elements of the *E.coli* translation initiation signal determined by using the bacteriophage T7 gene 1 translation start site. *Nucleic Acids Research* 21: 5705-5711.

Helser, T.L., Davies, J. E. and Dahlberg, J. E. (1971). Change in methylation of 16S ribosomal RNA associated with mutation to kasugamycin resistance in *Escherichia coli*. *Nature New Biology* 233: 12-14.

Hui, A. and de Boer, H. A. (1987). Specialized ribosome system: Preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in *Escherichia coli*. *Proceedings of the National Academy of Sciences, U.S.A.* 84: 4762-4766.

Hui, A., Hayflick, J., Dinkelspiel, K. and de Boer, H. A. (1984). Mutagenesis of the three bases preceding the start codon of the ß-galactosidase mRNA and its effect on translation in *Escherichia coli*. *EMBO Journal* 3: 623-629.

Iserentant, D. and Fiers, W. (1980). Secondary structure of mRNA and efficiency of translation initiation. *Gene* 9: 1-12.

Ivanov, I. G., Alexandrova, R., Dragulev, B. and AbouHaidar, M. G. (1995). A second putative mRNA binding site on the *Escherichia coli* ribosome. *Gene* 160: 75-79.

Ivanov, I. G., Alexandrova, R., Dragulev, B., Leclerc, D., Saraffova, A., Maximova, V. and AbouHaidar, M. G. (1992). Efficiency of the 5'-terminal sequence (Ω) of tobacco mosaic virus RNA for the initiation of eukaryotic gene translation in *Escherichia coli*. *European Journal of Biochemistry* 209: 151-156.

Ivanov, I. and Gigova, L. (1986). RNA colony hybridization method. *Gene* 46: 287-290.

Ivanov, I., Markova, N., Bachvarov, D., Alexcier, K., Sarattova, A., Maximova, V., Tsaneva, I. and Markov, G. (1989). Constitutive expression of a native human interferon-alpha 1 gene in *E.coli*. *International Journal of Biochemistry* 21: 983-985.

Jacob, W.F., Santer, M. and Dahlberg, A.E. (1987). A single base change in the Shine-Dalgarno region of 16S rRNA of *Escherichia coli* affects translation of many proteins. *Proceedings of the National Academy of Sciences, U.S.A.* 84: 4757-4761.

Jay, E., MacKnight, D., Lutze-Wallace, C., Harrison, D., Wishart, P., Lin, W.-Y., Asundi, V., Pomeroy-Cloney, L., Rommens, J., Eglington, L., Pawlar, J. and Jay, F. (1984a). Chemical synthesis of a biologically active gene for human immune interferon-gamma. *Journal of Biological Chemistry* 259: 6311-6317.

Jay, E., Rommens, J., Pomeroy-Cloney, L., MacKnight, D., Lutze-Wallace, C., Wishart, P., Harrison, D., Lin, W.-Y., Asundi, V., Dawood, M. and Jay, F. (1984b). High-level expression of a chemically synthesized gene for human interferon gamma using a prokaryotic expressive vector. *Proceedings of the National Academy of Sciences, U.S.A.* 81: 2290-2294.

Jay, E., Seth, A. K. and Jay, G. (1980). Specific binding of a chemically synthesized prokaryotic ribosome recognition site. *Journal of Biological Chemistry* 255: 3809-3812.

Jay, G., Khonry, G., Seth, A. K. and Jay, E. (1981). Construction of a general vector for efficient expression of mammalian proteins in bacteria: use a synthetic ribosome binding site. *Proceedings of the National Academy of Sciences, U.S.A.* 78: 5543-5548.

Johnston, H. M. and Roth J. R. (1981). DNA sequence changes of mutations altering attenuation control of the histidine operon of *Salmonella typhimurium*. *Journal of Molecular Biology* 145: 735-756.

Klock, G. and Hillen, W. (1986). Expression, purification and operator binding of the transposon Tn1721-encoded Tet repressor. *Journal of Molecular Biology* 189: 633-641.

Knight, J. A., Hardy, L. W., Rennell, D., Herrick, D. and Poteete, A. R., (1987). Mutations in an upstream regulatory sequence that increase expression of the bacteriophage T4 lysozyme gene. *Journal of Bacteriology* 169: 4630-4636.

Kozak, M. (1983). Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiology Review* 47: 1-45.

Lodish, H. F. (1969). Species specificity of polypeptide chain initiation. *Nature* 224: 867-870.

Loechel, S., Inamine, J. M. and Hu, P. (1991). A novel translation initiation region from *Mycoplasma genitalium* that functions in *Escherichia coli*. *Nucleic Acids Research* 19: 6905-6911.

Malhotra, A. and Harvey, S. C. (1994). A quantitative model of the *E.coli* 16S mRNA in 30S ribosomal subunit. *Journal of Molecular Biology* 270: 308-340.

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982). *Molecular cloning: a laboratory manual*. Cold Spring Harbour Laboratory, Cold Spring Harbour, N.Y.

McCarthy, J. E. G. and Brimacombe, R. (1994). Prokaryotic translation: the interactive pathway leading to initiation. *Trends in Genetics* 10: 402-407.

McCarthy, J.E.G. and Gualerzi, C. (1990). Translational control of prokaryotic gene expression. *Trends in Genetics* 6: 78-85.

McPheeters, D. S., Christensen, A., Young, E. T., Stormo, G. and Gold, L. (1986). Translational regulation of expression of the bacteriophage T4 lysozyme gene. *Nucleic Acids Research* 14: 5813-5817.

Moazed, D. and Noller, H.F. (1990). Binding of rRNA to the ribosomal A and P sites protects two distinct sets of nucleotides in 16S rRNA. *Journal of Molecular Biology* 211: 135-145.

Nierhaus, K.H., R. Adlung, T.P. Hausner, S. Schilling-Bartetzko, T. Twardowski, and F. Triana. (1993). The Allosteric Three-Site Model and the Mechanism of Action of both Elongation Factors EF-Tu and EF-G. in *The Translational Apparatus*. Nierhause, K. H., Franceschi, F., Subramanian, A. R., Erdmann, V. A. and Wittmann-Liebold, B. eds. New York: Plenum Press, pp. 263-272.

Noller, H. F., Green, R., Heilek, G., Hoffarth, V., Huttenhofer, A. Joseph, S., Lee, I., Lieberman, K., Mankin, A., Merryman, C., Powers, T., Puglisi, E. V., Samaha, R. R. and Weiser, B. (1995). Structure and function of ribosomal RNA. *Biochemistry and Cell Biology* 73: 997-1009.

Oakes, M. I. and Lake, J. A. (1990). DNA-hybridization electron microscopy: localization of five regions of 16S rRNA on the surface of 30S ribosomal subunits. *Journal of Molecular Biology* 211: 897-906.

Ofengand, J., Bakin, A. and Nurse, K. (1993). The functional role of conserved sequences of 16S ribosomal RNA in protein synthesis. in *The translational apparatus: structure, function, regulation, evolution*. Nierhaus, K. H., Franceschi, F., Subramanian, A. R., Erdmann, V. A. and Wittmann-Liebold, B. eds. Plenum Press, New York. pp. 489-500.

Olins, P. O. and Rangwala, S. H. (1989). A novel sequence element derived from bacteriophage T7 mRNA acts as an enhancer of translation of the *lacZ* gene in *Escherichia coli*. *Journal of Biological Chemistry* 264: 16973-16976.

Osbourn, J. K., Plaskitt, K. A., Watts, J. W. and Wilson, T. M. A. (1989). Tobacco mosaic virus coat protein and reporter gene transcripts containing the TMV origin-of-assembly sequence do not interact in double-transgenic tobacco plants: implications for coat protein-mediated protection. *Molecular Plant-Microbe Interaction* 2: 340-345.

Petersen, G. B., Stockwell, P. A. and Hill, D. F. (1988). Messenger RNA recognition in *Escherichia coli*: a possible second site of interaction with 16S ribosomal RNA. *EMBO Journal* 17: 3957-3962.

Pirrotta, V. (1979). Operators and promoters in the O_R region of phage 434. *Nucleic Acids Research* 6: 1495-1508.

Pohlner, J., Kramer, J. and Meyer, T.F. (1993). A plasmid system for high-level expression and *in vivo* processing of recombinant proteins. *Gene* 130: 121-126.

Ptashne, M., Backman, K., Humayun, M. Z., Jeffrey, A., Maurer, R., Meyer, B. and Sauer, R. T. (1976). Autoregulation and function of a repressor in bacteriophage lambda. *Science* 194: 156-161.

Ringquist, S., MacDonald, M., Gibson, T. and Gold, L. (1993). Nature of the ribosomal mRNA track: Analysis of the ribosome-binding sites containing different sequences and secondary structures. *Biochemistry* 32: 10254-10262.

Rommens, J., MacKnight, D., Pomeroy-Cloney, L. and Jay, E. (1983). Gene expression: Chemical synthesis and molecular cloning of a bacteriophage T5 (P25) early promoter. *Nucleic Acids Research* 11: 5921-5940.

Saito, H. and Richardson, C. (1981). Processing of mRNA by ribonuclease III regulates expression of gene 1.2 of bacteriophage T7. *Cell* 27: 533-542.

Scherer, G. F. E., Walkinshaw, M. D., Arnott, S. and Morre, D. J. (1980). The ribosome binding sites recognized by *E.coli* ribosomes have regions with signal character in both the leader and protein coding segments. *Nucleic Acids Research* 8: 3895-3907.

Schmitt, M., Kyriatsoulis, A. and Gassen, H. G. (1982). The context theory as applied to the decoding of the initiator tRNA by *Escherichia coli* ribosomes. *European Journal of Biochemistry* 125: 389-394.

Schoelz, J. E. and Zaittin, M. (1989). Tobacco mosaic virus RNA entres chloroplasts *in vivo*. *Proceedings of the National Academy of Sciences, U.S.A.* 86: 4496-4500.

Schwartz, M., Roa, M. and Debarbouille, M. (1981). Mutations that affect *lamB* gene expression at a posttranscriptional level. *Proceedings of the National Academy of Sciences, U.S.A.* 78: 2937-2941.

Shatsky, I. N., Bakin, A. V., Bogdanov, A. A. and Vasiliev, V. D. (1991). How does the mRNA pass through the ribosome? *Biochimie* 73: 937-945.

Shean, C. S. and Gottesman, M. E. (1992). Translation of the prophage lambda CI transcript. *Cell* 70: 513-522.

Shine, J. and Dalgarno, L. (1974). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences, U.S.A.* 71: 1342-1346.

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

Simons, R. W. and Kleckner, N. (1983). Translational control of IS10 transposition. *Cell* 34: 683-691.

Singer, M., and Berg, P. (1991). *Genes and Genomes*. Mill Valley, California: University Science Books, pp. 129-222.

Singer, B. S., Gold, L., Shinedling, S. T., Colkitt, M., Hunter, L. R., Pribnow, D. and Nelson, M. A. (1981). Analysis *in Vivo* of translational mutants of the rIIB cistron of bacteriophage T4. *Journal of Molecular Biology* 149: 405-432.

Sinha, N. D., Biernet, J., McManus, J. and Koster, H. (1984). Polymer support oligonucleotide synthesis XVIII: use of ß-cyanoethyl-N,N-dialkylamino-/N-morpholino phosphoramidite of deoxynucleosides for the synthesis of DNA fragments simplifying deprotection and isolation of the final product. *Nucleic Acids Research* 12: 4539-4557.

Sleat, D. E., Gallie, D. R., Jefferson, R. A., Bevan, M. W., Turner, P. C. and Wilson, T. M. A. (1987). Characterization of the 5'-leader sequence of tobacco mosaic virus RNA as a general enhancer of translation *in vitro*. *Gene* 66: 217-225.

Sleat, D. E., Hull, R., Turner, P. C. and Wilson, T. M. A. (1988). Studies on the mechanism of translational enhancement by the 5'-leader sequence of tobacco mosaic virus RNA. *European Journal of Biochemistry* 175: 75-86.

Sprengart, M. L., Fatscher, H. P. and Fuchs, E. (1990). The initiation of translation in *E.coli*: apparent base pairing between the 16srRNA and downstream sequences of the mRNA. *Nucleic Acids Research* 18: 1719-1723.

Sprengart, M. L., Fuchs, E. and Porter, A. G. (1996). The downstream box: and efficient and independent translation initiation signal in *Escherichia coli*. *EMBO Journal* 15: 665-674.

Steitz, J. A. (1969). Polypeptide chain initiation: nucleotide sequences of the three ribosomal binding sites in bacteriophage R17 RNA. *Nature* 224: 957-964.

Steitz, J.A. (1979). Genetic signals and nucleotide sequences in messenger RNA. in *Biological regulation and development*. Goldberger, R.F. ed. Plenum Publishing Corp., New York.

Steitz, J. A. and Jakes, K. (1975). How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. *Proceedings of the National Academy of Sciences, U.S.A.* 72: 4734-4738.

Steitz, J. A. and Steege, D. A. (1977). Characterization of two mRNA.rRNA complexes implicated in the initiation of protein biosynthesis. *Journal of Molecular Biology* 114: 545-558.

Stern, S., Weiser, B. and Noller, H. F. (1988). Model for the three-dimensional folding to 16S ribosomal RNA. *Journal of Molecular Biology* 204: 447-481.

Stormo, G.D. (1986). Translation Initiation. in *Maximizing Gene Expression*. Reznikoff, W. and Gold, L. eds. Boston: Butterworth Publisher, pp. 196-224.

Stroynowski, I., van Cleemput, M. and Yanofsky, C. (1982). Superattenuation in the tryptophan operon of *Serratia marcescens*. *Nature* 298: 38-41.

Tabor, S. and Richardson, C. C. (1987). DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proceedings of the National Academy of Sciences, U.S.A.* 84: 4767-4771.

Taniguchi, T. and Weissmann, C. (1979). *Escherichia coli* ribosomes bind to non-initiator sites of QB RNA in the absence of formylmethionyl-tRNA. *Journal of Molecular Biology* 128: 481-500.

Taniguchi, T. and Weissmann, C. (1978). Inhibition of QB RNA 70S ribosome initiation complex formation by an oligonucleotide complementary to the 3' terminal region of *E.coli* 16S ribosomal RNA. *Nature* 275: 770-772.

Tate, W.P., Adamski, F.M., Brown, C.M., Dalphin, M.E., Gray, J.P., Horsefield, J.A., McCaughan, K.K., Moffat, J.G., Powell, R.J., Timms, K.M. and Trotman, C.N.A. (1993). Translationa Stop Signals: Evolution, Decoding for Protein Synthesis and Recording for Alternative Events. in *The Translational Apparatus*. Nierhause, K.H., Franceschi, F., Subramanian, A.R., Erdmann, V.A. and Wittmann-Liebold, B. New York: Plenum Press, pp. 253-262.

Thummel, C. S., Burgess, T. L. and Tjian, R. (1981). Properties of simian virus 40 small t antigen over produced in bacteria. *Journal of Virology* 37: 683-697.

van Dieijen, G., van Knippenberg, P. H. and van Duin, J. (1976). The specific role of ribosomal protein S1 in the recognition of native phage RNA. *European Journal of Biochemistry* 64: 511-518.

van Duin, J., Kurland, C. G., Dondon, J. and Grunberg-Manago, M. (1975). Near neighbours of IF3 bound to 30S ribosomal subunits. *FEBS Letters* 59: 287-290.

van Knippenberg, P. H. (1990). Aspects of Translation Initiation in *Escherichia coli*.in *The ribosome: structure, function and evolution*. Hill, W. E., Moore, P. B., Dahlberg, A., Schlessinger, D., Garrett, R. A. and Warner, J. R. eds., American Society for Microbiology, Washington, DC, pp. 265-274.

van Wezenbeek, P. M. G. F., Hulsebos, T. J. M. and Schoenmakers, J. G. G. (1980). Nucleotide sequence of the filamentous bacteriophage M13 DNA genome: comparison with phage fd. *Gene* 11: 129-148.

Verschoor, A., Srivastava, S., Radermacher, M., Frank, J., Traut, R. R., Stoffler-Meilicke, M. and Glitz, D. (1993). Functional site determinations in three dimensions on eukaryotic and eubacterial ribosomes. in *The translational apparatus: structure, function, regulation, evolution*. Nierhaus, K. H., Franceschi, F., Subramanian, A. R., Erdmann, V. A. and Wittmann-Liebold, B. eds. Plenum Press, New York. pp. 411-419.

Watson, J.D., Hopkins, N.H., Roberts, J.W., Steitz, J.A. and Weiner, A.M. (1987). *Molecular Biology of Gene*. Menlo Park, California: The Benjamin/Cummings Publishing Company, pp. 382-430.

APPENDIX
LIST OF SYNTHETIC OLIGONUCLEOTIDES

CYMV18
5′-ATGCGAATCTTTGCTGGT

EcoF
5′-ATAGGCGTATCACGAGG

L36B
5′-AGCTTGTAATTGTAAAAGAGGGTGATATTGC

PMV18
5′-CGTTAATCTAGGGTTGTG

CYMVP
5′-TCGAGAAAACAAAACGAAAACAAACAAAATCTTCGTAAACACCCTCCAACAC
CTTTTCTTTTGCTTTTGTTTGTTTTAGAAGCATTTGTGGGAGGTTGTG

ACCATTCAATCCCGAATCGCCCGCAAACCAGCAAAGATTCGCATA
TGGTAAGTTAGGGCTTAGCGGGCGTTTGGTCGTTTCTAAGCGTATTCGA

L36
5′-TCGAGCAATATCACCCTCTTTTACAATTACA
CGTTATAGTGGGAGAAAATGTTAATGTTCGA

PMVP
5′-TCGAGAAAAGAAACACAAAGCAAAGCAAAGCAAAGCAACTCAAATAAACCAT
CTTTTCTTTGTGTTTCGTTTCGTTTCGTTTCGTTGAGTTTATTTGGTA

ATTTGGCCAAGGCACTTGGTAATCAAACGGGCACAACCCTAGATTAACGA
TAAACCGGTTCCGTGAACCATTAGTTTGCCCGTGTTGGGATCTAATTGCTTCGA

PP1′
5′-**AATT**GAAAAATTTATTTGCTTTCAGGAAAATTTTT
CTTTTTAAATAAACGAAAGTCCTTTTAAAAA

CTGTATAATAGATCTCGAGAAGGAGGTTTA
GACATATTATCTAGAGCTCTTCCTCCAAATTCGA

∈I
5′-**TCGA**TTAACTTTATTT
AATTGAAATAAATCGA

∈II
5′-**TCGA**TTAACTTTATTTACCTTATC
AATTGAAATAAATGGAATAG**TCGA**

∈SD
5′-**TCGA**TTAACTTTATCAAGGAGGTTTC
AATTGAAATAGTTCCTCCAAAG**TCGA**

Ω42 CONSTRUCT
5′-**TCGA**AGTCGAGTTGCTGACTTTACAATTACT
TCAGCTCAACGACTGAAATGTTAATGA**TCGA**

Ω43 CONSTRUCT
5′-**TCGA**CGCATTCTGACCTTTTACAATTACT
GCGTAAGACTGGAAAATGTTAATGA**TCGA**

NOTE: RESTRICTION SITES ARE UNDERLINED AND THE MODIFIED RESTRICTION OVERHANGS ARE IN BOLD LETTERS.