Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

A study of the regulation of aminoacyl tRNA synthetase gene expression in *Bacillus* by T-box regulatory elements

A Thesis Submitted to the University of Dublin for the Degree of Doctor of Philosophy

By

Niall Foy

Department of Genetics
University of Dublin
Trinity College
Dublin 2

October 2009

DECLARATION

This thesis has not previously been submitted to this or any other university for examination for a higher degree. The work presented here is entirely my own except otherwise acknowledged. This thesis may be lent, copied and made available for consultation within the university library upon request. It may be lent to other libraries for purposes of consultation.

ACKNOWLEDGEMENTS

First of all I'd like to thank my supervisor, Kevin Devine, for the support, advice, encouragement and motivation to keep going during my time in his lab. Thanks to Brian, Efi, Eric, Sebastian, Dave, Paola, Inga, Annette, Clodagh, Letal, Leagh and Alistair, for the banter, their help and putting up with my singing.

I'd like to thank Mam and Dad for looking after my mental health and being there for me when I really needed them. You guys are the best. Thanks as well to Sorcha, Ciaran and Gavin for all your support.

Thanks to Diarmuid and Aine, you guys got me through a really tough time. I can't thank you enough. Thanks to Suzanne for always being positive and Graham for the craic and helping me distract myself by lifing heavy things.

Thanks to all the lads, Pop, Gar, Scoop, Gonzo, Anitha, Melissa, El Gringo, Jonny and Magali. I'm lucky to have such great friends. Thanks to Kate for all the encouragement and listening to my ceaseless thesis based complaints.

Thanks to Gavin Conant for his help and expertise with the bioinformatic analysis.

If you're reading this and feel slighted that I haven't mentioned you, I suggest you buy me a pint so we can talk about it.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Figure 5.39 Putative stem loop structure present in the *gltX* promoter region

Summary

The T-box antitermination mechanism is commonly used for the regulation of aminoacyl tRNA synthetase gene expression in Gram positive bacteria. However, expression of lysyl tRNA synthetases is rarely controlled in this way. *Bacillus cereus* was found to contain a class I *lysK* gene which has a putative T-box regulatory element upstream (Ataide, *et al.*, 2005). In this work we characterized the T-box regulatory element of *lysK* from *B. cereus*. Our results showed that it is functional and responsive to reduced charging of $tRNA^{LYS}$. Our analysis of 891 sequenced bacterial genomes had shown that all 6 strains containing T-box regulated *lysS/K* gene also encoded another *lysS/K* gene in the genome. Work in *B. cereus* had previously shown that the T-box regulated *lysK* gene is expressed predominantly in stationary phase and does not charge $tRNA^{LYS}$ during exponential growth (Ataide, *et al.*, 2005). This led us to investigate if a strain of *B. subtilis,* containing a single *lysS/K* gene, whose expression is controlled by a T-box mechanism was viable. Our results showed that T-box regulation of a single *lysS/K* gene in *B. subtilis* is compatible with growth. As T-box regulation of lysyl tRNA synthetase genes was shown to occur rarely we investigated why this might be the case. We observed that lysine (AAA, AAG) and asparagine (AAC, AAU) are encoded by codons within a mixed codon box. This led us to suggest that the relative paucity of T-box regulation for *lysS/K* genes may be due to an inability to distinguish between reduced charging of $tRNA^{ASN}$ and $tRNA^{LYS}$ as they differ only at the wobble base of the anticodon. This proposed lack of specificity was confirmed by our results which showed that expression of a $P_{T\text{-}box\ lysK}$-*lacZ* fusion was induced by reduced charging of $tRNA^{ASN}$. We then investigated if, in *B. subtilis*, a feature of T-box regulation is a lack of specificity due to an inability to discriminate between tRNAs that are charged with amino acids from mixed codon boxes. For example, phenylalanine and leucine are partners in a mixed codon box. Therefore, we analysed the effect of leucine limitation on the expression of a $P_{T\text{-box}\ pheS}$-*gfp* fusion. Our analysis showed that induction of expression of T-box regulated AARS genes is specific to reduced charging of their cognate tRNA, indicating that the lack of specificity observed with the *lysK* T-box regulatory element is a special case. The analysis of T-box regulatory elements in *B. subtilis* did reveal however, that the T-box regulatory elements of *leuS* and *cysES* are not responsive to reduced charging of their cognate tRNA under the conditions tested and also showed that limitation for certain non-cognate amino acids can induce expression of aminoacyl tRNA synthetase genes by an unknown indirect mechanism.

ABBREVIATIONS

AARS	Aminoacyl tRNA synthetase
AMP	Adenosine monophosphate
Ap^R	Ampicillin resistant
ATP	Adenosine triphosphate
BCA	Bicinchoninic acid protein assay
BLAST	Basic Local Alignment Search Tool
BLM	Basal Limitation Medium
bp	Base pairs
Cm^R	Chloramphenicol resistant
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
DNAse	Deoxyribonuclease
DTT	Dithiothreitol
dNTP	Deoxynucleoside triphosphate
EDTA	Ethylenediaminetetra acetic acid
Em^R	Erythromycin resistant
GFP	Green Fluorescent Protein
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan^R	Kanamycin resistant
Kb	Kilobase pairs
LB	Luria Bertani
MCS	Multiple Cloning Site
mRNA	Messenger ribonucleic acid
OD	Optical density
ONPG	o-Nitrophenyl β-D-galactopyranoside
PCR	Polymerase Chain reaction
$Phleo^R$	Phleomycin resistant
RBS	Ribosome binding site
RNA	Ribonucleic acid
RNAse	Ribonuclease
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
$Spec^R$	Spectinomycin resistant

SSC	Salt sodium citrate
tRNA	Transfer ribonucleic acid
UTR	Untranslated region
UV	Ultra-violet

Chapter 1

Introduction

1.1 Introduction

Translation is the process by which mRNA is decoded by the ribosome to produce protein. The accuracy of this process is crucial to prevent the misincorporation of amino acids into elongating protein chains resulting in the production of nonsense proteins. In 1958, Crick put forward the "adaptor hypothesis", which suggested that an intermediary or "adaptor" nucleic acid would be required to link the genetic code to protein production (Crick, 1958). This hypothesis was confirmed in the same year by the discovery of tRNA by Hoagland *et al* (Hoagland, *et al.*, 1958). In order to carry out its function, the tRNA molecule must first be aminoacylated or charged with its cognate amino acid. The charging reaction is carried out by a group of enzymes known as the aminoacyl tRNA synthetases (AARSs). The accuracy of AARS function is essential for protein synthesis.

1.2 The aminoacyl tRNA synthetases

The aminoacyl tRNA synthetases are a family of enzymes whose cellular function is to attach amino acids to their cognate tRNA. The resultant aminoacyl-tRNA can then be used in translation (Grunberg-Manago, 1996; Woese, *et al.*, 2000; Ibba and Söll 2001). The reaction catalyzed by the AARS enzymes is a two-step process (Figure 1.1). The first step involves the formation of an activated aminoacyl adenylate complex, through cleavage of the terminal diphosphate of an ATP molecule and subsequent attachment of AMP to the amino acid (Hoagland, 1955). In the second step the amino acid is esterified to the terminal adenosine present at the 3' end of all tRNA molecules to form aminoacyl-tRNA (Hoagland, *et al.*, 1958). Following aminoacylation the charged tRNA can then function in the translation of mRNA on the ribosome. The accuracy of the tRNA charging process is essential to prevent misincorporation of incorrect amino acids into elongating protein chains on the ribosome (Loftfield, 1963). The AARS family of enzymes are divided into two separate groups termed class I and class II (Eriani, *et al.*, 1990). While both classes perform similar functions, they are in fact two entirely separate groups of enzymes sharing no structural similarities, having undergone convergent evolution to carry out the same task (Martinis and Schimmel; 2003; O Donaghue and Luthey-Schulten 2003). The division of the various synthetases into either class I or class II is conserved in all living things (Table 1.1).

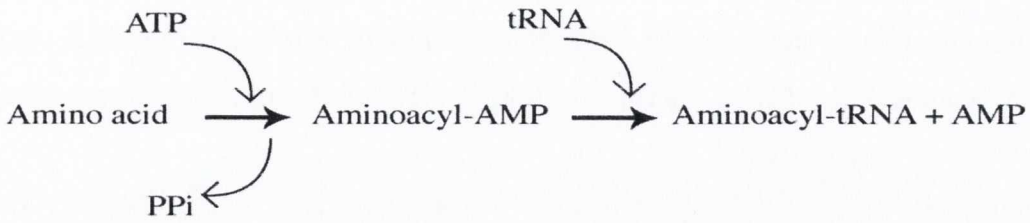


Figure 1.1 Diagram of the catalytic reaction carried out by the AARSs. The cognate amino acid and ATP react to form the enzyme bound aminoacyl adenylate. The aminoacyl adenylate is then transferred to the acceptor end of tRNA with the loss of AMP.

This so called "class rule" is broken only in the case of the lysyl tRNA synthetase which has both a class I and class II variant (the proteins are referred to as LysKRS and LysRS respectively in this study, the genes as *lysK* and *lysS*) (Jester, *et al.*, 2003; Ataide, *et al.*, 2005). The distribution of AARS enzymes into either class I or class II is based on their differing active site architectures. (Eriani, *et al.*, 1990 Martinis and Schimmel, 2003; O Donoghue and Luthey-Schulten, 2003).

Table 1.1

Class I aminoacyl tRNA synthetases	Class II aminoacyl tRNA synthetases
MetRS	AlaRS
ValRS	HisRS
IleRS	ProRS
LeuRS	ThrRZ
CysRS	SerRS
GluRS	GlyQS
GlnRS	PheST
ArgRS	AspRS
TrpRS	AsnRS
TyrRS	LysRS
LysKRS	

The active site of the class I enzyme is characterized by a dinucleotide binding domain known as a Rossman fold which contains an inner core of five anti-parallel β-sheets (O Donoghue and Luthey-Schulten, 2003; Martinis and Schimmel, 2003; Levengood, *et al.*, 2004). Within this active site are two highly conserved motifs known as the HIGH and KMSKS motifs which occupy opposite sides of the Rossman fold (Martinis and

Schimmel, 2003). The active site of the class II AARS enzyme is composed of seven anti-parallel β-sheets flanked by α helices (Martinis and Schimmel, 2003; Levengood, *et al.*, 2004; Schimmel, 2008). Within this active site there are a series of loosely conserved motifs designated motifs 1,2 and 3. These motifs vary in length from class II enzyme to class II enzyme and contain as little as one invariant residue (Eriani, *et al.*, 1990; Martinis and Schimmel, 2003) (Table 1.2).

Table 1.2 Motifs of the active site of class II AARSs

Motif 1	gΦxxΦxxPΦΦ
Motif 2	(F/Y/H)Rx(E/D)(4-12x)(R/H)xxxFxxx(D/E)
Motif 3	λxΦgΦgΦeRΦΦΦΦΦ

Φ=hydrophobic
x=variable residue
λ=small amino acid
Lowercase letters indicates amino acids that are partially conserved in the sequence

While there is little amino acid sequence conservation in motifs 1,2 and 3, significant structural conservation exists (O Donoghue and Luthey-Schulten, 2003). Similar structural conservation is seen at the HIGH motif of class I enzymes but not at the KMSKS motif (O Donoghue and Luthey-Schulten, 2003). A further indication of the different structures of the class I and class II enzymes is seen in their interactions with tRNA. The two classes of enzyme approach tRNA from different sides. Class I enzymes approach the tRNA from the minor groove side and aminoacylate the 2' OH of the terminal adenosine ribose moiety while the class II enzymes approach from the major groove side and aminoacylate the 3' OH of the terminal adenosine (Eriani, *et al.*, 1990; Levengood, *et al.*, 2004). The different active site architecture also results in the two classes of AARS binding ATP in different ways. The class I enzymes bind ATP in an extended conformation whereas the class II AARS bind ATP in a bent conformation (Levengood, *et al.*, 2004).

1.3 Editing of mis-acylated tRNA

Selection of the correct tRNA for charging by its cognate synthetase is a very precise process. It is facilitated by the unique structural features present on different tRNA molecules which allows the AARS enzyme to discriminate between them and unerringly select the correct tRNA for charging (Jakubowski & Goldman, 1992; Ibba & Soll, 2000). The selection of the correct amino acid for charging is less accurate. As many of the amino acids are quite similar in size and structure e.g. valine, leucine and

isoleucine, AARS enzymes can have difficulty in distinguishing between them and can often misactivate amino acids forming incorrect aminoacyl adenylates, which must be removed by a process referred to as editing. The editing process was first discovered for IleRS in *E. coli* (Baldwin & Berg, 1966). It was found that IleRS could produce either isoleucyl or valyl-adenylate in the presence of isoleucine and valine. However while the isoleucyl-adenylate became esterified to tRNA, the valyl-adenylate was hydrolyzed in the presence of $tRNA^{ILE}$. This result showed the existence of an editing mechanism to prevent charging of incorrect amino acids onto tRNA. Further research into this area has revealed that there are 4 main pathways through which editing is achieved. These are (i) the dissociation of an enzyme bound aminoacyl adenylate to produce free aminoacyl adenylate which is hydrolyzed in solution (ii) tRNA independent deacylation of an enzyme bound aminoacyl adenylate (iii) tRNA dependant hydrolysis of an enzyme bound aminoacyl adenylate without transient mischarging of tRNA and (iv) deacylation of an enzyme bound misacylated tRNA (Jakubowski & Goldman, 1992). Each AARS can employ different editing pathways for the removal of different amino acids. For example, IleRS edits cysteine by dissociation of the enzyme bound cysteinyl adenylate which is then hydrolyzed in solution, while valine is edited by tRNA-dependant hydrolysis of the enzyme bound valyl adenylate (Jakubowski & Fersht, 1981; Baldwin & Berg, 1966; Jakubowski & Goldman, 1992). There are some AARS enzymes for which no editing mechanism is necessary due to the specificity of their interaction with their cognate amino acid. Examples of such enzymes are CysRS and TyrRS (Jakubowski & Goldman, 1992).

1.4 Alternative mechanisms for tRNA charging

As there are twenty essential amino acids it is reasonable to assume the existence of twenty separate AARS enzymes in every living organism. However, the development of genome sequencing has demonstrated that this is only the case for the eukarya and some bacteria. The majority of bacteria, all known archaea and all eukaryotic organelles lack the AARS specific for glutamine (GlnRS) (Woese, *et al.*, 2000; Salazar, *et al.*, 2003). In addition, many archaea and bacteria lack a canonical asparaginyl tRNA synthetase whereas some methanogenic archaebacteria lack a canonical cysteinyl tRNA synthetase (Woese, *et al.*, 2000; Salazar, *et al.*, 2003; O Donoghue, *et al.*, 2005).

In organisms lacking either GlnRS ar AsnRS an indirect pathway to Gln-$tRNA^{GLN}$ and Asn-$tRNA^{ASN}$ formation is employed to ensure accurate translation of the genetic

code. This pathway employs a non-discriminating glutamyl or aspartyl tRNA synthetase which mischarges $tRNA^{GLN}$ or $tRNA^{ASN}$ with glutamate and aspartate respectively. These mischarged amino acids are then transamidated by the products of the *gatCAB* genes to form Gln-$tRNA^{GLN}$ and Asn-$tRNA^{ASN}$ (Woese, *et al.*, 2000; Salazar, *et al.*, 2003). An analysis of eukaryotic GlnRS and AsnRS enzymes indicates that they evolved through gene duplication of *gluS* and *aspS* genes which coded for non-discriminatong GluS and AspS enzymes (O Donoghue and Luthey-Schulten, 2003).

In *Methanococcus jannaschi* there is no class I CysRS enzyme. In order to charge $tRNA^{CYS}$ with cysteine the tRNA is first aminoacylated with O-phosphoserine (cysteine precursor) by the class II O-phosphoserine synthetase . This non-cognate Sep-$tRNA^{CYS}$ is then converted to cys-$tRNA^{CYS}$ by Sep-tRNA:Cys-tRNA synthase (O Donoghue, *et al.*, 2005).

1.5 The tRNA

The tRNAs are the molecules responsible for linking the genetic code with protein production. In general, tRNAs are 76 nucleotides long and end in a single stranded NCAA tetranucleotide to which amino acids are attached (Martinis and Schimmel, 2003). They are traditionally schematically represented as cloverleaf shaped molecules (Figure 1.2). This best illustrates their important structural features, namely, the TΨC (T pseudouridine C) loop, the D (dihydrouridine) loop and the anticodon loop. These loops are in turn attached to short helical regions. Also critical to the structure of tRNA are the variable loop and acceptor stem to which the amino acid is esterified (Martinis and Schimmel, 2003). The acceptor stem and TΨC stem and loop combine along with the D loop and anticodon loop to form an L-shaped tertiary structure as seen in figure (1.3). This tertiary structure is stabilized by highly specific base-pairing interactions between the different helical domains of the molecule (Martinis and Schimmel, 2003). tRNAs decode the genome by binding of their anticodon to codons on the mRNA in the ribosome. Generally the identity of the amino acid is specified by the first two nucleotides of the codon. The third base, or wobble base can vary between any of the four nucleotides depending on the amino acid. The existence of this variability at the third position means that frequently, non Watson-Crick base pairing interactions must

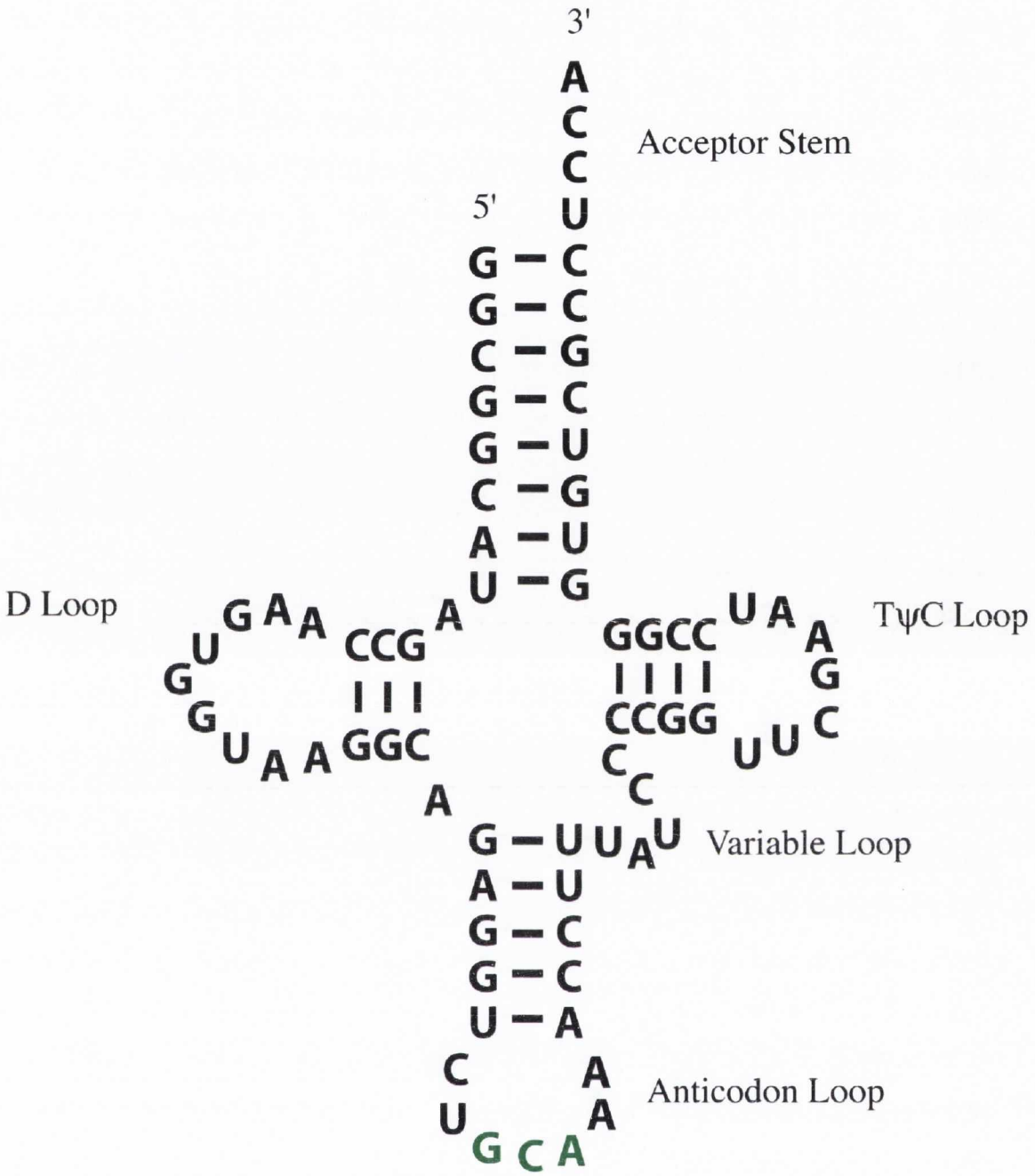


Figure 1.2. Secondary structure of *B. subtilis* $tRNA^{CYS}$. The important strucutural features are indicated. The anticodon is indicated in green.

occur to allow accurate decoding of mRNA. These non-canonical interactions are facilitated by the existence of tRNA modifications, which alter the structure of anticodons so they can accommodate non-cognate base pairs at the wobble position on the ribosome (Agris, *et al.*, 2007). These modifications along with the discriminator base (base 73 of tRNA) often function as specificity determinants for AARSs in the charging reaction (Agris, *et al.*, 2007). They have also been implicated in prevention of tRNA decay (Alexandrov, *et al.*, 2006). An even greater variability of RNA-RNA interactions, including pyrimidine-pyrimidine binding, can occur outside the ribosome (Agris, *et al.*, 2007). For example U-U pairings are found in structural RNAs. However such interactions are generally precluded on the ribosome through hydrogen bonding interactions with ribosomal proteins in the acceptor site. In a ribosomal context, the modified wobble hypothesis suggests that in addition to cognate nucleotide-nucleotide interactions, modified uracil bases can wobble to G and in the case of 5-oxyuridine modifications also to U and possibly C (Agris, *et al.*, 2007).

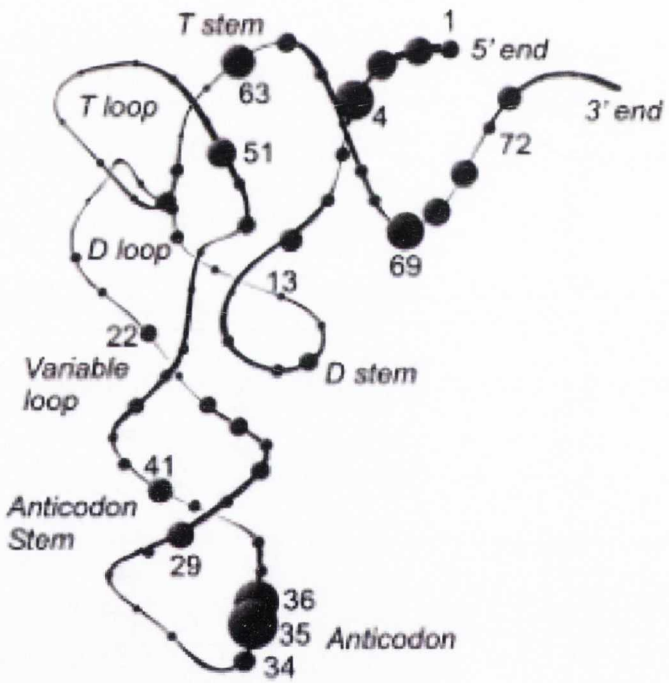


Figure 1.3 Tertiary structure of tRNA. Circles represent nucleotides involved in recognition by AARSs. The size of the circles is proportional to the documented frequency with which they are recognized by AARSs. This diagram was taken from (Ibba and Soll, 2000).

1.6 Regulation of expression of AARS genes

As essential components of the cellular translational machinery, production of the AARSs must be strictly controlled. Research into this area has revealed a variety of different mechanisms by which this control is achieved. A significant proportion of this research has been undertaken in *E. coli* and *B. subtilis* and has uncovered a number of examples of unique regulatory mechanisms.

The study of AARS regulation in *E. coli* has revealed that while a variety of regulatory mechanisms exist to coordinate the synthesis of individual AARSs in response to starvation for their own cognate amino acid, there is also a primary degree of regulation for most of the tRNA synthetases, known as metabolic regulation (Grunberg-Manago, 1996). This metabolic regulation results in a 2-3 fold increase in AARS production for a 5 fold increase in growth rate (Grunberg-Manago, 1996).

Early work on the regulation of expression of AARS genes showed that for 10 of the synthetases there was an increase in expression as a result of starvation for their cognate amino acid (Nass and Neidhardt, 1967). Since then it has become clear that while a similar derepression effect is observed for different synthetases, the mechanisms by which the increase in AARS production is achieved are quite different. Regulation of AARS expression has been shown to occur at the level of transcription initiation, translational repression and transcription anti-termination. Below are examples of these types of regulation

In *E. coli*, PheST synthesis is regulated by a mechanism linking translational attenuation to transcription anti-termination (Mayaux, *et al.*, 1985; Grunberg-Manago, 1996; Ryckelynck, *et al.*, 2005). The upstream untranslated leader region (UTR) of the *pheST* genes contains a number of hairpin secondary structures. One of these is a Rho-independent terminator that can refold into an alternative anti-terminator structure. Also contained in the UTR are a number of phenylalanine codons, three of which are consecutive. Under conditions of phenylalanine starvation, uncharged $tRNA^{PHE}$ enters the ribosome acceptor site causing the ribosome to stall at the phenylalanine codons. This promotes the formation of the anti-terminator structure allowing transcriptional read-through into the *pheST* genes. When there is an adequate supply of phenylalanine in the cell, the ribosome can translate the run of phenylalanine codons and the terminator structure is favoured, preventing read-through of transcription (Figure 1.4).

Regulation of alanyl tRNA synthetase is *E. coli* provides a good example of direct feedback inhibition of transcription initiation. AlaRS controls the level of transcription of its own gene by directly binding a pair of palindromic sequences that overlap the

transcription initiation site (Putney and Schimmel, 1981; Grunberg-Manago, 1996). Thus, high levels of AlaRS in the cell prevent transcription of the *alaRS* genes.

Regulation of *thrS* expression in *E. coli* is by a feedback inhibition mechanism based on molecular mimicry (Moine, et al., 1988; Ryckelynck, *et al.*, 2005). Control in this case is exerted at a translational level (Lestienne, *et al.*, 1984). The UTR of the *thrS* gene can fold into four well defined domains. The first domain contains the Shine-Dalgarno sequence and the *thrS* start codon. Domains 2 and 4 are stem-loop structures which resemble the anticodon stem-loop and the acceptor stem of $tRNA^{THR}$ respectively (Springer, *et al.*, 1985; Moine, *et al.*, 1990). Domains 2 and 4 are linked by domain 3 which is unstructured (Figure 1.5). The ribosome recognizes domains 1 and 3 whereas ThrRS recognizes and binds to domains 2 and 4. This leads to competition between ThrS and the 30S ribosomal subunit for *thrS* mRNA binding (Moine, *et al.*, 1990). High levels of ThrRS binding of its own mRNA prevent translation. The level of uncharged $tRNA^{THR}$ also plays a role in this regulation mechanism as it acts as a competitor for ThrRS binding (Romby, et al., 1992; Romby, *et al.*, 1996). The role of domain two of the *thrS* mRNA as a molecular mimic of the $tRNA^{THR}$ anticodon stem loop was confirmed by experiments which mutated the threonine anticodon of domain two to a methionine anticodon, thus succeeding in switching control of *thrS* mRNA translation to MetRS binding of the mRNA (Graffe, *et al.*, 1992).

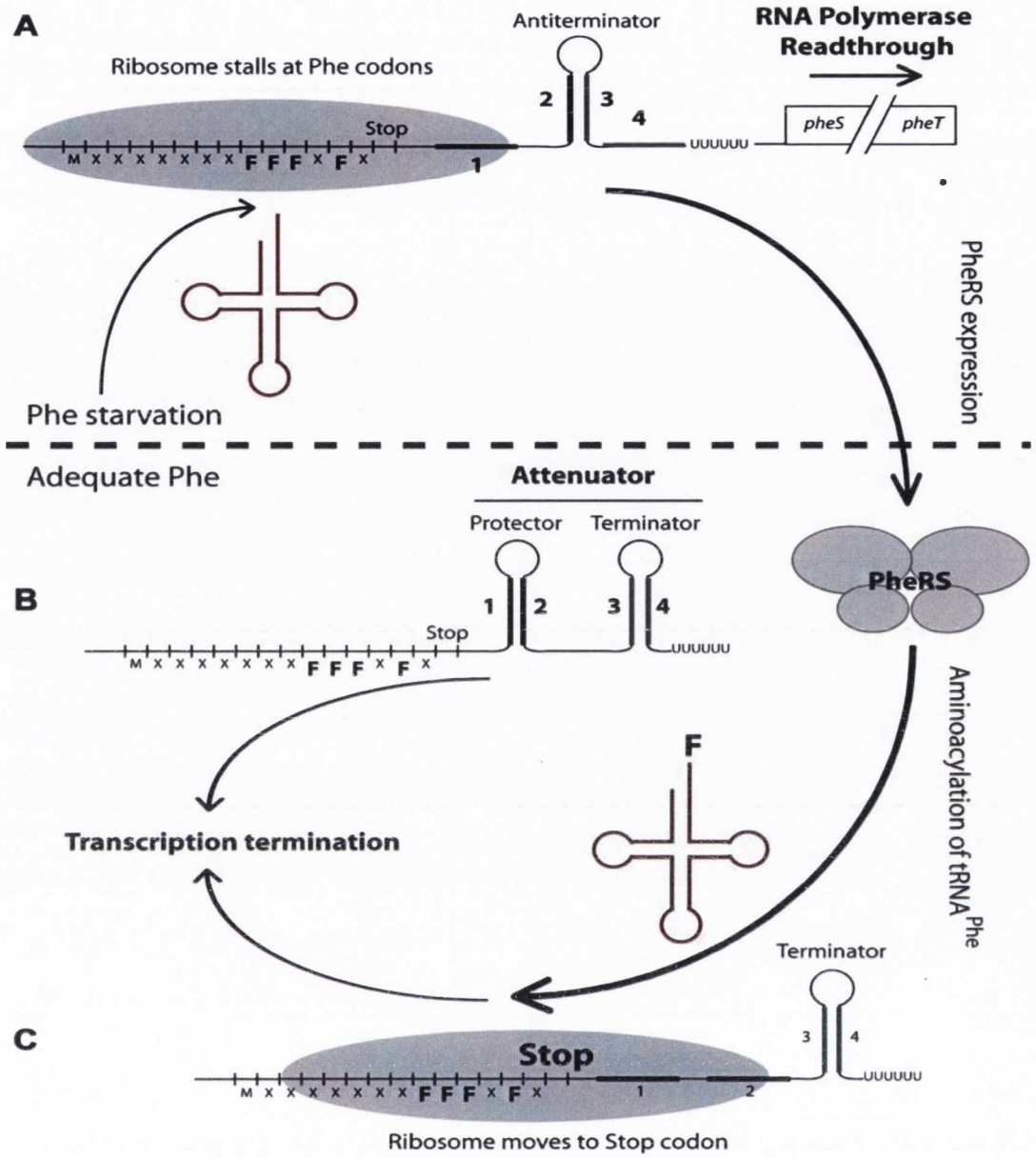


Figure 1.4 Transcriptional attenuation mechanism of the *pheST* genes of *E.coli*. Upstream of the pheST genes are two stem-loop strucutures and a number of phenylalanine codons. In conditions of phenylalanine starvation, the ribosome stalls at the phenylalanine codons promoting formation of the anti-terminator structure and transcriptional read-through. The resultant increase in PheRS expression increases the level of $tRNA^{PHE}$ charging. This prevents ribosome stalling at the run of phenylalanine codons in the leader region favouring the terminator structure and termination of transcription. This figure is taken from Ryckelynck, *et al*, (2005).

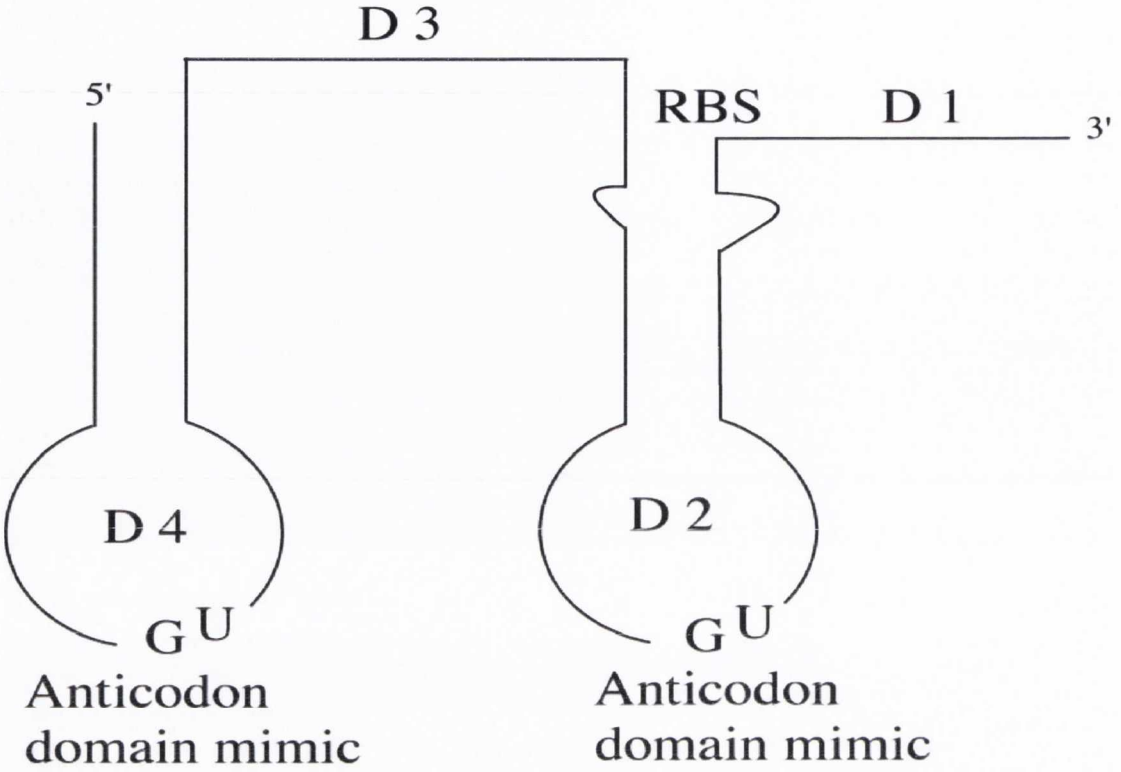


Figure 1.5 Feedback attenuation of translation mechanism of *thrS* in *E. coli*. The *thrS* leader region contains four well defined domains, two of which can fold into stem-loop structures which resemble the anticodon stem and loop of $tRNA^{THR}$. Domain one contains the Shine-Dalgarno sequence for ribosome binding. The ribosome recognizes and binds domains 1 and 3 whereas ThrRS recognizes domains 2 and 4. This overlap in binding sites leads to competition between the ribosome and ThrRS thus regulating the translation of *thrS* mRNA [Springer, *et al.*, (1985), Grunberg-Manago, (1996)]. This image is adapted from Ryckelynck, *et al*, (2005).

1.7 Riboswitches and their role in genetic regulation

The term riboswitch was coined by Ronald Breaker in 2002 to describe an RNA regulatory element that can control regulation of gene expression without the aid of a protein cofactor (Winkler, *et al.*, 2002). These ribswitches have been shown to regulate gene expression by mechanisms of transcription antitermination, attenuation, translation initiation and mRNA splicing (Blouin, *et al.*, 2009). As RNA lacks the range of functional groups available to proteins to interact with different cellular metabolites, they must form complex tertiary structures to form ligand binding sites [Muhlbacher, *et al.*, (2007), Blouin, *et al.*, (2009)]. These complex secondary and tertiary RNA structures are found in the untranslated leader regions of a number of different genes involved in varying biosynthetic pathways in both gram positive and gram negative bacterial species. Examples of biosynthetic pathways regulated by riboswitches are those for adenine, guanine, glucosamine-6-phosphate, lysine, S-adenosylmethionine (SAM) and AARSs [Muhlbacher, *et al.*, (2007), Blouin, *et al.*, (2009), Grundy and Henkin, (1998), Grundy, et al., (2003)].

Many of these riboswitches regulate biosynthetic pathways that were first described in *B. subtilis*. It has been shown that approximately 4% of all genes in B. subtilis are regulated by riboswitch mechanisms (Garst, *et al.*, 2008). Two examples are the L-box regulon controlling lysine biosynthesis and the S-box controlling synthesis of methionine and SAM.

1.8 T-boxes and AARS regulation in *B. subtilis*

It is clear from the above examples that regulation of AARS gene expression in *E. coli* occurs by a variety of different modes. However, in *B. subtilis*, one type of regulation appears to control the expression of the majority of AARS genes. This method of regulation is known as the T-box anti-termination mechanism.

The T-box anti-termination mechanism is one of the most common regulatory elements found in Gram positive bacteria. It was first identified in *B. subtilis*, upstream of the genes encoding threonyl and tyrosyl tRNA synthetase (Putzer, *et al.*, 1992; Henkin, *et al.*, 1992; Grundy and Henkin, 1993; Henkin, 1994). Since its initial discovery in *Bacillus* species, the T-box regulatory element has been found in bacterial phyla including the *Firmicutes*, the *Chloroflexi*, the *Deinococcus-Thermus*, the *δ-proteobacteria* and the *Actinobacteria*, located upstream of genes involved in tRNA charging, amino acid biosynthesis, amino acid transport and a number of regulatory proteins involved in amino acid metabolism (Gutierrez-Preciado, *et al.*, 2009).

The term T-box regulatory element refers to a 200-300 nucleotide untranslated RNA leader region containing a number of stem-loop structures, conserved motifs, an intrinsic Rho-independent terminator and the T-box [Putzer, *et al.*, (1992), Henkin, *et al.*, (1992), Grundy and Henkin, (1993)]. The T-box is a 14bp nucleotide region that is highly conserved across all T-box regulatory elements (Figure 1.6).

Figure 1.6 The T-box motif. The relative conservation of nucleotides is indicated by size. N represents the variable nucleotide of the T-box motif which co-varies with the discriminator base of the T-box regulatory element's cognate tRNA. This figure is adapted from Wels, *et al*, (2008).

The majority of the studies on the T-box regulatory element have been carried out in *B. subtilis*. Analysis of the *B. subtilis* genome shows the presence of the T-box upstream of 14 of the 24 genes encoding AARSs, indicating that these genes are regulated by a T-box antitermination mechanism (Pelchat and Lapointe, 1999). The T-box mechanism functions in the regulation of AARS genes by sensing the level of charging of its cognate tRNA (Grundy, *et al.*, 1994). This is achieved by the interaction of the tRNA with the T-box regulatory element. This interaction induces re-folding of the terminator structure to form the less stable anti-terminator, which allows read-through of transcription into the downstream genes. The T-box regulatory element functions as follows: When the levels of tRNA charging are high, the tRNA can interact with the T-box regulatory element through the anticodon-specifier codon interaction. However, the acceptor end of the tRNA is occluded by the presence of the tRNAs cognate amino acid. Therefore the tRNA cannot interact with the T-box region to stabilize the formation of the anti-terminator. The terminator structure is preferred and transcription is halted at the terminator. In conditions when the level of tRNA charging is low, the acceptor end is not blocked by the presence of an amino acid. The acceptor end of the tRNA can then interact with the T-box region and stabilize anti-terminator formation. This allows transcription to proceed into the downstream genes (Figure 1.7).

There are two main points of contact of the tRNA with the leader mRNA. These are via the tRNAs anticodon which interacts with the specifier codon of the T-box regulatory

element and the tRNAs acceptor stem which interacts with the T-box loop (Grundy, *et al.*, 1994).

1.8.1 The specifier codon

The specifier codon is the primary recognition determinant for the cognate tRNA of the T-box regulatory element. It is complementary to the anticodon of the cognate tRNA. Recognition of the cognate tRNA of any given T-box regulatory element is therefore expected to occur by a simple codon-anticodon interaction. The specifier codon is located in a single stranded loop present on the 3' side of Stem I of the T-box regulatory element, the size of which varies in different T–box regulatory elements (Grundy and Henkin, 1993; Grundy, *et al.*, 1994) (Figure 1.8). It has been observed that the specifier codon is always separated from the stem I helical domain by one nucleotide. This nucleotide tends to be an A or a G allowing the potential for base-pairing with the universally conserved U34 residue 5' to the anticodon in tRNA. The one nucleotide separation of the specifier codon from the helical domain may also be important for correct presentation of the specifer codon for tRNA binding [Henkin, (1994). Yousef, *et al.*, (2005)]. A comparative analysis of T-box regulatory elements from a number of sequenced genomes has also indicated a bias for the presence of a C in the third position of the specifier codon. This preference is not influenced by codon usage but may be selected based on those tRNA isoacceptors most sensitive to the shortage of their cognate amino acid (Grundy and Henkin, 1994; Wels, *et al.*, 2008; Elf, *et al.*, 2003). Work carried out in *E. coli* has shown that where there are two tRNA isoacceptors, with different ratios of total tRNA concentration to codon frequency, the isoacceptor with the lowest ratio will totally lose its charging under amino acid starvation conditions whereas the isoacceptor with the higher amount of total tRNA relative to codon frequency will retain a level of charging which in some cases can approach 100%. This is because the level of supply of uncharged tRNA to satisfy the demand imposed by codon frequency is satisfied at a lower level of charging for the isoacceptor with the lower ratio of total tRNA to codon frequency. Accordingly, a codon with a lower ratio of total tRNA concentration to codon frequency is more sensitive to amino acid starvation and should be selected for use in a regulatory system that must respond to amino acid depletion (Elf, *et al.*, 2003). This may indicate that in *B. subtilis*, the bias for a C at the third position of the specifier codon is a result of the fact that the tRNAs most sensitive to starvation for their cognate amino acid have a G at the 5' position of the anticodon.

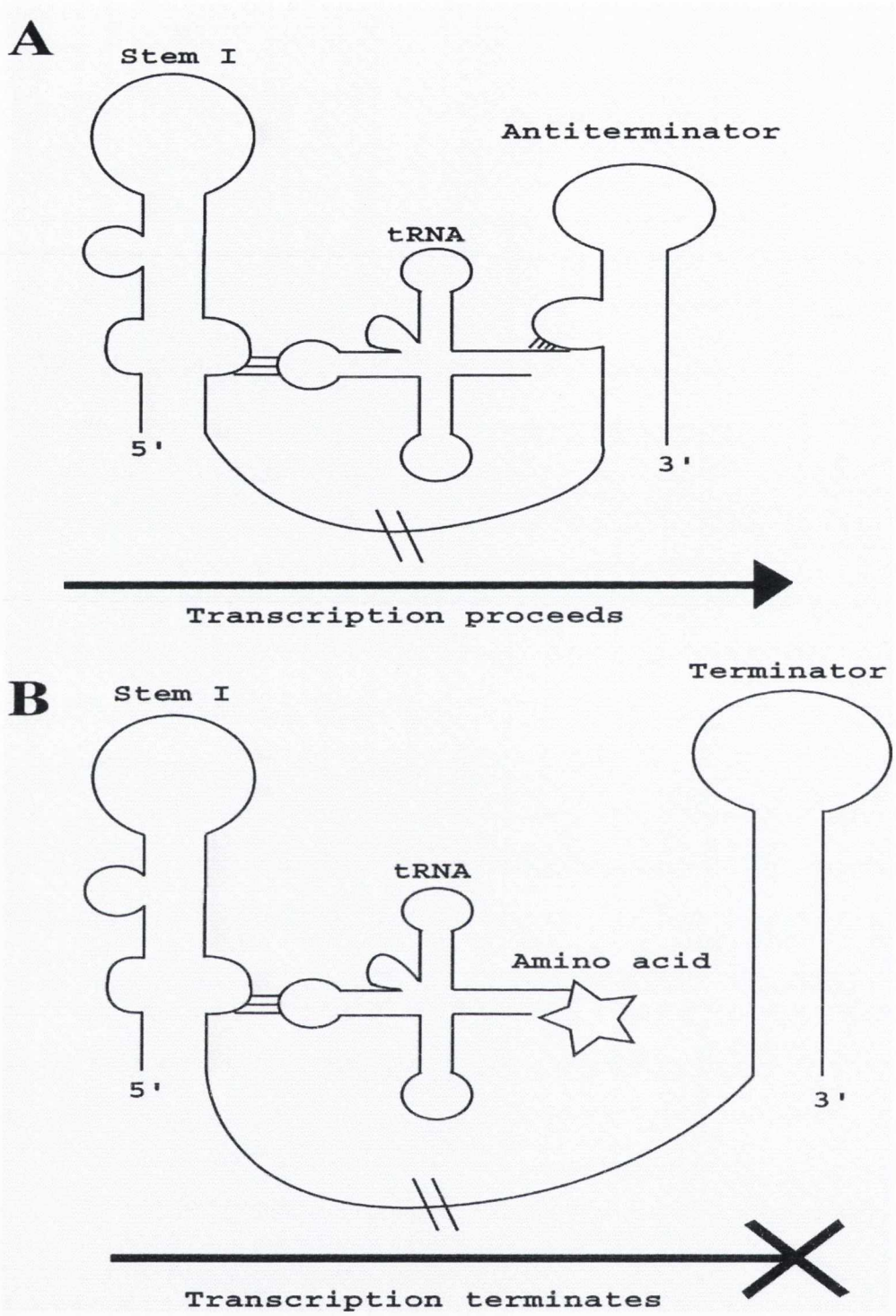


Figure 1.7 Schematic representation of the interaction of both charged and uncharged tRNA with a T-box regulatory element. (A) Interaction of uncharged tRNA with a T-box regulatory element. The tRNA interacts through basepairing of its anticodon with a complementary specifier codon present in stem I and through basepairing of its acceptor end with a complementary sequence that forms part of the T-box sequence present in a loop on the 5' side of the antiterminator. This interaction of uncharged tRNA with the T-box regulatory element allows transcription to proceed. (B) Interaction of charged tRNA with a T-box regulatory element. The tRNA interacts through basepairing of its anticodon with a complementary specifier codon present in stem I. However, the presence of an amino acid on the acceptor end of the tRNA prevents basepairing with the T–box sequence. The thermodynamically more stable terminator structure is preferred and transcription is terminated.

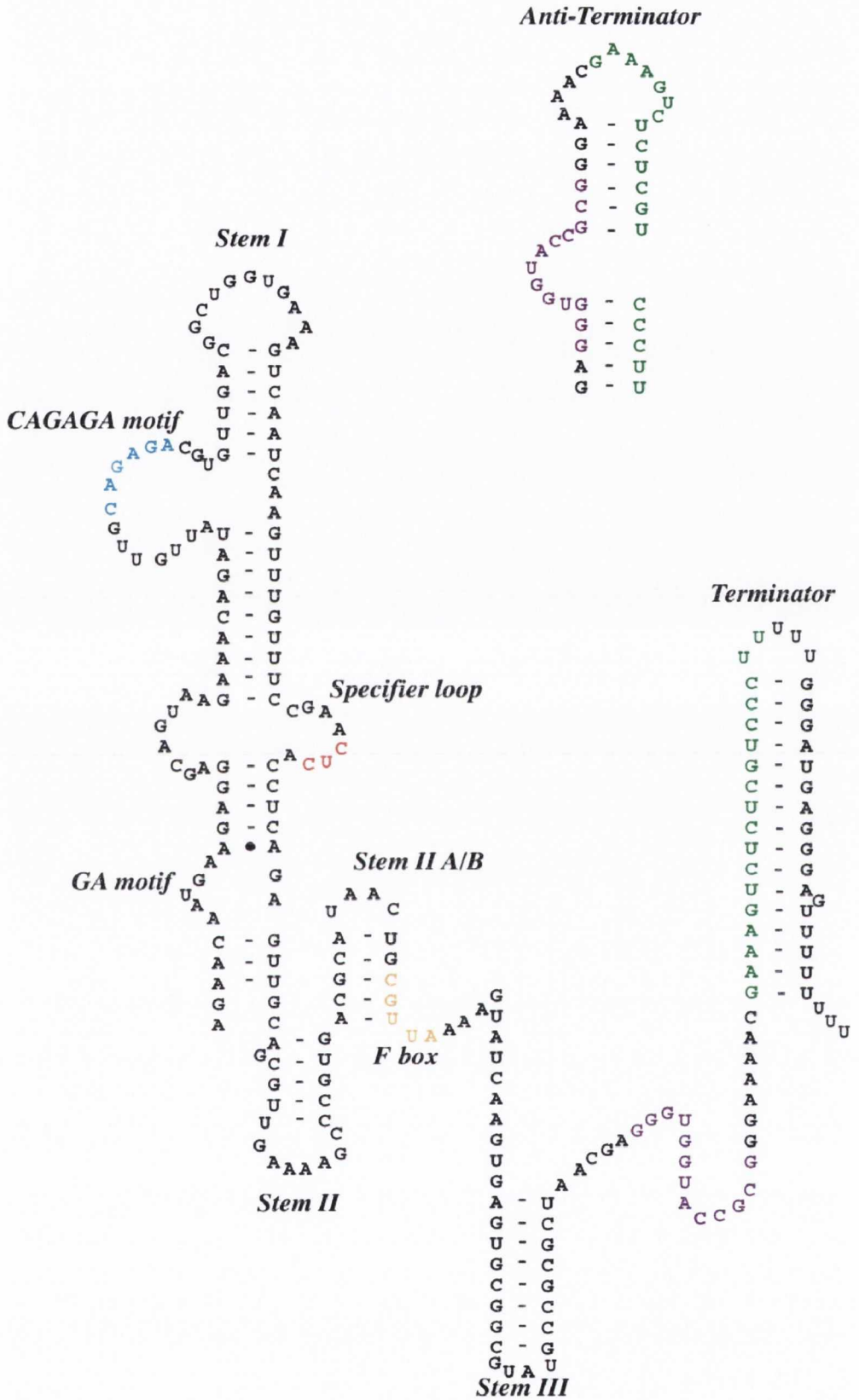


Figure 1.8 Proposed secondary structure of the upstream leader region of the leucyl tRNA synthetase gene of *B. subtilis*. The T-box sequence is highlighted in purple. Sequences that can form part of either the terminator or antiterminator are highlighted in green. The specifier codon is highlighted in red. Other important structural features are indicated

Alternatively, it has been proposed that the high frequency of a C at the third position of the anticodon is due to the increased stability that a G-C interaction, relative to other base-pairing possibilities, contributes to the anticodon-specifier codon interaction (Henkin, 1994).

The importance of a cognate specifier codon-anticodon interaction in tRNA recognition in T-box regulatory elements has been shown by a number of mutational and biochemical analyses. Fluorescence quenching experiments carried out on the *B. subtilis glyQS* stem I specifier loop *in vitro* have shown that quenching of fluorescence from the GGC specifier codon was only achieved when a reconstituted anticodon stem and loop (ASL) bearing a glycine CCG anticodon was used. An ASL carrying a phenylalanine anticodon did not quench fluorescence showing that the non-cognate phenylalanine ASL could not interact with the *glyQS* specifier loop (Nelson, *et al.*, 2006). Mutational analyses of T-box regulatory elements *in vivo* have also demonstrated the importance of the specifier codon. Work on the *tyrS* T-box regulatory element of *B. subtilis* showed that mutation of the specifier codon from a tyrosine AAC codon to a UUC phenylalanine codon switched the specificity of the T–box regulatory element from tyrosine to phenylalanine albeit with a lowered induction of transcription relative to the levels seen in the wild-type T-box system (Grundy and Henkin, 1993). Mutations of this nature do not always result in a switch in tRNA specificity. A similar experiment in which the specifier codon of the *thrS* T-box regulatory element of *B. subtilis* was switched from ACC to an AAC tyrosine codon or UUC phenylalanine codon did not result in a switch of tRNA specificity (Henkin, 1994; Putzer, *et al.*, 1995; Grundy, *et al.*, 1997). Attempts to switch the specificity of the *tyrS* T-box regulatory element to arginine by mutation of the specifier codon to a CGC arginine codon were also unsuccessful. These results show that while the interaction of the tRNA anticodon with the specifier codon is necessary for antitermination, it is not sufficient. Other features of the tRNA or T-box regulatory element are important for efficient antitermination and that the relative importance of these varies from system to system. These features will be addressed in a later section.

1.8.2 The antiterminator T-box loop

The specifier codon-tRNA anticodon interaction is necessary to select for the correct tRNA for interaction with the T-box regulatory element. A second point of interaction is necessary to achieve antitermination. This second point of interaction occurs between the tRNA acceptor end and the T-box loop present on the 5' side of the antiterminator structure (Grundy, *et al.*, 1994). The antiterminator is comprised of two helical regions

separated by a 7 nucleotide single stranded loop region (Figure 1.7). The first 4 bases of the single stranded loop region are made up of the invariant 5' UGGN 3' of the T-box motif. These four nucleotides bind to the conserved 5' NCCA 3' unpaired nucleotides present at the end of virtually every tRNA (Grundy, *et al.*, 1994). The N variable nucleotide of the T-box motif co-varies with the discriminator base of the T-boxe's cognate tRNA. Studies carried out on the discriminator base-T-box interaction have shown it to be vital for antitermination to occur (Grundy, *et al.*, 1994; Gerdeman, *et al.*, 2002). Work performed on the *tyrS* leader region of *B. subtilis in vivo* showed that mutation of the variable base in the T-box loop region from a U to an A results in a 10 fold decrease in expression as a result of starvation for charged $tRNA^{TYR}$, which has an A at the discriminator base, due to the A-A mismatch between the tRNA discriminator base and the variable base of the T-box. The importance of this interaction for successful antitermination was demonstrated further by a gel shift analysis carried out using a *tyrS* antiterminator model RNA. Studies of the ability of $tRNA^{TYR}$ and $tRNA^{TYR}$ acceptor stem models to bind to the antiterminator loop showed that binding of the discriminator base of the tRNA to the variable base of the T-box was required for gel shifts to occur (Gerdeman, *et al.*, 2002). However, the results of these analyses have also shown that some mismatches between the tRNA discriminator base and the variable base of the T-box have less of an effect in preventing antiterminator stabilization. Mutation of the variable base of the *tyrS* T-box to a C allowed induction of antitermination albeit with a three fold reduction from wild-type. This is despite the A-C mismatch between the discriminator base of $tRNA^{TYR}$ and the variable nucleotide of the T-box (Grundy, *et al.*, 1994). These results suggest that there are other aspects of the tRNA acceptor end-antiterminator T-box loop interaction, which are important for antitermination. An investigation of the antiterminator T-box loop and the T-box motif itself have shown that it is the overall tertiary structure of the antiterminator T-box loop that is important for antitermination. Mutational analyses of conserved nucleotides of the T-box motif involved in base-pairing interactions in the helical domains of the antiterminator have confirmed this. Many of the highly conserved bases of the T-box can be switched with only minor effects on antitermination as long as base-pairing is maintained (Grundy, *et al.*, 2002). There were some exceptions to this. The second two basepairs of the helix 3' to the antiterminator T-box loop could not be altered without producing a major negative effect on expression of a *tyrS-lacZ* fusion. In addition to this, the highly conserved 5' ACC 3' nucleotides of the antiterminator T-box loop were shown to be very important in facilitating the T-box-tRNA interaction. All mutations of these three nucleotides resulted in severely reduced expression from *tyrS-lacZ* fusions

indicating that interaction of the antiterminator with uncharged tRNA was severely reduced. Only two mutations of these nucleotides allowed a response of *tyrS-lacZ* fusions to uncharged $tRNA^{TYR}$. These were ACC to AUC and AAC. Both of these sets of mutations resulted in reduced expression but allowed induction of expression in response to reduced $tRNA^{TYR}$ charging (Grundy, *et al.*, 2002). Elucidation of the solution structure of a model *tyrS* antiterminator RNA by Gerdeman *et al,* (2002) showed that the conserved ACC residues exhibit extensive stacking interactions. These stacking interactions produce an 80 degree kink between the two helices of the antiterminator. This kink is important in presentation of the T-box loop for interaction with the tRNA. The UGGN nucleotides of the antiterminator T-box loop which interact with the tRNA acceptor end are unstructured and this results in the UGGN nucleotides having a variety of different conformations for the tRNA to sample. It is believed that the stacking interactions of the ACC nucleotides of the loop are important to provide some structure to the loop and limit the number of available conformations (Gerdeman, *et al.*, 2003). Binding of Mg^{2+} ions have also been shown to be important in the stabilization of antiterminator structure (Jack, *et al.*, 2008). Analysis of the effect of mutations on the terminal loop of the antiterminator showed that this loop region does not appear to have any role in interaction with tRNA (Grundy, *et al.*, 2002).

1.8.3 The importance of tRNA and T-box tertiary structure in antitermination

As mentioned earlier, other interactions between the tRNA and T-box regulatory element, besides the tRNA anticodon-specifier codon and tRNA acceptor end-T-box loop interactions, are important for efficient antitermination. This was initially suggested by experiments carried out to switch the specificity of certain T-box systems to that of another tRNA, which were only successful in some cases. These results indicated that there are other features of the tRNA-T-box regulatory element interaction that are required for antitermination (Grundy, *et al.*, 1994; Putzer, *et al.*, 1995). Genetic and biochemical analyses of the *glyQS* and *tyrS* T-box regulatory elements of *B. subtilis* and the *trpE* T-box regulatory element of *Lactococcus lactis* have shown that the tertiary structure of the tRNA and T-box leader regions in addition to the anticodon-specifier codon and T-box-acceptor end interactions are essential for antitermination (Grundy, *et al.*, 2000; van de Guchte, *et al.*, 2001; Yousef, *et al.*, 2003; Yousef, *et al.*, 2005; Fauzi, *et al.*, 2005).

Analysis of the effect of mutations on the ability of $tRNA^{TYR}$ to induce expression of a *tyrS* leader-*lacZ* fusion *in vivo* showed that mutations in the helical regions of the tRNA which did not affect base-pairing were generally tolerated. However, those mutations that

affected the overall tertiary structure of the tRNA abolished antitermination (Grundy, *et al.*, 2000). Further confirmation of the importance of these interactions was shown by *in vitro* analyses of the ability of $tRNA^{GLY}$ to interact with the *glyQS* leader region. These experiments demonstrated that the tRNA D arm – T arm interaction was necessary for antitermination, consistent with the data previously obtained from the analysis of the *tyrS* leader region.

RNase protection experiments also indicated that there are other sites of interaction between the tRNA and the T-box leader region in addition to the anticodon-specifier codon and the acceptor stem-T-box loop interaction (Yousef, *et al.*, 2005). It was shown using the *glyQS* T–box regulatory element that 4 regions are cleavage protected in the presence of $tRNA^{GLY}$. As expected, the specifier codon and the antiterminator T-box loop were protected. Also protected were the single stranded region between stem I and stem III (note there is no stem II in the *glyQS* leader) and the single stranded region between stem III and the antiterminator (Yousef, *et al.*, 2005).

Investigation of the *trpE* T-box regulatory element from *L. lactis* provides the most convincing evidence to date of interactions of the tRNA with the leader region beyond those that occur at the specifier loop and antiterminator T-box loop. It was observed that there is significant sequence complementarity between the $tRNA^{TRP}$ and the *trpE* leader region. A stretch of 21 nucleotides of the *trpE* leader is complementary to the D-arm and anticodon arm of $tRNA^{TRP}$ while another 18 nucleotides is complementary to the T-arm of $tRNA^{TRP}$. It was shown that $tRNA^{TRP}$ could still induce antitermination in a *trpE* leader-*lacZ* fusion even when the specifier codon had been switched from UGG (trp) to AAC (asn) albeit producing a lower level of expression. This was believed to be due to the interactions of the D-arm and T-arm with the *trpE* leader. This was shown convincingly by production of chimeric tRNAs. $tRNA^{ASN}$ had been shown to induce expression of the *trpE* leader-lacZ fusion, containing an AAC asparagine specifier codon poorly in response to asparagine starvation. Replacement of the D-arm and T-arm of $tRNA^{ASN}$ with those of $tRNA^{TRP}$ greatly increased *lacZ* expression proving that the D-arm and T-arm interactions of $tRNA^{TRP}$ with the *trpE* leader, were vital for antitermination (van de Guchte, *et al.*, 2001). A sequence analysis of other T-box regulatory elements and their cognate tRNAs shows that similar sequence complementarity exists in other T-box regulatory elements though it is not as extensive as that seen in the *L. lactis trpE* system (van de Guchte, *et al.*, 2001). Investigation of the possible role of this complementarity in antitermination in the *B. subtilis glyQS* system showed that mutation of nucleotides in the T-loop and D-loop of $tRNA^{GLY}$ did not have any negative effect on antitermination as would be expected if they

had an important role in the tRNA-leader interaction. A similar result was found in an analysis of $tRNA^{TYR}$ interactions with antiterminator model RNAs (Fauzi, *et al.*, 2005). It was found that there was a correlation between structural elements of tRNA and their ability to bind model antiterminator RNAs that was independent of sequence complementarity. These results show that while all the T-boxes investigated experimentally respond to the level of their cognate uncharged tRNA, there may be differences from system to system in how this interaction is achieved.

1.8.4 Other important structural features of T-box leader regions

In addition to stem I containing the specifier loop, and the terminator and antiterminator structures, there are other conserved features of T-box regulatory elements that have been shown to be important for function. These include the F-box, the CAGAGA motif, the GA motif and stem III of the T-box leader region.

The F-box is made up of the five highly conserved nucleotides 5'CGUUA3'. This motif is located in the helical region of the stem IIA/B pseudoknot (Figure 1.8). Mutational analysis of this motif resulted in reductions in basal and induced expression of a *B. subtilis tyrS* leader-*lacZ* fusion whether or not base-pairing of the stem IIA/B pseudoknot was affected suggesting an important role for this motif in antitermination (Rollins, *et al.*, 1997).

The CAGAGA motif (also known as the AG box) is located in an interior loop of stem I (Figure 1.8). Again mutational analyses of this motif shows that it has an important role in antitermination though its function is not clear (Rollins, *et al.*, 1997).

The best characterized of the three motifs mentioned above is the GA motif. This motif is very common in T-boxes found in the *Bacilli* and *Clostridia* but is rarely seen in T-boxes of the *Staphylococci* and *Streptococci* (Winkler, *et al.*, 2001). The GA motif is located at the base of stem I. It is made up of two short helices separated by an asymmetric internal loop with a 5+2 nucleotide arrangement (Figure 1.9). Helix A is generally 5 basepairs in length while helix B is generally 3 basepairs in length. Between these two helices is the asymmetric internal loop. Within this loop are two oppositely oriented GA dinucleotide sequences which are highly conserved. Mutational analysis of this motif suggests it is functionally important for *tyrS* leader antitermination (Winkler, *et al.*, 2001). Mutation of bases within the helices reduced expression of a *tyr*S leader-*lacZ* fusion under inducing conditions while mutation of the conserved GA nucleotides abolished antitermination. Alteration of the 5+2 nucleotide arrangement by addition of nucleotides to the loop also reduced antitermination.

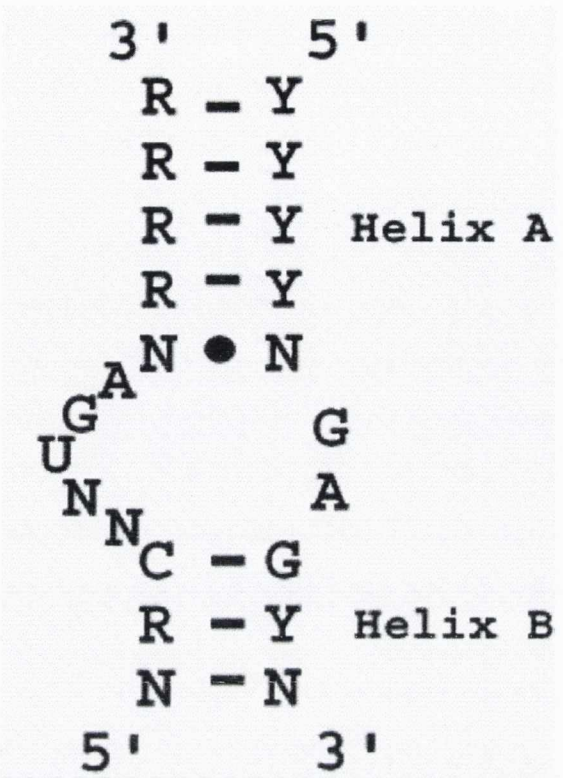


Figure 1.9 The GA motif. R represents purines. Y represents pyrimidines. N represents any base. The weak nucleotide-nucleotide interaction at the bottom of helix A is represented by a dot.

The GA motif shares structural and sequence similarities to other RNA elements found in 16S and 23S rRNA in bacteria. It also displays significant similarities to RNA elements found in humans and yeast that are known to interact with proteins involved in rRNA processing. It is not yet clear if the GA motif found in T-box regulatory elements has a protein binding function (Winkler, *et al.*, 2001). However, there is evidence to suggest that protein cofactors may have a role in stabilization of T-box leader structure in some cases. For example, work on the *thrS* system of *B. subtilis* carried out *in vitro* showed a requirement for the polyamine spermidine, to stabilize the leader RNA secondary structure. Experiments *in vivo*, however showed that spermidine was not necessary for anti-termination, possibly indicating the replacement of spermidine with an as yet unidentified protein factor (Luo, *et al.*, 1998; Putzer, *et al.*, 2002).

1.8.5 The role of transcriptional pausing in antitermination

Investigation of the kinetics of transcription of the *B. subtilis glyQS* T-box leader region *in vitro* identified a number of transcriptional pause sites with a major one at stem III of the

leader region (Grundy and Henkin, 2004). Mutation of the sequence of stem III showed that the sequence itself was not important to trigger transcriptional pausing consistent with a general lack of sequence conservation in this region of T-box regulatory elements (Grundy and Henkin, 2004). It is thought that the role of transcriptional pausing at the stem III loop may be to sense the level of cognate tRNA charging in the cell before commiting to termination or antitermination (Grundy and Henkin, 2004; Grundy, *et al.*, 2005). Pausing at stem III may allow for complete folding of stem I such that the cognate tRNA can interact with the specifier codon. It was noted, that during the *in vitro* transcription assays on the *glyQS* leader that uncharged $tRNA^{GLY}$ could be added at any time prior or during the pause in transcription at stem III and induce antitermination. However, after the transcription elongation complex had left the pause site addition of uncharged tRNA had no effect and termination occurred quickly (Grundy and Henkin, 2004).

1.8.6 The role of RNA processing in antitermination

Another important feature of T-box regulation is the role of RNA processing (Condon, *et al.*, 1996a). Investigation of the *thrS* T-box regulatory element of *B. subtilis* revealed an RNA processing site 9 nucleotides upstream of the transcriptional terminator (Condon, *et al.*, 1996a). It was shown that processing occurred at this site in both inducing and non-inducing conditions but it was induced to a higher level by threonine starvation. Cleavage at this site resulted in an increase in *thrS* mRNA half-life. Under inducing conditions, when the antiterminator is formed, the cleavage site is located in the terminal loop of the antiterminator (Figure 1.10). Cleavage under these conditions and subsequent refolding of the RNA would produce a stem loop structure at the 5' end of the *thrS* transcript. It has been shown in *E. coli* that the presence of stable secondary structures at either end of a transcript increases mRNA stability. The enzymes responsible for cleavage of the *thrS* transcript in *B. subtilis* have recently been identified. They are functional homologues of RNase E from *E. coli* which was shown to cleave *B. subtilis thrS* mRNA transcripts *in vivo* and *in vitro* (Condon, et al., 1997). However, they have no sequence similarity. They have been named RNase J1 and RNase J2 (Even, *et al.*, 2005). Evidence of RNA processing has been seen in other T-box systems suggesting that it may play an important role in the regulation in AARS regulation [Condon, *et al.*, (1996b), Pelchat and Lapointe, (1999)b].

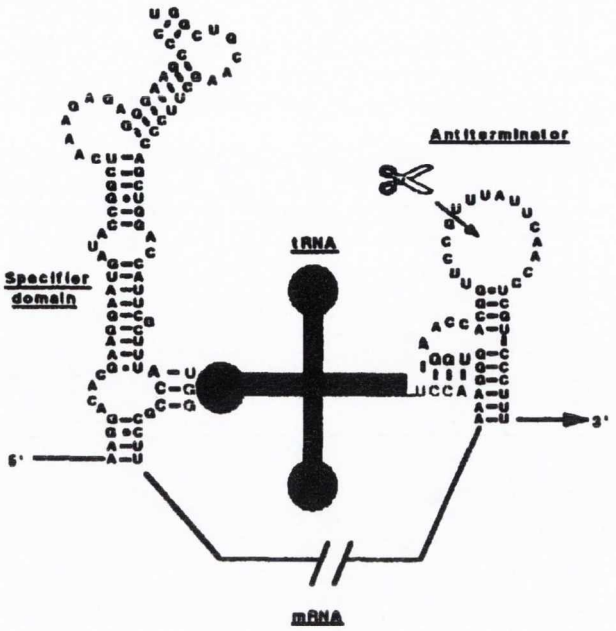
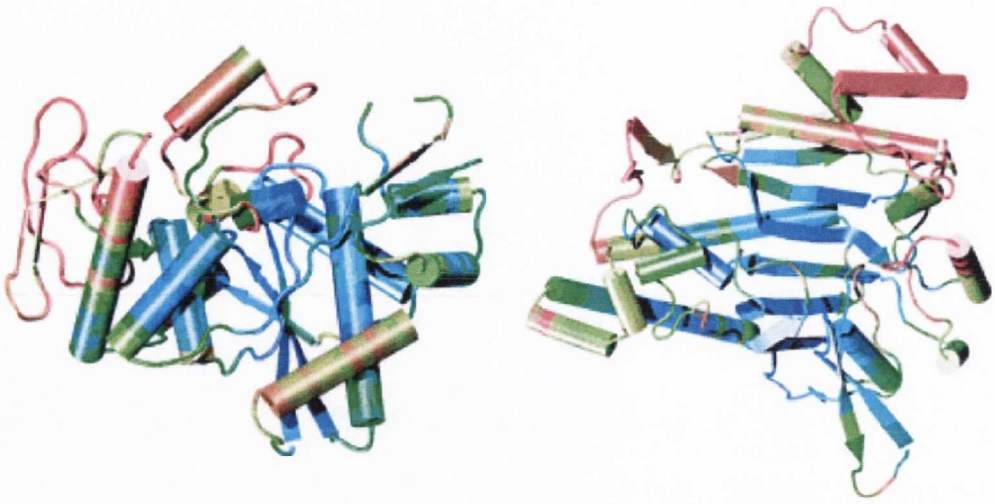


Figure 1.10 Schematic representation of anti-termination of the *thrS* leader region with RNA processing at the anti-terminator. The cleavage site is indicated by the scissors and arrow. This figure is taken from Condon, *et al.*, (1996b).

1.9 The class I and class II lysyl tRNA synthetase

As mentioned earlier there are both class I and class II variants of lysyl tRNA synthetase (LysKRS and LysRS) (Figure 1.11). These two enzymes carry out the same function of $tRNA^{LYS}$ charging but are structurally unrelated. Despite the structural differences between the active sites of LysKRS and LysRS, both enzymes recognize the same regions on the substrate $tRNA^{LYS}$, namely the anticodon, the acceptor stem and the discriminator base (Levengood, *et al.*, 2004). However, they differ in their catalytic approach to lysine activation and subsequent $tRNA^{LYS}$ charging. LysRS initiates $tRNA^{LYS}$ charging using only lysine and ATP to form an enzyme bound aminoacyl adenylate. The class I enzyme however requires tRNA binding to alter the conformation of the active site to allow aminoacyl adenylate synthesis (Levengood, *et al.*, 2004).



Class I Lysyl tRNA Synthetase Class II Lysyl tRNA synthetase

Figure 1.11 Structural models of the class I and class II lysyl tRNA synthetases. Regions highlighted in blue are areas of high conservation. Regions highlighted in red are areas of low conservation. This figure is taken from O'Donoghue and Luthey-Schulten, (2003).

The difference in active site architecture may be a contributing factor in the distribution of the two different classes of LysRS enzyme (Jester, *et al.*, 2003). The class I LysKRS is found in all archaea and some bacteria, while the class II LysRS is found in most bacteria and all eukarya (Jester, *et al*, 2003; Ataide, *et al.*, 2005). The different active site structures of the two enzymes influence their ability to interact with lysine analogues (Jester, *et al.*, 2003; Wang, *et al.*, 2006). The active site of LysKRS is more compact than that of LysRS and it employs both electrostatic and hydrophobic interactions to discriminate against naturally occurring lysine analogues such as S-(2-aminoethyl)-L-cysteine (thialysine). This difference in active site structure makes LysKRS 200 times less sensitive to thialysine than LysRS (Jester, *et al.*, 2003; Wang, *et al.*, 2006). While the larger active site of LysRS results in less specificity towards lysine than LysKRS, it does allow the class II enzyme to catalyze $tRNA^{LYS}$ charging at a faster rate (Wang, *et al.*, 2006). These differences between the two enzymes may have influenced their distribution in the tree of life

While most bacteria contain either LysRS or LysKRS there are an isolated number of cases where both enzymes are present in the same organism. The most well known examples are the *Methanosarcinae* and *Bacillus cereus* (Ataide, *et al.*, 2005).

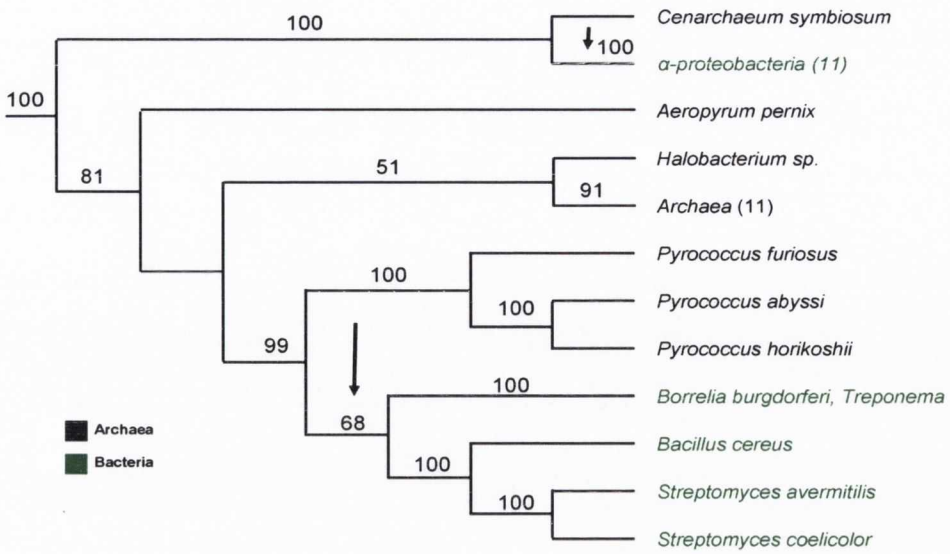


Figure 1.12 LysKRS gene tree. Maximum likelihood phylogenetic tree for 17 archael and 17 bacterial LysKRS sequences. Archael sequences are indicated in black. Bacterial sequences are indicated in green. The tree is rooted using four GluRS sequences from *E. coli, C. jejuni, T. maritima* and *B. acidophila*. Bootstrap values greater than 50% are indicated. Arrows indicate plausible paths for lateral gene transfer events. This figure is taken from (Shaul, *et al.*, 2006).

In the case of *B. cereus,* work by Shaul *et al* has indicated a recent acquisition of LysKRS by *B. cereus* from a Pyrococcal source (Figure 1.12) (Shaul, *et al.*, 2006). The function of the LysKRS enzyme in *B. cereus* is not entirely clear however it has been shown to be essential, and expressed predominantly in stationary phase (Ataide, *et al.*, 2005). In addition it was shown that the class I LysKRS was able to function in concert with the class II enzyme to aminoacylate a novel tRNA species known as $tRNA^{OTHER}$(Ataide, *et al.*, 2005). The function of $tRNA^{OTHER}$ has yet to be accurately established, however it has been shown not to be essential for the survival of *B. cereus* (Foy and Devine, unpublished results). Of interest also is the presence of a T-box regulatory element upstream of the *lysK* gene of *B. cereus* encoding the class I LysKRS enzyme (Ataide, *et al.*, 2005). This is the first recorded instance of a lysyl tRNA synthetase gene whose expression is regulated by a T-box mechanism.

1.10 Objectives of this study

In this study we investigated the functionality of a rare T-box regulatory element, namely that of *B. cereus lysK*. An analysis of 891 sequenced bacterial genomes showed that there are only six lysyl tRNA synthetase genes regulated by T-box mechanisms. This

relatively low level of T-box regulation for this species of tRNA synthetase led us to ask (i) is this a functional regulatory element, i.e. does it respond in to changes in the level of charged $tRNA^{LYS}$; (ii) is a T-box regulated *lysS* gene in *B. subtilis* compatible with normal growth and (iii) what feature of the *lysS* T-box regulatory element make it suboptimal for regulation of *lysS* in *B. subtilis*. Results from these experiments led us to investigate (i) the possibility of a lack of specificity in tRNA recognition of the *asnS* T-box regulatory element in *B. cereus*; (ii) potential lack of specificity of tRNA recognition in T-box regulatory elements whose cognate tRNA synthetases charge tRNAs with amino acids from mixed codon boxes in *B. subtilis*.

Chapter 2

Materials And Methods

2.1 Bacterial strains and growth conditions

Bacillus subtilis strains 168, JH642, 1A10, 1A31, 1A58, 1A75, 1A79, 1A765 and *B. cereus* 14579 were the strains used in this study. Cells were grown routinely at 37°C in LB media, Spizizen's minimal media (Anagnostopoulos and Spizizen, 1961) or Basal Limitation media (Stulke, *et al.*, 1993). All cloning was carried out in *E. coli* strains TG1 and TP611. Transformation of *B. subtilis* and *E. coli* was performed as described in [Harwood and Cutting, (1990), Sambrook, *et al.*, (1989)] Antibiotics were added to the media at the following concentrations: 100μg/ml ampicillin, 100μg/ml spectinomycin, 3μg/ml chloramphenicol, 2μg/ml erythromycin, 1.5μg/ml Phleomycin, 10μg/ml kanamycin. IPTG was added to cultures at various concentrations indicated in the text. *B. subtilis* and *E. coli* strains were stored as glycerol stocks at a 25% v/v concentration at -70°C. *B. cereus* electroporations were performed at 300V, 800mA and 4kΩ (Hornstra, *et al.*, 2005).

2.2 General Methods

Chromosomal DNA was isolated from *B. subtilis* and *B. cereus* using the Edge Biosystems chromosomal DNA purification kit according to the manufacturerer's protocol. Plasmid DNA was isolated either by a modified boiling lysis method (Sambrook, *et al.*, 1989) and further purified through DNA binding columns using the Concert Rapid PCR Purification System (Invitrogen, Carlsbad, Ca) according to the manufacturers instructions, or by using the Genelute Plasmid miniprep kit (Sigma Aldrich, St Louis, MO, USA). PCR amplification was achieved using either Taq polymerase (Invitrogen, Carlsbad, Ca) or high fidelity KOD polymerase (Calbiochem-Novabiochem Corp. USA) as described in the manufacturers protocol. Sequencing for construct and strain verification was carried out by MWG Biotech-Germany (Ebersburg, Germany) and GATC Biotech (Konstanz, Germany). Standard DNA manipulations and cloning procedures were carried out as described previously (Sambrook *et al.*, 1989). PCR mutagenesis was carried out as for Stratagene Quikchange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA) or by standard overlapping PCR (Sambrook, *et al.*, 1989). Ligation independant cloning (LIC) (Aslanidis & de Jong, 1990) was carried out by incubating 0.2pmol of each fragment with 2.5mM dTTP and T4 DNA polymerase. Plasmid pEB15 was linearized using *Sma*I, gel purified and treated with T4 DNA polymerase and 2.5mM dATP. A mix of approximately 5ng of prepared vector and 15ng of insert DNA was used to transform *E. coli*. Resulting plasmids were transformed into *B. subtilis*. PCR primers

were purchased from MWG Biotech-Germany (Ebersburg, Germany) and Sigma-Aldrich (St. Louis, MO, USA). For Southern blot analysis the DNA was transferred to positively charged biodyne membrane (Pall Gelman, Ann Arbor, MI, USA) by vacuum blotting. The DNA was then UV crosslinked by exposure to 150mJ UV. Dig labeled probes (Roche, East Sussex, UK) were prepared as per manufacturers protocol and hybridized to the filter using high concentration SDS buffers. Subsequent washes and detection were carried out using the Dig detection kit (Roche, East Sussex, UK). Analysis of IPTG dependance of expression was performed by growing relevant strains in LB containing 1mM IPTG until an OD_{600} of approximately 0.5 and then plating 10 fold dilutions of the culture onto LB plates containing either 1mM IPTG or no IPTG. IPTG dependence of expression was inferred by growth on plates containing IPTG and no growth on plates without IPTG.

2.3 Detection of tRNA charging by Northern analysis

Cell lysates of exponentially growing cultures were harvested by centrifugation (5000 rpm, 5 minutes, 4°C) and all subsequent procedures were carried out at 4°C. The cell pellet was snap frozen in an ethanol/dry ice bath. The cell pellet was resuspended in 0.3ml 0.3M sodium acetate, 10mM EDTA, pH 4.5 mixed with 0.3g glass beads. Phenol:chloroform (0.3ml., pH4.7) was added to each tube. The cells were disrupted using a bead beater. Samples were then centrifuged and the aqueous layer was collected. The phenol:chloroform extraction was repeated. The RNA was then ethanol precipitated twice. The first pellet was resuspended in 60µl 0.3M sodium acetate (pH 4.5) and then reprecipitated using ethanol. The pellet was air dried while on ice. The RNA was resuspended in 100µl 10mM sodium acetate and quantified spectrophotometrically. 2µg of total RNA was fractionated on a 14% polyacrylamide gel (61mm X 81mm gel, pH5) with 0.3M (pH5) and 7M urea. The electrophoresis buffer was 0.3M sodium acetate pH5). Gels were run using the Biorad Mini Protein 3 apparatus (Hercules, CA, USA) at 50V for 24 hours at 4°C. The electrophoresis buffer was changed every 7 hours to maintain an acid pH. The separated RNA samples were electroblotted onto positively charged 0.45µM Biodyne B nylon membranes (Pall, East Hills, NY, USA) using 10mM Tris acetate (pH7.8), 5mM sodium acetate, 0.5mM EDTA buffer at 44 volts for 2 hours at 4°C. The RNA was then UV crosslinked to the membrane. An oligonucleotide probe complementary to nucleotides 26-51 of *B. subtilis* $tRNA^{LYS}$ was labeled with either the DIG oligonucleotide Tailing Kit (Roche, East Sussex, UK) or biotin (New England Biolabs, USA). Filters were pre-hybridized for 1

hour at 42°C with 22mg/ml salmon sperm DNA and hybridized with the labeled oligonucleotide probe to the membranes at 42°C overnight in standard hybridization buffer according to the manufacturers instructions. Filters were wshed as follows: one wash for 15 minutes at 42°C with 6xSSC containing 0.1% SDS and 3x15 minute washes at 42°C using 4xSSC containing 0.1% SDS (Sambrook, *et al.*, 1989). Chemiluminescent detection was performed according to the manufacturers instructions for both the DIG labeling kit (Roche, East Sussex, UK) and the NEB blot phototope kit (New England Biolabs, USA). Deacylation of tRNA samples was achieved by incubation for 30 minutes at 70°C after addition of an equal volume of 0.1M Tris pH9.5, 0.1 M NaCl. Densitometry was performed using Gelworks software package (UVP, Upland, CA, USA).

2.4 β-galactosidase assay to measure promoter activity

β-galactosidase activity was measured as described previously (Miller, 1972). The promoter fusion strains were streaked onto LB plates from frozen stocks to obtain isolated colonies, which were used to inoculate 5.0ml overnight cultures, and grown overnight. For strains, which required IPTG for growth, the promoter fusion strains were lawned onto LB plates containing 1mM IPTG and grown overnight at room temperature to achieve an exponentially growing culture after 16 hours. These cells were then used to inoculate 10ml starter cultures which were grown at 1mM IPTG to an OD_{600} of 0.2. These starter cultures were then used to inoculate growth cultures. For strains grown in minimal media 50ml overnight cultures were inoculated from plates at 1p.m. and grown overnight. Overnight cultures were then used to inoculate minimal media growth cultures at an OD_{600} of 0.1. Samples were collected by centrifugation at 5000rpm for 5 minutes at 4°C and decanting the supernatant. Pellets were resuspended in Z buffer (Miller, 1972) containing 100μg/ml of lysozyme, 10μg/ml DNase and 1mM dithiothreitol (DTT) and then lysed in eppendorf tubes by incubating at 37°C for 25 minutes. The lysed samples were spun in a microfuge at 13,000 rpm for 5 minutes. The supernatant was transferred to a fresh eppendorf and was used for both the β-galactosidase assay and protein quantification. Protein quantification was performed using the BCA assay according to the manufacturers instructions (Biorad, Hercules, CA, USA) using IgG as the protein standard (Bradford, 1976). Aliquots of lysate were added to 100μl reaction mixtures, in a 96 well plate, containing o-nitrophenyl-D-galactopyranoside (ONPG) as a substrate. This plate was incubated at 28°C. The

reaction was stopped by the addition of 60µl of 1M Na_2CO_3 and the time of the reaction recorded. The OD_{420} was read and β-galactosidase units determined using the following formula

$$\beta\text{-gal units} = \frac{(OD_{420})(3.3)}{(\text{mg/ml protein})(\text{ml sample})(\text{minutes of reaction})(0.00486)}$$

2.5 Green Fluorescent Protein assay of promoter fusion

Reporter strains containing *gfp* transcriptional fusions were grown overnight in BLM at 37°C. Overnight cultures were diluted back to OD_{600} 0.02 in a final volume of 100µl in a 96 well plate (Nunc, black optical bottom). Plates were covered with lids to prevent evaporation. Cultures were incubated at 37°C with constant shaking in a Synergy Biotek II platereader for 10 hours. Fluorescence readings (emission 528/20nm) and OD_{600} were measured every 10 minutes.

Wild-type *B. subtilis* was grown on each plate in triplicate to establish background fluorescence. Natural fluorescence values of wild type *B. subtilis* (average of 3 cultures) at each OD_{600} reading was subtracted from raw fluorescence values of reporter strains at the same OD_{600}. This allowed us to establish accurate GFP values for the reporterstrains at each time point. Specific GFP activity for each strain was established by dividing GFP values at each time point by the OD_{600}.

2.6 Analysis of sequenced bacterial genomes for the presence of T-box motifs upstream of AARS genes

We searched for aminoacyl tRNA synthetase genes in 891 bacterial genomes downloaded from GenBank by BLASTP comparison to the corresponding AARS genes from *Bacillus subtilis*. We used homology to the *E. coli* glutaminyl tRNA synthetase to search for T-box regulation in this gene. Only hits with E-values less than 10^{-10} were retained. For each such hit, we extracted the region from the end of the previous gene to the start codon for the gene with the significant hit. We then searched for matches to the consensus T-box motif TGGNACCGCG. We allowed up to two mismatches ('N') in the last six positions. Sequences containing potential T-box sequences were then examined manually for their ability to form terminator and mutually exclusive anti-terminator structures.

2.7 Generation times

Generation times were calculated using time points from the logarithmic phase of the growth curve. The generation times were established using the following formula,

Generation time (min.) = 2(t) / [logb-logB]

t = time between intervals in minutes

b = final OD_{600}

B= starting OD_{600}

2.8 Strain construction

Strain NF33.1 [*trpC2 lys amyE*::pBCJ307.5 ($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R]

Strain NF33.1 was constructed by integration of the plasmid pBCJ307.5 into the amylase locus of *B. subtilis* 1A765 and a Cm^R colony was characterized. Integration at the amylase locus was confirmed by starch test and PCR.

Strain NF37.1 [*trpC2 pheA1 amyE*::pBCJ307.5 ($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R]

Strain NF37.1 was constructed by integration of the plasmid pBCJ307.5 into the amylase locus of *B. subtilis* JH642 and a Cm^R colony was characterized. Integration at the amylase locus was confirmed by starch test and PCR.

Strain NF52.9 [*trpC2 amyE*::pNF48.44 ($P_{T\text{-}box}$-*lysK*) $Spec^R$]

A 1.95kb fragment containing the *B. cereus lysK* promoter, leader region and structural gene was amplified by PCR from chromosomal DNA using primers 36F and oNJF9R. This fragment was digested with EcoR1 and cloned into the corresponding site of pBCJ102.1 (Jester B, personal communication) resulting in the plasmid pNF30.18. A 2567bp fragment including flanking terminator sequences was amplified from pNF30.18 using the pBluescript primers T7 and M13 reverse. This fragment was then 5' phosphorylated using T4 polynucleotide kinase (Promega) and cloned into the EcoRV site of pDG1730 to produce the plasmid pNF48.44. Strain NF52.9 was constructed by integration of the plasmid pNF48.44 into the amylase locus of *B. subtilis* 168 and a $Spec^R$ colony was characterized.

Strain NF53.4 [*trpC2* (P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R]

A 516bp fragment containing the upstream region of the *asnS* gene of *B. subtilis* was amplified by PCR from chromosomal DNA using primers oNJF16F and oNJF16R. This fragment was digested with HindIII and cloned into the corresponding site of the plasmid pMUTINXZ (this is a pMUTIN4 based plasmid with the *lacZ* gene removed; Vagner, *et al.*, 1998; Jester B, personal gift) resulting in the plasmid pNF40.10. Strain NF53.4 was constructed by integration of the plasmid pNF40.10 at the *asnS* locus of *B. subtilis* 168. Following this the non-integrating plasmid pMAP65 was added to the strain to reduce non-

IPTG induced readthrough from the P_{Spac} promoter. An Em^R,$Phleo^R$ and IPTG dependant colony was characterized.

Strain NF54.13 [*trpC2 amyE*::pNF48.44($P_{T\text{-}box}$-*lysK*) Δ*lysS*Kan^R]

Strain NF54.13 was constructed by integration of the plasmid pBCJ144.3 at the *lysS* locus of *B. subtilis*. This integration removes a central portion of the *lysS* gene. A $Spec^R$,Kan^R resistant colony was characterized.

Strain NF58.1 [*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R)]

Strain NF60.6 was constructed by integration of the plasmid pNF40.10 into strain BCJ363.1 (*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ* Cm^R).

Strain NF60.6 [*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R]

Strain NF60.6 was constructed by transforming the strain NF58.1 with the plasmid pMAP65. Successful transformation of NF58.1 with pMAP65 was comfirmed by PCR and tests for IPTG dependance of *asnS* expression.

Strain NF87.3 [(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R]

Strain NF87.3 was constructed by integration of p2XZ.34 (Jester B, personal communicaton). Plasmid pMAP65 was added to the resultant strain to produce strain NF87.3.

Strain NF89.2 [*trpC2 amyE*::pNF65.1($P_{T\text{-}box\ lysK\ AAU}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R]

A 150bp fragment containing the specifier loop region of the *B. cereus lysK* T-box leader region was amplified from chromosomal DNA using the primers 36F and oNJF26R. Another 250bp fragment containing the specifier loop region of the *B. cereus lysK* T-box leader region was amplified from chromosomal DNA using the primers oNJF26F and 36R. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers 36F and 36R to reconstitute the *B. cereus lysK* T-box leader region with the specifier codon mutated from AAA to AAU. This fragment was then digested with EcoRI and BamHI and cloned into the corresponding sites of pDG268 (Antoniewski, *et al.*, 1990) to create the plasmid pNF65.1.

Plasmid pNF65.1 was then integrated into the amylase locus of strain NF53.4 to create strain NF89.2. Strains were verified by starch test, PCR and IPTG dependance of *asnS* expression analysis.

Strain NF92.1 [*trpC2 amyE*::pNF83.1($P_{T\text{-}box\ lysK\ AAG}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R]

A 150bp fragment containing the specifier loop region of the *B. cereus lysK* T-box leader region was amplified from chromosomal DNA using the primers 36F and oNJF25R. Another 250bp fragment containing the specifier loop region of the *B. cereus lysK* T-box leader region was amplified from chromosomal DNA using the primers oNJF25F and 36R. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers 36F and 36R to reconstitute the *B. cereus lysK* T-box leader region with the specifier codon mutated from AAA to AAG. This fragment was then digested with EcoRI and BamHI and cloned into the corresponding sites of pDG268 (Antoniewski, *et al.*, 1990) to create the plasmid pNF83.1. Plasmid pNF83.1 was then integrated into the amylase locus on strain NF53.4 to create strain NF92.1. Strains were verified by starch test, PCR and IPTG dependance of *asnS* expression analysis.

Strain NF93.3 [*trpC2 amyE*::pNF84.1($P_{T\text{-}box\ lysK\ AAC}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R]

A 150bp fragment containing the specifier loop region of the *B. cereus lysK* T-box leader region was amplified from chromosomal DNA using the primers 36F and oNJF27R. Another 250bp fragment containing the specifier loop region of the *B. cereus lysK* T-box leader region was amplified from chromosomal DNA using the primers oNJF27F and 36R. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers 36F and 36R to reconstitute the *B. cereus lysK* T-box leader region with the specifier codon mutated from AAA to AAC. This fragment was then digested with EcoRI and BamHI and cloned into the corresponding sites of pDG268 (Antoniewski, *et al.*, 1990) to create the plasmid pNF84.1. Plasmid pNF84.1 was then integrated into the amylase locus on strain NF53.4 to create strain NF93.3. Strains were verified by starch test, PCR and IPTG dependance of *asnS* expression analysis.

Strain NF95.1 [*trpC2 amyE*::pNF65.1($P_{T\text{-}box\ lysK\ AAU}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R]

Plasmid pNF65.1 was integrated into the amylase locus of strain NF87.3 to create strain NF95.1. Strains were verified by starch test, PCR and IPTG dependance of *lysS* expression analysis.

Strain NF98.3 [*trpC2 amyE*::pNF83.1($P_{T\text{-}box\ lysK\ AAG}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R]

Plasmid pNF83.1 was integrated into the amylase locus of strain NF87.3 to create strain NF98.3. Strains were verified by starch test, PCR and IPTG dependance of *lysS* expression analysis.

Strain NF99.2 [*trpC2 amyE*::pNF84.1($P_{T\text{-}box\ lysK\ AAC}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R]

Plasmid pNF84.1 was integrated into the amylase locus of strain NF87.3 to create strain NF99.2. Strains were verified by starch test, PCR and IPTG dependance of *lysS* expression analysis.

Strain NF113.5/113.6 [*trpC2 lysS*::pNF112.5 ($P_{T\text{-}box\ lysK}$*lysS*2 Cm^R)]

A 423bp fragment containing the *B. cereus lysK* leader region was amplified from chromosomal DNA using the primers 36F and oNJF15R (contains a tag which has homology with the 5' end of the *B. subtilis lysS* gene). Another 672bp fragment corresponding to the first 672bp of the *B. subtilis lysS* gene was amplified from chromosomal DNA using the primers oNJF15F (contains a tag which has homology with the 3' end of the *B. cereus lysK* leader region) and 3R/2. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers 36F and 3R/2 to produce a product with the *B. cereus lysK* leader region fused to the 5' 672bp of the *B. subtilis lysS*. This product was then digested with EcoRI and BamHI and cloned into the corresponding sites on pBCJ102.1 (Jester, personal communication) to produce the plasmid pNF112.5. Insertion of the PCR fragment into the EcoRI and BamHI sites results in the $P_{T\text{-}box\ lysK}$*lysS*2 fusion being flanked by two transcriptional terminators. The plasmid pNF112.5 was then integrated onto the chromosome of *B. subtilis* 168 at the *lysS* locus to produce the strains NF113.5 and NF113.6.

Strain NF117.1 [*trpC2 lys lysS*::pNF112.5 ($P_{T\text{-}box\ lysK}$*lysS*2 Cm^R)]

Plasmid pNF112.5 was integrated onto the genome of *B. subtilis* 1A765 to create the strain NF117.1

Strain NF162.1 [*trpC2 amyE*::pNF141.2($P_{T\text{-}box\ lysK\ AAU\ G\text{-}C}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R]

A 150bp fragment containing the specifier loop region of the *B. cereus lysK* T-box leader region was amplified from pNF65.1 plasmid DNA using the primers 36F and oNJF38R. Another 250bp fragment containing the specifier loop region of the *B. cereus lysK* T-box leader region was amplified from chromosomal DNA using the primers oNJF38F and 36R. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers 36F and 36R to reconstitute the *B. cereus lysK* T-box leader carrying a G-C mutation at the second position of the specifier loop. This fragment was then digested with EcoRI and BamHI and cloned into the corresponding sites of pDG268 (Antoniewski, *et al.*, 1990) to create the plasmid pNF141.2. Plasmid pNF141.2 was then integrated into the amylase locus of strain NF53.4 to create strain NF162.1. Strains were verified by starch test, PCR and IPTG dependance of *asnS* expression analysis.

Strain NF163.1 [*trpC2 amyE*::pNF141.2($P_{T\text{-}box\ lysK\ AAU\ G\text{-}C}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R]

Plasmid pNF141.2 was integrated into the amylase locus of strain NF87.3 to create strain NF163.1. Strains were verified by starch test, PCR and IPTG dependance of *lysS* expression analysis.

Strain NF174.1 [*trpC2 amyE*::pNF139.2($P_{T\text{-}box\ lysK\ G\text{-}C}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R]

A 150bp fragment containing the specifier loop region of the *B. cereus lysK* T-box leader region was amplified from pBCJ307.5 plasmid DNA using the primers 36F and oNJF35R. Another 250bp fragment containing the specifier loop region of the *B. cereus lysK* T-box leader region was amplified from chromosomal DNA using the primers oNJF35F and 36R. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers 36F and 36R to reconstitute the *B. cereus lysK* T-box leader carrying a G-C mutation at the second position of the

specifier loop. This fragment was then digested with EcoRI and BamHI and cloned into the corresponding sites of pDG268 (Antoniewski, *et al.*, 1990) to create the plasmid pNF139.2. Plasmid pNF139.2 was then integrated into the amylase locus of strain NF53.4 to create strain NF174.1. Strains were verified by starch test, PCR and IPTG dependance of *asnS* expression analysis.

Strain NF175.7 [*trpC2 amyE*::pNF161.11($P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R]

A 493bp fragment containing the *asnS* leader region from *B. cereus* 14579 was amplified from chromosomal DNA using the primers NJF42F and oNJF42R. The resultant PCR fragment was digested with EcoRI and BamHI and cloned into the corresponding sites on pDG268 (Antoniewski, *et al.*, 1990) to create the plasmid pNF161.11. Plasmid pNF161.11 was then integrated into the amylase locus of strain NF53.4 to create strain NF175.7. Strains were verified by starch test, PCR and IPTG dependance of *asnS* expression analysis.

Strain NF177.1 [*trpC2 amyE*::pNF161.11($P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R]

Plasmid pNF161.11 was integrated into the amylase locus of strain NF87.3 to create strain NF177.1. Strains were verified by starch test, PCR and IPTG dependance of *lysS* expression analysis.

Strain NF178.7 [*trpC2 amyE*::pNF139.2($P_{\text{T-box}\ lysK\ G\text{-}C}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R]

Plasmid pNF139.2 was integrated into the amylase locus of strain NF87.3 to create strain NF178.7. Strains were verified by starch test, PCR and IPTG dependance of *lysS* expression analysis.

Strain NF179.2 [*trpC2 amyE*::pNF165.2 ($P_{L.\ bulgaricus\ \text{T-box}\ asnS}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R)]

A 350bp fragment containing the *asnS* leader region from *Lactobacillus bulgaricus delbrueckii* was amplified from chromosomal DNA using the primers NJF40F and oNJF40R. The resultant PCR fragment was digested with EcoRI and BamHI and cloned into the corresponding sites on pDG268 (Antoniewski, *et al.*, 1990) to create the plasmid pNF165.2. Plasmid pNF165.2 was then integrated into the amylase locus of strain NF53.4

to create strain NF179.2. Strains were verified by starch test, PCR and IPTG dependance of *asnS* expression analysis.

Strain NF187.1 [*trpC2 amyE*::pNF161.11($P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ*) Cm^R]

Plasmid pNF161.11 was integrated into the amylase locus of *B. subtilis* 168. Strains were verified by starch test and PCR.

Strain NF196.1 [*trpC2 amyE*::pNF144.2($P_{\text{T-box}\ lysK\ AAC\ G\text{-}C}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R]

A 150bp fragment containing the specifier loop region of the *B. cereus lysK* T-box leader region was amplified from pNF84.1 plasmid DNA using the primers 36F and oNJF37R. Another 250bp fragment containing the specifier loop region of the *B. cereus lysK* T-box leader region was amplified from chromosomal DNA using the primers oNJF37F and 36R. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers 36F and 36R to reconstitute the *B. cereus lysK* T-box leader carrying a G-C mutation at the second position of the specifier loop. This fragment was then digested with EcoRI and BamHI and cloned into the corresponding sites of pDG268 (Antoniewski, *et al.*, 1990) to create the plasmid pNF144.2. Plasmid pNF144.2 was then integrated into the amylase locus of strain NF53.4 to create strain NF196.1. Strains were verified by starch test, PCR and IPTG dependance of *asnS* expression analysis.

Strain NF198.1 [*trpC2 amyE*::pNF144.2($P_{\text{T-box}\ lysK\ AAC\ G\text{-}C}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R]

Plasmid pNF144.2 was integrated into the amylase locus of strain NF87.3 to create strain NF198.1. Strains were verified by starch test, PCR and IPTG dependance of *asnS* expression analysis.

Strain NF205.1 [*trpC2 lysS*::pNF112.5 ($P_{T\text{-}box\ lysK}$*lysS* Cm^R) pBCJ307.5($P_{\text{T-box}\ lysK}$-*lacZ*) Cm^R]

Chromosomal DNA was prepared from strain BCJ363.1 and used to transform strain NF113.5 to create strain NF205.1. Strains were verified by starch test and PCR.

Strain NF206.1 [*trpC2 amyE*::pNF48.44($P_{T\text{-}box}$-*lysK*) Δ*lysS*Kan^R pBCJ307.5($P_{T\text{-box }lysK}$-*lacZ*) Cm^R]

Plasmid pBCJ307.5 was integrated into strain NF53.14 to create strain NF206.1. Strains were verified by starch test and PCR.

Strain NF232.1 [*lysK*::pNF222.1($P_{T\text{-box lysK}}$-lacZ) Cm^R]

A 910bp fragment containing the upstream leader region of *B. cereus* 14579 *lysK* was amplified from chromosomal DNA using the primers 36F/2 and 36R. The PCR product was digested with EcoRI and BamHI and cloned into the corresponding sites on the plasmid pDG268 (Antoniewski, *et al.*, 1990) to produce the plasmid pNF222.1. Plasmid pNF222.1 was then integrated into the *lysK* locus of *B. cereus* 14579 to create the strain NF232.1. Strains were verified by PCR.

Strain NF257.1 [*cysC1 trpC2*::pNF251.1 ($P_{T\text{-box }trpS}$-*gfp*) Spec^R]

The plasmid pEB15 (Botella, personal communication) was digested with EcoRI, treated with DNA polymerase I Klenow fragment and then re-ligated to destroy the EcoRI site. The resultant plasmid was called pNF244.1. The pBluescript2SK(-) (Strategene, La Jolla, California) multiple cloning site was then amplified with the primers T7 and 1201. The PCR product was treated with T4 polynucleotide kinase (Promega) and cloned into the SmaI site of pNF244.1 to create the plasmid pNF245.1. A 718bp fragment containing the upstream leader region of the *trpS* gene of *B. subtilis* 168 was amplified from chromosomal DNA using the primers oNJF43F and oNJF43R. The PCR product was digested with EcoRI and BamHI and cloned into the corresponding sites on pNF245.1 to create the plasmid pNF251.1. Plasmid pNF251.1 was integrated into the *trpS* locus of *B. subtilis* 1A79 to create the strain NF257.1.

Strain NF259.1 [*asnS*::pNF223.8($P_{T\text{-box }asnS}$-lacZ) Cm^R]

A 1600bp fragment containing the upstream leader region of *B. cereus* 14579 *asnS* was amplified from chromosomal DNA using the primers oNJF42F and oNJF50R. The PCR product was digested with EcoRI and BamHI and cloned into the corresponding sites on the plasmid pDG268 (Antoniewski, *et al.*, 1990) to produce the plasmid pNF223.8. Plasmid pNF223.8 was then integrated into the *asnS* locus of *B. cereus* 14579 to create the strain NF592.1. Strains were verified by PCR.

Strain NF260.1 [*pheA1 trpC2*::pNF249.4 ($P_{T\text{-}box\ leuS}$-*gfp*) $Spec^R$]

A 472bp fragment containing the upstream leader region of the *leuS* gene of *B. subtilis* 168 was amplified from chromosomal DNA using the primers oNJF47F and oNJF47R. The PCR product was digested with BamHI and cloned into the corresponding site on pNF245.1 to create the plasmid pNF249.4. Plasmid pNF249.4 was integrated into the *leuS* locus of *B. subtilis* JH642 to create the strain NF260.1.

Strain NF261.1 [*ilvA1 leuB8 metB5*::pNF249.4 ($P_{T\text{-}box\ leuS}$-*gfp*) $Spec^R$]

Plasmid pNF249.4 was integrated into the *leuS* locus of *B. subtilis* 1A75 to create the strain NF261.1.

Strain NF265.1 [*pheA1 trpC2 pheS*::pNF255.1 ($P_{T\text{-}box\ pheS}$-*gfp*)

A 457bp fragment containing the upstream leader region of the *pheS* gene of *B. subtilis* 168 was amplified from chromosomal DNA using the primers oNJF19F and oNJF19R. The PCR product was digested with EcoRI and BamHI and cloned into the corresponding sites on pNF245.1 to create the plasmid pNF255.1. Plasmid pNF255.1 was integrated into the *pheS* locus of *B. subtilis* JH642 to create the strain NF265.1.

Strain NF266.1 [*ilvA1 leuB8 metB5*::pNF255.1 ($P_{T\text{-}box\ pheS}$-*gfp*) $Spec^R$]

Plasmid pNF255.1 was integrated into the *pheS* locus of *B. subtilis* 1A75 to create the strain NF266.1.

Strain NF267.1 [*ilvA1 leuB8 metB5*::pNF253.1 ($P_{T\text{-}box\ ileS}$-*gfp*) $Spec^R$]

A 424bp fragment containing the upstream leader region of the *ileS* gene of *B. subtilis* 168 was amplified from chromosomal DNA using the primers oNJF45F and oNJF45R. The PCR product was digested with EcoRI and BamHI and cloned into the corresponding sites on pNF245.1 to create the plasmid pNF253.1. Plasmid pNF253.11 was integrated into the *ileS* locus of *B. subtilis* 1A75 to create the strain NF267.1

Strain NF277.1 [*trpC2*::pNF251.1 ($P_{T\text{-}box\ trpS}$-*gfp*) $Spec^R$]

Plasmid pNF251.1 was integrated into the *trpS* locus of *B. subtilis* 168 to create the strain NF277.1

Strain NF282.1 [*hisH2 trpC2*::pNF250.13 ($P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp*) $Spec^R$]

A 530bp fragment containing the upstream leader region of the *leuS* gene of *B. subtilis* 168 was amplified from chromosomal DNA using the primers oNJF46F and oNJF46R. The PCR product was digested with EcoRI and cloned into the corresponding site on pNF245.1 to create the plasmid pNF250.13. Plasmid pNF250.13 was integrated into the *hisS* locus of *B. subtilis* 1A58 to create the strain NF282.1.

Strain NF290.1 [*trpC2*::pNF250.13 ($P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp*) $Spec^R$ (P_{spac}-*gatCAB* Em^R) pMAP65 (*penP-lacI* Phl^R)]

A 590bp fragment containing the upstream leader region of the *gatCAB* genes of *B. subtilis* 168 was amplified from chromosomal DNA using the primers oNJF48F and oNJF48R. The PCR product was digested with EcoRI and BamHI and cloned into the corresponding sites of pMUTINXZ (Jester, personal communication) to produce the plasmid pNF213.1. Plasmid pNF213.1 was integrated into the *gatCAB* locus of *B. subtilis* 168 and plasmid pMAP65 (Petit, *et al.*, 1998) was added to produce the strain NF215.1. Strain NF215.1 was then transformed with chromosomal DNA from strain NF282.1 to create strain NF290.1. Strains were verified by IPTG dependance test and PCR.

Strain NF292.1 [*hisA1 thr5 trpC2*::pNF249.4 ($P_{T\text{-}box\ leuS}$-*gfp*) $Spec^R$]

Plasmid pNF249.4 was integrated into the *leuS* locus of *B. subtilis* 1A10 to create the strain NF292.1.

Strain NF293.1 [*trpC2 lys*::pNF253.1 ($P_{T\text{-}box\ ileS}$-*gfp*) $Spec^R$]

Plasmid pNF253.1 was integrated into the *ileS* locus of *B. subtilis* 1A765 to create the strain NF293.1.

Strain NF295.1 [*hisA1 trpC2 thr5*::pNF255.1 ($P_{T\text{-}box\ pheS}$-*gfp*) $Spec^R$]

Plasmid pNF255.1 was integrated into the *pheS* locus of *B. subtilis* 1A10 to create the strain NF295.1.

Strain NF302.1 [*ilvA1 leuB8 metB5*::pNF251.1 ($P_{T\text{-}box\ trpS}$-*gfp*) $Spec^R$]

Plasmid pNF251.1was integrated into the *trpS* locus of *B. subtilis* 1A75 to create the strain NF302.1.

Strain NF305.1 [*pheA1 trpC2 pheS*::pNF289.1 ($P_{T\text{-}box\ pheS\ STOP}$-*gfp*) $Spec^R$]

A 186bp fragment containing the specifier loop region of the *B. subtilis pheS* T-box leader region was amplified from pNF255.1 plasmid DNA using the primers oNJF19F and oNJF58F. Another 271bp fragment containing the specifier loop region of the *B. subtilis pheS* T-box leader region was amplified from pNF255.1 plasmid DNA using the primers oNJF19R and oNJF58R. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers oNJF19F and oNJF19R to reconstitute the *B. subtilis pheS* T-box leader with the specifier codon mutated from UUC to UAA. This fragment was then digested with EcoRI and BamHI and cloned into the corresponding sites of pNF245.5 to create the plasmid pNF289.1 Plasmid pNF289.1 was then integrated into the *pheS* locus of *B. subtilis* JH642 to create strain NF305.1.

Strain NF306.1 [*ilvA1 leuB8 metB5*::pNF288.2 ($P_{T\text{-}box\ leuS\ STOP}$-*gfp*) $Spec^R$]

A 170bp fragment containing the specifier loop region of the *B. subtilis leuS* T-box leader region was amplified from pNF249.4 plasmid DNA using the primers oNJF47F and oNJF57R. Another 302bp fragment containing the specifier loop region of the *B. subtilis leuS* T-box leader region was amplified from pNF249.4 plasmid DNA using the primers oNJF57F and oNJF47R. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers oNJF47F and oNJF47R to reconstitute the *B. subtilis leuS* T-box leader with the specifier codon mutated from CUC to UAA. This fragment was then digested with BamHI and cloned into the corresponding site of pNF245.5 to create the plasmid pNF288.2 Plasmid pNF288.2 was then integrated into the *leuS* locus of *B. subtilis* 1A75 to create strain NF306.1.

Strain NF307.1 [*ilvA1 leuB8 metB5*::pNF287.1 ($P_{T\text{-}box\ ileS\ STOP}$-*gfp*) $Spec^R$]

A 206bp fragment containing the specifier loop region of the *B. subtilis ileS* T-box leader region was amplified from pNF253.1 plasmid DNA using the primers oNJF45F and oNJF56R. Another 218bp fragment containing the specifier loop region of the *B. subtilis*

ileS T-box leader region was amplified from chromosomal DNA using the primers oNJF56R and oNJF45R. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers oNJF45F and oNJF45R to reconstitute the *B. subtilis ileS* T-box leader with the specifier codon mutated from AUC to UAA. This fragment was then digested with EcoRI and BamHI and cloned into the corresponding sites of pNF245.5 to create the plasmid pNF287.1 Plasmid pNF287.1 was then integrated into the *ileS* locus of B. subtilis 1A75 to create strain NF307.1.

Strain NF310.1 [*trpC2*::pNF250.13 ($P_{\text{T-box } hisS\text{-}aspS}$-*gfp*) $Spec^R$ (P_{spac}-*aspS* Em^R) pMAP65 (*penP-lacI* Phl^R)]

A 570bp fragment containing the upstream leader region of the *aspS* genes of *B. subtilis* 168 was amplified from chromosomal DNA using the primers oNJF49F and oNJF49R. The PCR product was digested with EcoRI and BamHI and cloned into the corresponding sites of pMUTINXZ (Jester, personal communication) to produce the plasmid pNF214.1. Plasmid pNF214.1 was integrated into the *aspS* locus of *B. subtilis* 168 to produce strain NF235.1. Plasmid pMAP65 (Petit, *et al.*, 1998) was added to produce the strain NF309.1. Strain NF309.1 was then transformed with chromosomal DNA from strain NF282.1 to create strain NF310.1. Strains were verified by IPTG dependance test and PCR.

Strain NF321.1 [*trpC2 lys*::pNF287.1 ($P_{\text{T-box } ileS \text{ STOP}}$-*gfp*) $Spec^R$]

Plasmid pNF287.1 was integrated onto the chromosome of *B. subtilis* 1A765 at the *ileS* locus to produce strain NF321.1.

Strain NF331.1 [*trpC2 amyE*::pNF327.1 ($P_{gltX\text{-T-box } cysES \text{ UGG}}$-*gfp*) $Spec^R$]

A 190bp fragment containing the specifier loop region of the *B. subtilis cysES* T-box leader region was amplified from pNF252.1 plasmid DNA using the primers oNJF44F and oNJF64R. Another 215bp fragment containing the specifier loop region of the *B. subtilis cysES* T-box leader region was amplified from pNF252.1 plasmid DNA using the primers oNJF64F and oNJF44R. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers oNJF44F and oNJF44R to reconstitute the *B. subtilis cysES* T-box leader with the specifier codon mutated from UGC to UGG. This fragment was then digested with EcoRI and

BamHI and cloned into the corresponding sites of pNF245.5 to create the plasmid pNF312.1

A 125bp fragment containing the promoter region of the *B. subtilis gltX* gene was amplified from chromosomal DNA using the primers oNJF67F and oNJF67R. This fragment was then digested with EcoRI and cloned into the corresponding site on pNF312.1 to create the plasmid pNF326.1.

A 530bp fragment containing the *gltX* promoter and mutated *cysES* T-box leader region was amplified from pNF326.1 plasmid DNA using the primers oNJF68F and oNJF68R which contain tags for ligation independent cloning (LIC). The resultant PCR fragment was cloned into the corresponding LIC sites on the plasmid pBP122 (Bisicchia, personal communication) to produce the plasmid pNF327.1. Plasmid pNF327.1 was then integrated into the amylase locus of *B. subtilis* 168 to produce strain NF331.1. Strains were verified by starch test and PCR.

Strain NF332.1 [*trpC2 amyE*::pNF328.1 ($P_{gltX\text{-T-box }cysES\text{ UAA}}$-*gfp*) $Spec^R$]

A 190bp fragment containing the specifier loop region of the *B. subtilis cysES* T-box leader region was amplified from pNF252.1 plasmid DNA using the primers oNJF44F and oNJF62R. Another 215bp fragment containing the specifier loop region of the *B. subtilis cysES* T-box leader region was amplified from pNF252.1 plasmid DNA using the primers oNJF62F and oNJF44R. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers oNJF44F and oNJF44R to reconstitute the *B. subtilis cysES* T-box leader with the specifier codon mutated from UGC to UAA. This fragment was then digested with EcoRI and BamHI and cloned into the corresponding sites of pNF245.5 to create the plasmid pNF308.1

A 125bp fragment containing the promoter region of the *B. subtilis gltX* gene was amplified from chromosomal DNA using the primers oNJF67F and oNJF67R. This fragment was then digested with EcoRI and cloned into the corresponding site on pNF308.1 to create the plasmid pNF324.11.

A 530bp fragment containing the *gltX* promoter and mutated *cysES* T-box leader region was amplified from pNF324.11 plasmid DNA using the primers oNJF68F and oNJF68R which contain tags for ligation independent cloning (LIC). The resultant PCR fragment was cloned into the corresponding LIC sites on the plasmid pBP122 (Bisicchia, personal

communication) to produce the plasmid pNF328.1. Plasmid pNF328.1 was then integrated into the amylase locus of *B. subtilis* 168 to produce strain NF332.1. Strains were verified by starch test and PCR.

Strain NF333.1 [*trpC2 amyE*::pNF329.1 ($P_{gltX\text{-T-box } cysES}$-*gfp*) $Spec^R$]

A 125bp fragment containing the promoter region of the *B. subtilis gltX* gene was amplified from chromosomal DNA using the primers oNJF67F and oNJF67R. This fragment was then digested with EcoRI and cloned into the corresponding site on pNF252.1 to create the plasmid pNF323.1.

A 530bp fragment containing the *gltX* promoter and *cysES* T-box leader region was amplified from pNF323.1 plasmid DNA using the primers oNJF68F and oNJF68R which contain tags for ligation independent cloning (LIC). The resultant PCR fragment was cloned into the corresponding LIC sites on the plasmid pBP122 (Bisicchia, personal gift) to produce the plasmid pNF329.1. Plasmid pNF329.1 was then integrated into the amylase locus of *B. subtilis* 168 to produce strain NF333.1. Strains were verified by starch test and PCR.

Strain NF334.1 [*trpC2 amyE*::pNF330.1 ($P_{gltX\text{-T-box } cysES\ UGU}$-*gfp*) $Spec^R$]

A 190bp fragment containing the specifier loop region of the *B. subtilis cysES* T-box leader region was amplified from pNF252.1 plasmid DNA using the primers oNJF44F and oNJF63R. Another 215bp fragment containing the specifier loop region of the *B. subtilis cysES* T-box leader region was amplified from pNF252.1 plasmid DNA using the primers oNJF63F and oNJF44R. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers oNJF44F and oNJF44R to reconstitute the *B. subtilis cysES* T-box leader with the specifier codon mutated from UGC to UGU. This fragment was then digested with EcoRI and BamHI and cloned into the corresponding sites of pNF245.5 to create the plasmid pNF311.1

A 125bp fragment containing the promoter region of the *B. subtilis gltX* gene was amplified from chromosomal DNA using the primers oNJF67F and oNJF67R. This fragment was then digested with EcoRI and cloned into the corresponding site on pNF311.1 to create the plasmid pNF325.1.

A 530bp fragment containing the *gltX* promoter and mutated *cysES* T-box leader region was amplified from pNF325.1 plasmid DNA using the primers oNJF68F and oNJF68R which contain tags for ligation independent cloning (LIC). The resultant PCR fragment was cloned into the corresponding LIC sites on the plasmid pBP122 (Bisicchia, personal communication) to produce the plasmid pNF330.1. Plasmid pNF330.1 was then integrated into the amylase locus of *B. subtilis* 168 to produce strain NF334.1. Strains were verified by starch test and PCR.

Strain NF335.2 [*cysC1 trpC2 amyE*::pNF327.1 ($P_{gltX\text{-T-box }cysES\text{ UGG}}$-*gfp*) $Spec^R$]

Plasmid pNF327.1 was integrated into at the amylase locus of *B. subtilis* 1A79 to create strain NF335.2. Strains were verified by starch test and PCR.

Strain NF336.1 [*cysC1 trpC2 amyE*::pNF328.1 ($P_{gltX\text{-T-box }cysES\text{ UAA}}$-*gfp*) $Spec^R$]

Plasmid pNF328.1 was integrated at the amylase locus of *B. subtilis* 1A79 to create strain NF336.1. Strains were verified by starch test and PCR.

Strain NF337.1 [*cysC1 trpC2 amyE*::pNF329.1 ($P_{gltX\text{-T-box }cysES}$-*gfp*) $Spec^R$]

Plasmid pNF329.1 was integrated at the amylase locus of *B. subtilis* 1A79 to create strain NF337.1. Strains were verified by starch test and PCR.

Strain NF338.1 [*cysC1 trpC2 amyE*::pNF330.1 ($P_{gltX\text{-T-box }cysES\text{ UGU}}$-*gfp*) $Spec^R$]

Plasmid pNF330.1 was integrated at the amylase locus of *B. subtilis* 1A79 to create strain NF338.1. Strains were verified by starch test and PCR.

Strain NF343.1 [*trpC2 amyE*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$]

A 125bp fragment containing the *gltX* promoter region of *B. subtilis* was amplified from pNF327.1 plasmid DNA using the primers oNJF68F and oNJF72R. This fragment was then cloned into the plasmid pBP122 (Bisicchia, personal gift) by ligation independent cloning to create the plasmid pNF339.1. The plasmid pNF339.1 was then integrated into the amylase locus of *B. subtilis* 168 to produce the strain NF343.1. Strains were verified by starch test and PCR.

Strain NF344.1 [*pheA1 trpC2 amyE*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$]

The plasmid pNF339.1 was integrated into the amylase locus of *B. subtilis* JH642 to produce the strain NF344.1. Strains were verified by starch test and PCR.

Strain NF345.1 [*cysC1 trpC2 amyE*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$]

The plasmid pNF339.1 was integrated into the amylase locus of *B. subtilis* 1A79 to produce the strain NF345.1. Strains were verified by starch test and PCR.

Strain NF347.1 [*trpC2 amyE*::pNF346.1 ($P_{gltX\text{-T-box }cysES\text{ UGA}}$-*gfp*) $Spec^R$]

A 315bp fragment containing the specifier loop region of the *B. subtilis cysES* T-box leader region was amplified from pNF329.1 plasmid DNA using the primers oNJF68F and oNJF70R. Another 215bp fragment containing the specifier loop region of the *B. subtilis cysES* T-box leader region was amplified from pNF329.1 plasmid DNA using the primers oNJF70F and oNJF68R. The products of these PCR reactions were cleaned through DNA binding columns, cleaned and used as template in a third PCR reaction using the primers oNJF68F and oNJF68R to reconstitute the *B. subtilis cysES* T-box leader with the specifier codon mutated from UGC to UGA. This fragment was then cloned into pBP122 (Bisicchia, personal gift) by ligation independent cloning to produce the plasmid pNF346.1 Plasmid pNF346.1 was then integrated into the amylase locus of *B. subtilis* 168 to produce strain NF347.1. Strains were verified by starch test and PCR.

Strain NF351.1 [*ilvA1 leuB8 metB5 amyE*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$]

The plasmid pNF339.1 was integrated into the amylase locus of *B. subtilis* 1A75 to produce the strain NF351.1. Strains were verified by starch test and PCR.

Strain NF352.1 [*lys trpC2 amyE*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$]

The plasmid pNF339.1 was integrated into the amylase locus of *B. subtilis* 1A765 to produce the strain NF352.1. Strains were verified by starch test and PCR.

Table2.1 Oligonucleotides used in this study

3R/2 -(BamHI)- **CGGGATCC**GCGTTGTGGTGAGTGATAAACGG

oNJF9R –(EcoRI)- **CGGAATTC**CTTGCGTTGCTTTTAACTGTGTT

oNJF15F-GGTGAGTAATATATGAGTCAAGAAGAGCATAACC

oNJF15R- TTGACTCATATATTACTCACCTCAATTATTTTTTG

oNJF16F–(HindIII)- **CCCAAGCTT**CTTTCTTGTTTTGGAGGGAAATATG

oNJF16R–(HindIII)-**CCCAAGCTT**CCTTCAGGTGCGCTTCCAGTC

oNJF21R- CCGCTTTGAGACAGATATG

oNJF19F- (EcoRI)- **CGGAATTC**GGCTGCCATCCTCGTGTATC

oNJF19R- (BamHI)- **CGGGATCC**CCTTCAATGAGCTTGCCG

24- ($tRNA^{LYS}$ probe)- TTCGACCCTCTGATTAAAAGTCAGAT

oNJF25F- CTTAATGAAAAAGGACTTGGAGCTGCGC

oNJF25R- GCAGCTCCAAGTCCTTTTTCATTAAGATG

oNJF26F- CTTAATGAAAAATGACTTGGAGCTGCGC

oNJF26R- GCAGCTCCAAGTCATTTTTCATTAAGATG

oNJF27F- CTTAATGAAAAACGACTTGGAGCTGCGC

oNJF27R- GCAGCTCCAAGTCGTTTTTCATTAAGATG

oNJF35F- GAAAGAGCGGTATACATCTTAATCAAAAAAGACTTGGAGCTGCGC

oNJF35R- GCGCAGCTCCAAGTCTTTTTTGATTAAGATGTATACCGCTCTTTC

36F- (EcoRI)- **CGGAATTC**CGATTTCGATTCGGAAAAGTTTG

36F/2- (EcoRI)- **CGGAATTC**CTCTAAATATGGAACAACACAGC

36R- (BamHI)- **CGGGATCC**GCTACTTCATACGCCCAATGC

oNJF37F- GAAAGAGCGGTATACATCTTAATCAAAAACGACTTGGAGCTGCGC

oNJF37R- GCGCAGCTCCAAGTCGTTTTTGATTAAGATGTATACCGCTCTTTC

oNJF38F- GAAAGAGCGGTATACATCTTAATCAAAAATGACTTGGAGCTGCGC

38R- TAATGAGGCACTTCTCG

oNJF38R- GCGCAGCTCCAAGTCATTTTTGATTAAGATGTATACCGCTCTTTC

oNJF40F-(EcoRI)- **CCGGAATTC**ATCTTGACAAAAGTTAAATCCGTA

oNJF40R- (BamHI)- **CGCGGATCC**ATCTCTTGCCTCTGCTAAAAATC

oNJF42F- (BamHI)- **CCGGGATCC**CATAGTGAAATTCTCCTTTGCTC

oNJF42R- (EcoRI)- **CCGGAATTC**GAAGAGAAATTCGGTGCGG

oNJF43F- (EcoRI)- **CCGGAATTC**CAGCTCCTTACGGCTAAATC

oNJF43R- (BamHI)- **CGCGGATCC**TGAAAACATCAACCTCATTTC

oNJF44F- (EcoRI)- **CCGGAATTC**CCGATTCGTGTTGCTGTAAC

oNJF44R- (BamHI)- **CGCGGATCC**ATGCTTCCCCCCGTTTC

oNJF45F- (EcoRI)- **CCGGAATTC**GGGATCATCTCCTTGAGTATG

oNJF45R- (BamHI)- **CGCGGATCC**CGGCATTAAGAGCGTGTC

oNJF46F-(EcoRI)- **CCGGAATTC**ACGCAGATCCCGATACG

oNJF46R-(EcoRI)-**CCGGAATTC**GGCAGAATATCCTGTGTTCC

oNJF47F- (BamHI)- **CGCGGATCC**GTCCTCAATCTCCGCCC

oNJF47R- (BamHI)- **CGCGGATCC**GCCAATATGTCTGCCATTTC

oNJF48F-(EcoRI)- **CCGGAATTC**CGGAGAAATATGTGGAGGTG

oNJF48R- (BamHI)- **CGCGGATCC**GCGCAGCCCTTTTGTTAC

oNJF49F-(EcoRI)- **CCGGAATTC** GGGATAAGGAGTGAAACGC

oNJF49R- (BamHI)- **CGCGGATCC**CGGCACCAAGTAGTCACG

oNJF50F- (NotI)- **ATAAGAATGCGGCCGC**GAGCAAAGGAGAATTTCACTATG

oNJF50R- (BamHI)- **CGCGGATCC**TTCGTGCTCTGTTTGTAAGTC

oNJF54- GCCCAGAAACGACCAGC

oNJF55- GTAACAGAAGCACATAACCGTG

oNJF56F- GAATCATGTAGTAAAGCCCTTTAG

oNJF56R- CTAAAGGGCTTTACTACATGATTC

oNJF57F- CTCTGAGGTTTATTCGGAAAC

oNJF57R- GTTTCCGAATAAACCTCAGAG

oNJF58F- CAAGAGGTTTAGCCAATTTC

oNJF58R- GAAATTGGCTAAACCTCTTG

oNJF59F -(EcoRI)- **CCGGAATTC**CTCCCCAATGTGCAATC

oNJF62F- GGGTTTTAGAAGTAAGCCCTTAAG

oNJF62R- CTTAAGGGCTTACTTCTAAAACCC

oNJF63F- GGGTTTTAGAAGTGTGCCCTTAAG

oNJF63R- CTTAAGGGCACACTTCTAAAACCC

oNJF64F- GGGTTTTAGAAGTGGGCCCTTAAG

oNJF64R- CTTAAGGGCCCACTTCTAAAACCC

oNJF65F- GCGACAGTACTGATACAAAAAGG

oNJF65R- CGGGATCTTGATCGAACAC

oNJF67F- -(EcoRI)- **CCGGAATTC**CTTGGATTTACAGGCCG**G**

oNJF67R- -(EcoRI)- **CCGGAATTC**CTCAGGCAAAAATGTATGAATC

oNJF68F- (LIC)- **CCGCGGGCTTTCCCAGC**CTTGGATTTACAGGCCGG

oNJF68R- (LIC)- **GTTCCTCCTTCCCACC**ATGCTTCCCCCCGTTTC

oNJF70F- GGGTTTTAGAAGTGAGCCCTTAAG

oNJF70R- CTTAAGGGCTCACTTCTAAAACCC

oNJF72R- (LIC)- **GTTCCTCCTTCCCACC** GTGCATAACGGACGCGTAC

Kan1-GGCGCTTCATAGAGTAATTCTG

Kan2- GATAGGGGTCCCGAGC

PxR-TCCGTCGCTATTGTAACCAGTTC

Spec1- GCCAGTCACGTTACGTTATTAG

Spec2- GAAAACAATAAACCCTTGCATATG

AmyF- GAGGGATTTTTGACTCCGAAG

AmyR- GATGGTTTCTTTCGGTAAGCTT

All primers are listed 5' (left) to 3' (right). Added clamps and restriction sites are indicated in bold

Table 2.2 Strains and Plasmids

Strains or Plasmid	Relevant characteristic/Genotype	Source/Reference
E. coli strains		
TG-1	*SupE hsdΔ5 thi Δ(lac-proAB)* F'[*traD*36 *pro AB*$^{+}$ lacI^{q} lacZΔM15]	Gibson, (1984)
Tp611	*RecBC hsdR^{-}M^{-} cyab10pcn*	Glaser, P, *et al.*, (1993)
B. cereus strains		
14579	Wild type isolate	Lawrence and Ford. (1916)
NF232.1	*lysK*::pNF222.1($P_{T\text{-box lysK}}$-lacZ) Cm^{R}	
NF259.1	*asnS*::pNF223.8($P_{T\text{-box asnS}}$-lacZ) Cm^{R}	
B. subtilis strains		
168	*trpC2*	Laboratory Stock
JH642	*pheA1 trpC2*	Laboratory Stock
1A10	*hisA1 thr5 trpC2*	BGSC
1A31	*hisA1 thr5 trpC2*	BGSC
1A58	*hisH2 trpC2*	BGSC
1A75	*ilvA1 leuB8 metB5*	BGSC
1A79	*cysC1 trpC2*	BGSC
1A765	*trpC2 lys*	BGSC
NF33.1	*trpC2 lys amyE*::pBCJ307.5 ($P_{T\text{-box }lysK}$-*lacZ*) Cm^{R}	pBCJ307.5→168
NF37.1	*trpC2 pheA1 amyE*::pBCJ307.5 ($P_{T\text{-box }lysK}$-*lacZ*) Cm^{R}	pBCJ307.5→JH642
NF52.9	*trpC2 amyE*::pNF48.44 ($P_{T\text{-box}}$-*lysK*)	pNF48.44→168

	$Spec^R$	
NF53.4	*trpC2* (P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF40.10→168
NF54.13	*trpC2 amyE*::pNF48.44($P_{T\text{-}box}$-*lysK*) Δ*lysS*Kan^R	pBCJ144.3→NF52.9
NF58.1	[*trpC2 amyE*::pBCJ307.5($P_{T\text{-box }lysK}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R)	pNF40.10 →BCJ363.1
NF60.6	*trpC2 amyE*::pBCJ307.5($P_{T\text{-box }lysK}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R	pMAP65 →NF58.1
NF87.3	(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R	p2xz.34 →168
NF89.2	*trpC2 amyE*::pNF65.1($P_{T\text{-box }lysK\ AAU}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF65.1→NF53.4
NF92.1	*trpC2 amyE*::pNF83.1($P_{T\text{-box }lysK\ AAG}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF83.1→NF53.4
NF93.3	*trpC2 amyE*::pNF84.1($P_{T\text{-box }lysK\ AAC}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF84.1→NF53.4
NF95.1	*trpC2 amyE*::pNF65.1($P_{T\text{-box }lysK\ AAU}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF65.1→NF87.3
NF98.3	*trpC2 amyE*::pNF83.1($P_{T\text{-box }lysK\ AAG}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF83.1→NF87.3
NF99.2	*trpC2 amyE*::pNF84.1($P_{T\text{-box }lysK\ AAC}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF84.1→NF87.3

NF113.5	*trpC2 lysS*::pNF112.5 ($P_{T\text{-box } lysK}$*lysS*2 Cm^R	pNF112.5→168
NF117.1	*trpC2 lys lysS*::pNF112.5 ($P_{T\text{-box } lysK}$*lysS*2 Cm^R	pNF112.5→1A765
NF162.1	*trpC2 amyE*::pNF141.2($P_{T\text{-box } lysK\ AAU\ G\text{-}C}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF141.2→NF53.4
NF163.1	*trpC2 amyE*::pNF141.2($P_{T\text{-box } lysK\ AAU\ G\text{-}C}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF141.2→NF87.3
NF174.1	*trpC2 amyE*::pNF139.2($P_{T\text{-box } lysK\ G\text{-}C}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF139.2→NF53.4
NF175.7	*trpC2 amyE*::pNF161.11($P_{B.\ cereus\ T\text{-box}\ asnS}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF161.11→NF53.4
NF177.1	*trpC2 amyE*::pNF161.11($P_{B.\ cereus\ T\text{-box}\ asnS}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF161.11→NF87.3
NF178.7	*trpC2 amyE*::pNF139.2($P_{T\text{-box } lysK\ G\text{-}C}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF139.2→NF87.3
NF179.2	*trpC2 amyE*::pNF165.2 ($P_{L.\ bulgaricus\ T\text{-box}\ asnS}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF165.2→NF53.4
NF187.1	*trpC2 amyE*::pNF161.11($P_{B.\ cereus\ T\text{-box}\ asnS}$-*lacZ*) Cm^R	pNF161.11→168
NF196.1	*trpC2 amyE*::pNF144.2($P_{T\text{-box } lysK\ AAC\ G\text{-}C}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF144.2→NF53.4

NF198.1	*trpC2 amyE*::pNF144.2($P_{T\text{-}box\ lysK\ AAC\ G\text{-}C}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl	pNF144.2→NF87.3
NF205.1	*trpC2 lysS*::pNF112.5 ($P_{T\text{-}box\ lysK}$*lysS*2 Cm^R) pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R	pNF112.5→BCJ363.1
NF206.1	*trpC2 amyE*::pNF48.44($P_{T\text{-}box}$-*lysS*1) Δ*lysS*Kan^R pBCJ307.5($_{PT\text{-}boxlysK}$-*lacZ*) Cm^R	pBCJ307.5→NF54.13
NF218.1	*trpC2 amyE*::pNF208.1($P_{T\text{-}box\ cysES}$-*lacZ*) Cm^R	pNF208.1→168
NF257.1	cysC1 *trpC2*::pNF251.1 ($P_{T\text{-}box\ trpS}$-*gfp*) $Spec^R$	pNF251.1→1A79
NF258.1	*cysC1 trpC2*::pNF252.1 ($P_{T\text{-}box\ cysES}$-*gfp*) $Spec^R$	pNF252.1→1A79
NF260.1	*pheA1 trpC2*::pNF249.4 ($P_{T\text{-}box\ leuS}$-*gfp*) $Spec^R$	pNF249.4→JH642
NF261.1	*ilvA1 leuB8 metB5*::pNF249.4 ($P_{T\text{-}box\ leuS}$-*gfp*) $Spec^R$	pNF249.4→1A75
NF265.1	*pheA1 trpC2 pheS*::pNF255.1 ($P_{T\text{-}box\ pheS}$-*gfp*) $Spec^R$	pNF255.1→JH642
NF266.1	*ilvA1 leuB8 metB5*::pNF255.1 ($P_{T\text{-}box\ pheS}$-*gfp*) $Spec^R$	pNF255.1→1A75
NF267.1	*ilvA1 leuB8 metB5*::pNF253.1 ($P_{T\text{-}box\ ileS}$-*gfp*) $Spec^R$	pNF253.1→1A75
NF277.1	*trpC2*::pNF251.1 ($P_{T\text{-}box\ trpS}$-*gfp*) $Spec^R$	pNF251.1→168
NF278.1	*trpC2*::pNF252.1 ($P_{T\text{-}box\ cysES}$-*gfp*) $Spec^R$	pNF252.1→168
NF282.1	*hisH2 trpC2*::pNF250.13 ($P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp*) $Spec^R$	pNF250.13→1A58
NF290.1	*trpC2*::pNF250.13 ($P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp*) $Spec^R$ (P_{spac}-*gatCAB* Em^R) pMAP65	pNF250.13→NF215.1

	(*penP-lacI* Phl^R)]	
NF292.1	*hisA1 thr5 trpC2*::pNF249.4 ($P_{T\text{-}box\ leuS}$-*gfp*) $Spec^R$	pNF249.4 →1A10
NF293.1	*trpC2 lys*::pNF253.1 ($P_{T\text{-}box\ ileS}$-*gfp*) $Spec^R$	pNF253.1 →1A765
NF295.1	*hisA1 trpC2 thr5*::pNF255.1 ($P_{T\text{-}box\ pheS}$-*gfp*) $Spec^R$	pNF255.1 →1A10
NF302.1	*ilvA1 leuB8 metB5*::pNF251.1 ($P_{T\text{-}box\ trpS}$-*gfp*) $Spec^R$	pNF251.1 →1A75
NF304.1	*cysB3 hisA1 trpC2*::pNF252.1 ($P_{T\text{-}box\ cysES}$-*gfp*) $Spec^R$	pNF252.1 →1A31
NF305.1	*pheA1 trpC2 pheS*::pNF289.1 ($P_{T\text{-}box\ pheS\ STOP}$-*gfp*) $Spec^R$	pNF289.1 →JH642
NF306.1	*ilvA1 leuB8 metB5*::pNF288.2 ($P_{T\text{-}box\ leuS\ STOP}$-*gfp*) $Spec^R$	pNF288.2 →1A75
NF307.1	*ilvA1 leuB8 metB5*::pNF287.1 ($P_{T\text{-}box\ ileS\ STOP}$-*gfp*) $Spec^R$	pNF287.1 →1A75
NF310.1	*trpC2*::pNF250.13 ($P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp*) $Spec^R$ (P_{spac}-*aspS* Em^R) pMAP65 (*penP-lacI* Phl^R	pNF250.13 →NF309.1
NF321.1	*trpC2 lys*::pNF287.1 ($P_{T\text{-}box\ ileS\ STOP}$-*gfp*) $Spec^R$	pNF287.1 →1A765
NF331.1	*trpC2 amyE*::pNF327.1 ($P_{gltX\text{-}T\text{-}box\ cysES\ UGG}$-*gfp*) $Spec^R$	pNF327.1 →168
NF332.1	*trpC2 amyE*::pNF328.1 ($P_{gltX\text{-}T\text{-}box\ cysES\ UAA}$-*gfp*) $Spec^R$	pNF328.1 →168
NF333.1	*trpC2 amyE*::pNF329.1 ($P_{gltX\text{-}T\text{-}box\ cysES}$-*gfp*) $Spec^R$	pNF329.1 →168
NF334.1	*trpC2 amyE*::pNF330.1 ($P_{gltX\text{-}T\text{-}box\ cysES\ UGU}$-*gfp*) $Spec^R$	pNF330.1 →168
NF335.2	*cysC1 trpC2 amyE*::pNF327.1 ($P_{gltX\text{-}T\text{-}box}$	pNF327.1 →1A79

	cysES UGG-*gfp*) $Spec^R$	
NF336.1	[*cysC1 trpC2 amyE*::pNF328.1 ($P_{gltX\text{-T-box}}$ *cysES* UAA-*gfp*) $Spec^R$	pNF328.1 →1A79
NF337.1	*cysC1 trpC2 amyE*::pNF329.1 ($P_{gltX\text{-T-box}}$ *cysES*-*gfp*) $Spec^R$	pNF329.1 →1A79
NF338.1	*cysC1 trpC2 amyE*::pNF330.1 ($P_{gltX\text{-T-box}}$ *cysES* UGU-*gfp*) $Spec^R$	pNF330.1 →1A79
NF343.1	*trpC2 amyE*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$	pNF339.1 →168
NF344.1	*pheA1 trpC2 amyE*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$	pNF339.1 →JH642
NF345.1	*cysC1 trpC2 amyE*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$	pNF339.1 →1A79
NF347.1	*trpC2 amyE*::pNF346.1 ($P_{gltX\text{-T-bocx}}$ *cysES* UGA-*gfp*) $Spec^R$	pNF346.1 →168
NF351.1	*ilvA1 leuB8 metB5 amyE*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$	pNF339.1 →1A75
NF352.1	*lys trpC2 amyE*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$	pNF339.1 →1A765
BCJ363.1	*trpC2 amyE*::pBCJ307.5 ($P_{\text{T-box } lysK}$-*lacZ*) Cm^R	Jester, personal gift
BCJ367.1	*trpC2 amyE*::pBCJ307.5 ($P_{\text{T-box } lysK}$-*lacZ*) Cm^R (P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R)]	Jester, personal gift
Plasmids		
pBCJ102.1	pBluescript based vector containing transcription terminator cassettes Ap^R	Jester, personal gift
pBCJ144.3	*B. subtilis lysS* knockout vector Kan^R	Jester, personal gift
pBCJ307.5	Transcriptional fusion vector for *B.*	Jester, personal gift

	cereus lysK Cm^R	
pDG268	Vector for integrating *lacZ* promoter fusions at the *amyE* locus by double crossover Ap^R Cm^R	Antoniewski, *et al.*, 1990
pDG1730	Vector for integration at the *amyE* locus in *B. subtilis* $Spec^R$	Guerot-Fleury, A-M *et al.*, (1996)
p2X.34	Vector for placing *B. subtilis lysS* under IPTG control Em^R	Jester, personal gift
pMUTIN-XZ	pMUTIN4 with the *lacZ* gene removed Em^R	Jester, personal gift
Pbluescript2 KS(-)	Cloning Vector Ap^R	Stratagene, La Jolla CA
pMAP65	Muticopy *B. subtilis* vector *penP-lacI* Phl^R Kan^R	Petit, *et al.*, 1998
pBP122	Vector for integrating *gfp* promoter fusions at the *amyE* locus by double crossover Ap^R $Spec^R$	Bissicchia, personal gift
pEB15	LIC vector for making *gfp* transcriptional fusions in *B. subtilis*	Botella, personal gift
pNF30.18	*B. cereus lysK* and leader region in pBCJ102.1 Ap^R	This Work
pNF40.10	516bp upstream of *asnS* in pMUTIN-XZ	This Work
pNF48.44	1.95kb of *B. cereus lysK* and leader region in pDG1730	This Work
pNF65.1	*lysK* leader region with specifier codon mutated to AAU in pDG268	This Work
pNF83.1	*lysK* leader region with specifier codon mutated to AAG in pDG268	This Work
pNF84.1	*lysK* leader region with specifier codon mutated to AAC in pDG268	This Work
pNF112.5	A 423bp fragment of *lysK* leader and	This Work

	fused to 672bp of the 5' end of *lysS* in pBCJ102.1	
pNF139.2	*lysK* leader region with G-C mutation in the specifier loop in pDG268	This Work
pNF141.2	*lysK* leader region with G-C mutation in the specifier loop with AAU specifier codon in pDG268	This Work
pNF144.2	*lysK* leader region with G-C mutation in the specifier loop with AAC specifier codon in pDG268	This Work
pNF161.11	A 493bp fragment upstream of *B. cereus asnS* in pDG268	This Work
pNF165.2	A 350bp fragment upstream of *L. delbrueckii bulgaricus asnS* in pDG268	This Work
pNF208.1	A 405bp fragment upstream of *cysES* of *B. subtilis* in pDG268	This Work
pNF213.1	A 590bp fragment upstream of the *gatCAB* genes of *B. subtilis* in pMUTIN-XZ	This Work
pNF214.1	A 570bp fragment upstream of the *aspS* genes of *B. subtilis* in pMUTIN-XZ	This Work
pNF222.1	A 910bp fragment upstream of *B. cereus lysK* in pDG268	This Work
pNF223.8	A 1600bp fragment upstream of *B. cereus asnS* in pDG268	This Work
pNF244.1	pEB15 with the EcoRI site destroyed	This Work
pNF245.1	pEB15 with the pBluescript MCS inserted at the SmaI site	This Work
pNF249.4	A 472bp fragment upstream of *B. subtilis leuS* in pNF245.5	This Work

pNF250.13	A 590bp fragment upstream of *B. subtilis hisS* in pNF245.5	This Work
pNF251.1	A 718bp fragment upstream of *B. subtilis trpS* in pNF245.5	This Work
pNF252.1	A 405bp fragment upstream of *B. subtilis cysES* in pNF245.5	This Work
pNF253.1	A 424bp fragment upstream of *B. subtilis ileS* in pNF245.5	This Work
pNF255.1	A 457bp fragment upstream of *B. subtilis pheS* in pNF245.5	This Work
pNF287.1	A 424bp fragment upstream of *B. subtilis ileS* with UAA specifier codon in pNF245.5	This Work
pNF288.2	A 472bp fragment upstream of *B. subtilis leuS* with UAA specifier codon in pNF245.5	This Work
pNF289.1	A 457bp fragment upstream of *B. subtilis pheS* with UAA specifier codon in pNF245.5	This Work
pNF308.1	A 405bp fragment upstream of *B. subtilis cysES* with UAA specifier codon in pNF245.5	This Work
pNF311.1	A 405bp fragment upstream of *B. subtilis cysES* with UGU specifier codon in pNF245.5	This Work
pNF312.1	A 405bp fragment upstream of *B. subtilis cysES* with UGG specifier codon in pNF245.5	This Work
pNF323.1	pNF252.1with the *gltX* promoter fragment cloned in	This Work
pNF324.11	pNF308.1 with the *gltX* promoter	This Work

	inserted at the EcoRI site	
pNF325.1	pNF311.1 with the *gltX* promoter inserted at the EcoRI site	This Work
pNF326.1	pNF312.1 with the *gltX* promoter inserted at the EcoRI site	This Work
pNF327.1	pBP122 with the *gltX* promoter-*cysES* T-box fragment from pNF326.1 cloned in	This Work
pNF328.1	pBP122 with the *gltX* promoter-*cysES* T-box fragment from pNF324.11 cloned in	This Work
pNF329.1	pBP122 with the *gltX* promoter-*cysES* T-box fragment from pNF323.1 cloned in	This Work
pNF330.1	pBP122 with the *gltX* promoter-*cysES* T-box fragment from pNF325.1 cloned in	This Work
pNF339.1	This is the *gltX* promoter fragment cloned into pBP122	This Work
pNF346.1	The *gltX* promoter –*cysES* leader fragment with the specifier mutated to UGA cloned into pBP122	This Work

Chapter 3

Investigation of the T-box transcriptional regulatory element of the class I lysyl tRNA synthetase of *Bacillus cereus*

3.1 Distribution of T-box regulated AARS genes amongst the sequenced bacterial genomes

The T-box anti-termination mechanism is a system of transcriptional regulation commonly used in the bacteria (particularly *Firmicutes*) to control the production of enzymes involved in tRNA charging, amino acid biosynthesis, synthesis of cofactors, nucleotides and metal ions (Gutiérrez-Preciado, *et al.*, 2009). In *Bacillus subtilis*, 14 out of 24 aminoacyl tRNA synthetases (AARSs) are transcriptionally regulated by a T-box mechanism. Among those AARS not regulated in this manner in *B. subtilis* is the class II synthetase *lysS*. We had previously noted the presence of a putative T-box regulatory element upstream of the class I *lysK* gene of *B. cereus* 14579 (Ataide, *et al.*, 2005). This was the first instance of T-box regulation observed for a lysyl tRNA synthetase gene. This discovery prompted us to investigate the sequenced bacterial genomes in order to find all the T–box regulated AARS genes so we could (i) establish the most commonly regulated and most rarely T-box regulated AARS genes and (ii) analyze the distribution of the T-box mechanism of regulation amongst the different bacterial phyla and (iii) assess the number of *lysS/K* genes that are regulated by a T-box mechanism.

To investigate this we undertook an analysis of all the fully sequenced bacterial genomes available in Genbank. We searched for all the AARS genes present on a given genome and analysed their upstream leader sequences for the presence of a T-box regulatory element. From this analysis we compiled a list of bacterial genomes containing T-box regulatory elements upstream of AARS genes and also a list of the instances of T-box regulation for the individual AARS genes. The full data set showing the results of this analysis is available on the disk accompanying this thesis. In total we found 976 T-box regulatory elements upstream of AARS genes in 891 fully sequenced bacterial genomes.

Figure 6.1 shows a bar graph illustrating the number of instances of T-box regulation for the different AARS genes. The data shows that the *ileS* gene is the most frequently T-box regulated AARS gene with 146 cases of T-box regulation of this gene amongst the 891 fully sequenced bacterial genomes analysed. The *ileS* gene was followed in terms of frequency of T-box regulation by *pheS* with 103 instances of T-box regulation and *thrS* with 94. The lowest instances of T-box regulation are found for *aspS* and *gltX* with only one instance each found in *Lactobacillus reuteri* and *Symbiobacterium thermophilum* repectively. It should be noted that *aspS* is co-regulated by a T-box mechanism with *hisS* in 60 of the bacterial genomes analyzed. It is also interesting to note that there are only 4 instances of T-box regulation of *hisS,* without the co-regulation of *aspS* by the same T-box

regulatory element. Theses four bacteria are *Geobacillus thermodenitrificans, Lactococcus lactis, Lactococcus lactis cremoris* MG1363 and *Lactococcus lactis cremoris* SK11.

We were also curious about the distribution of T-box regulated AARS genes among the different bacterial phyla. We wished to assess the number of instances of T-box regulation of AARS genes outside the *Firmicutes*. We used the same data set to address this. Results from this analysis are presented in figure 6.2.

The data shows that the vast majority of T-box regulated AARS genes are found in bacteria from the *Firmicute* phylum. There are 103 bacterial species from this phylum that contain AARS genes whose upstream leader regions contain T-box regulatory elements. There are also a large number (39) of bacterial species from the phylum *Actinobacteria* which contain AARS genes transcriptionally regulated by putative T-box regulatory elements. The other phyla which contain bacteria with T-box regulated AARS genes are the *Chloroflexi* (5; *Dehalococcoides CBDB1, Dehalococcoides* BAV1, *Dehalococcoides ethenogenes, Roseiflexus* RS-1, *Roseiflexus castenholzii*), the *Deinococcus-Thermus* (3; *Deinococcus geothermalis, Thermus thermophilus* HB27, *Thermus thermophilus* HB8), the *ß-proteobacteria* (*Verminephrobacter eiseniae*) and the *Other bacteria* (*Syntrophomonas wolfei*).

The class of bacterial species which contain the greatest number of AARS genes regulated by T-box regulatory elements are the *Bacilli*. The species *B. anthracis* and *B. cereus* contain the highest number of T-box regulated AARS genes with 14 (see attached Microsoft excel spreadsheet).

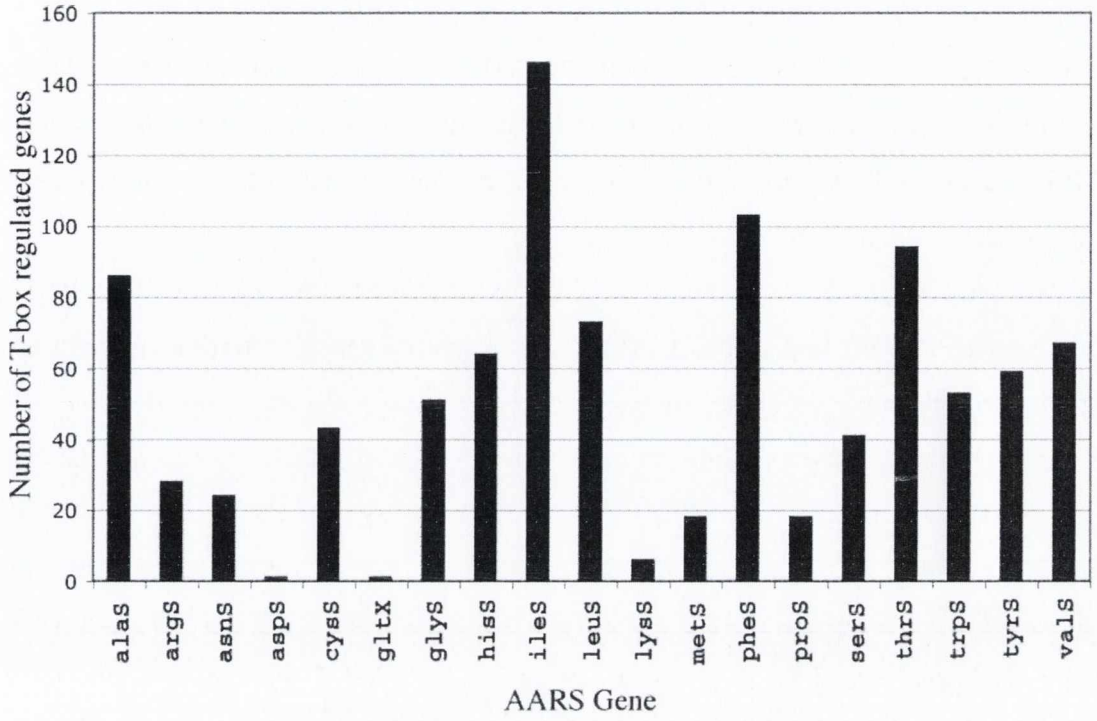


Figure 3.1 Analysis of the number of cases of T-box regulation for each AARS gene. The number of cases of T-box regulation for each AARS gene is indicated. Note that *aspS* is co-regulated with *hisS* by a single T-box regulatory element in 60 bacterial species.

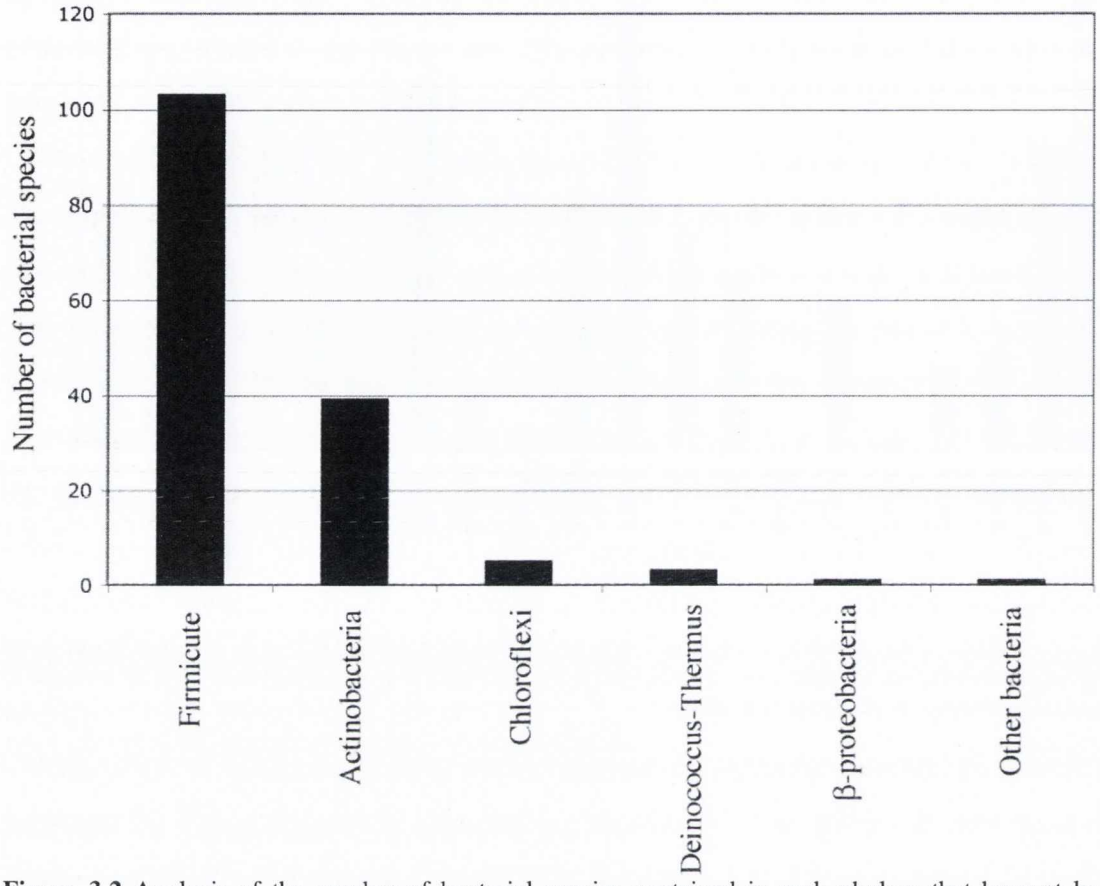


Figure 3.2 Analysis of the number of bacterial species contained in each phylum that have at least one AARS gene transcriptionally regulated by a T-box regulatory element.

3.2 Regulation of lysyl tRNA synthetases by the T-box anti-termination mechanism occurs very rarely

Of the 891 complete bacterial genome sequences analysed, only 6 *lysS/K* genes had a putative T-box regulatory element (Table 3.1). All 6 of these species encode more than one lysyl tRNA synthetase. Only *aspS* and *gltX* have less instances of T-box regulation. Five of the six species containing T-box regulated *lysS* genes were found to contain both a class I and a class II *lysS*. *Clostridium beijerinckii* however contains two copies of the class II *lysS* gene. These two copies share only 45% amino acid sequence identity. Of six putative T-box regulated *lysS* genes identified, four are class I LysKRS while those of *C. beijerinckii* and *S. thermophilum* are class II LysRS. The four *lysK* T-box regulatory elements are highly conserved at the nucleotide sequence level (Figure 3.3). Such similarity was not observed in the sequence of the putative regulatory elements of the *lysS* genes of either *C. beijerinckii* or *S. thermophilum* (data not shown).

3.3 Is the *lysK* T-box regulatory element of *B. cereus 14579* functional?

Our investigation of AARS genes that have putative T-box regulatory elements located upstream revealed that there are only six bacterial species that contain *lysS/K* genes with putative T-box regulatory elements upstream. We also noted that in these strains, two copies of the *lysS/K* gene are encoded. This may indicate that one copy is used for housekeeping functions in while the other is retained to perform another function. Analysis of the relative expression levels of the two *lysS* genes contained in the *B. cereus* genome showed that the class II *lysS* gene is expressed during exponential growth while the class I *lysK* is expressed predominantly during stationary phase indicating that the *lysK* gene does not perform the housekeeping function of tRNA charging during exponential growth (Ataide *et al.*, 2005). The *lysK* gene has a 380bp untranslated leader region which has many of the features associated with the T-box transcriptional anti-termination mechanism (Figure 3.4). These include the consensus T-box sequence, a series of stem loop RNA structures, a number of conserved nucleotides and an AAA specifier codon that indicates that this putative T-box mechanism responds to the level of $tRNA^{LYS}$ charging in the cell.

To assess whether the *B. cereus* 14579 *lysK* T-box element is functional, i.e. that it responds to the level of $tRNA^{LYS}$ charging in the cell, we created a transcriptional fusion of the *B. cereus* strain 14579's, *lysK* promoter and T-box leader region, to the *lacZ* reporter gene to create the plasmid pBCJ307.5. This construct was then integrated

into the chromosome in single copy of several genetic backgrounds of *B. subtilis* by double crossover at the amylase locus and expression levels monitored.

Strain BCJ363.1 ($P_{T\text{-}box\ lysK}$-*lacZ*) was used to examine expression of the *lysK* promoter and T-box region during growth in minimal media. This strain was grown in Spizizen's minimal media and the expression levels monitored (Figure 3.5). Accumulation of ß-galactosidase was low. Approximately 10-15 units throughout the growth cycle indicating

Table 3.1 Distribution of T-box regulated class I and class II *lysS* genes in completely sequenced genomes

Species	**Number of *lysS/K* genes**	**Description**
Bacillus cereus 14579	2	One class II *lysS* One T-box regulated class I *lysK*
Bacillus cereus 10987	2	One class II *lysS* One T-box regulated class I *lysK*
Bacillus thuringiensis Al Hakam	2	One class II *lysS* One T-box regulated class I *lysK*
Bacillus thuringiensis Konkukian	2	One class II *lysS* genes One T-box regulated class I *lysK*
Clostridium beijerinckii	2	Two class II *lysS* One T-box regulated
Symbiobacterium thermophilum	2	One class I *lysK* One T-box regulated class II *lysS*

a low level of promoter activity.

The expression profile of the ($P_{T\text{-}box\ lysK}$-*lacZ*) fusion under conditions of lysine starvation was then examined. To achieve this the plasmid pBCJ307.5 was integrated into the *amyE* locus of the lysine auxotroph strain 1A765 to create the strain NF33.1. In order to establish the level of lysine required to achieve lysine starvation of this strain, cultures were grown in decreasing lysine levels and the results are presented in figure 3.6. The data shows that as the level of lysine in the culture is decreased, the growth plateaus at a lower OD_{600}. We concluded that a lysine concentration of 20µg/ml yielding starvation at OD_{600} 1.0 was optimal for these experiments.

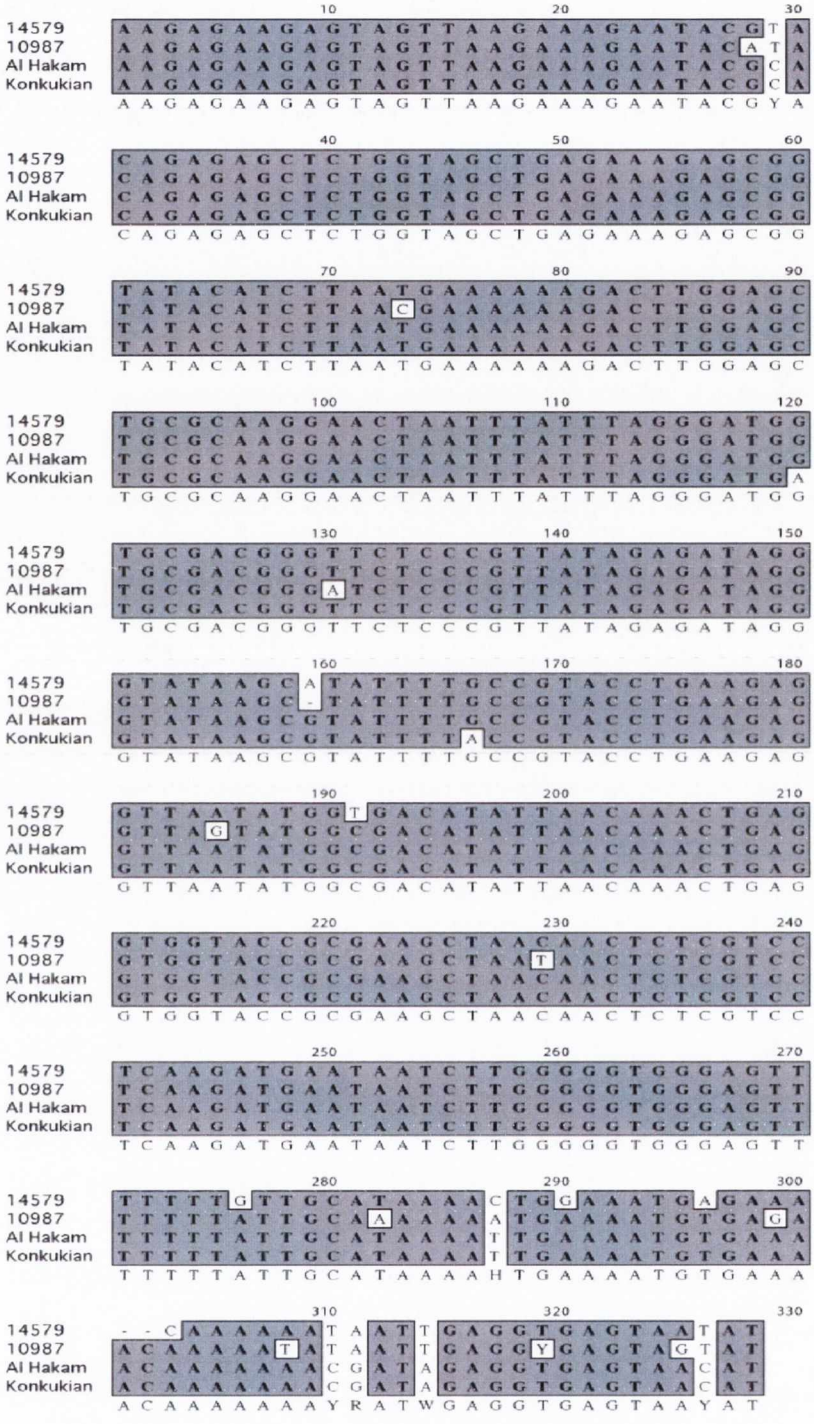


Figure 3.3 Clustal W alignment of the *lysK* leader region of the B. cereus 14579 and 10987 and B. thuringiensis species Al Hakam and Konkukian. The gray shaded regions indicate regions of nucleotide sequence identity

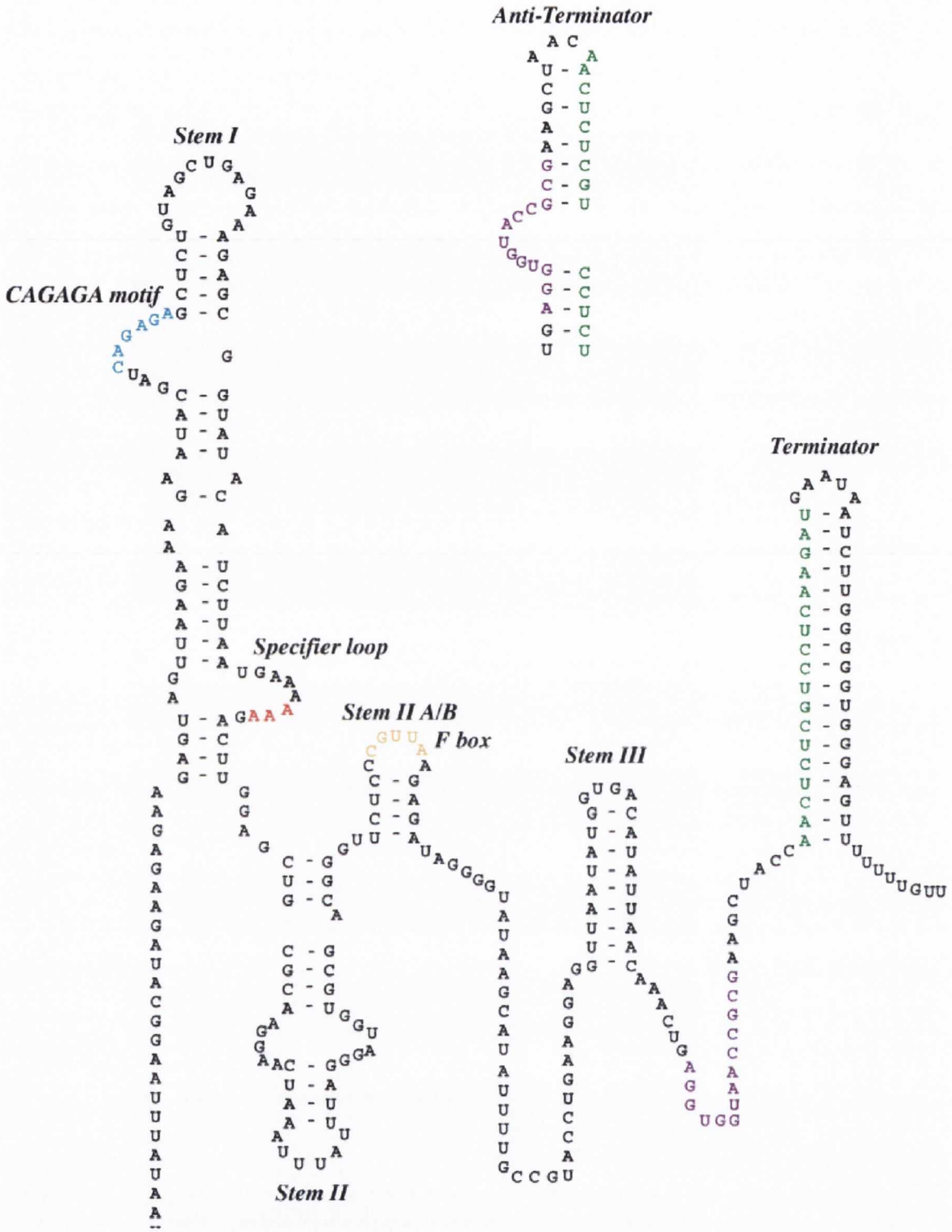


Figure 3.4 Proposed secondary strucuture of the *lysK* T-box regulatory element of *B. cereus*. The specifier codon (AAA) is indicated in red. Important motifs and structural regions are indicated. The T-box sequence is indicated in purple. Adapted from Ataide., *et al,* 2005.

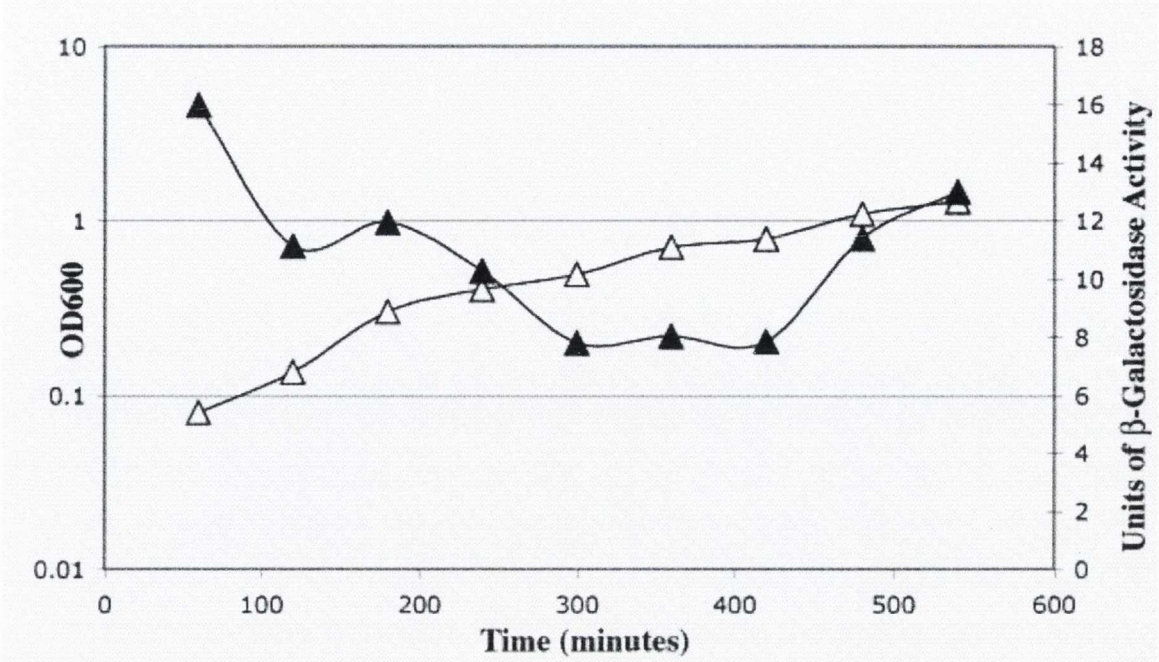
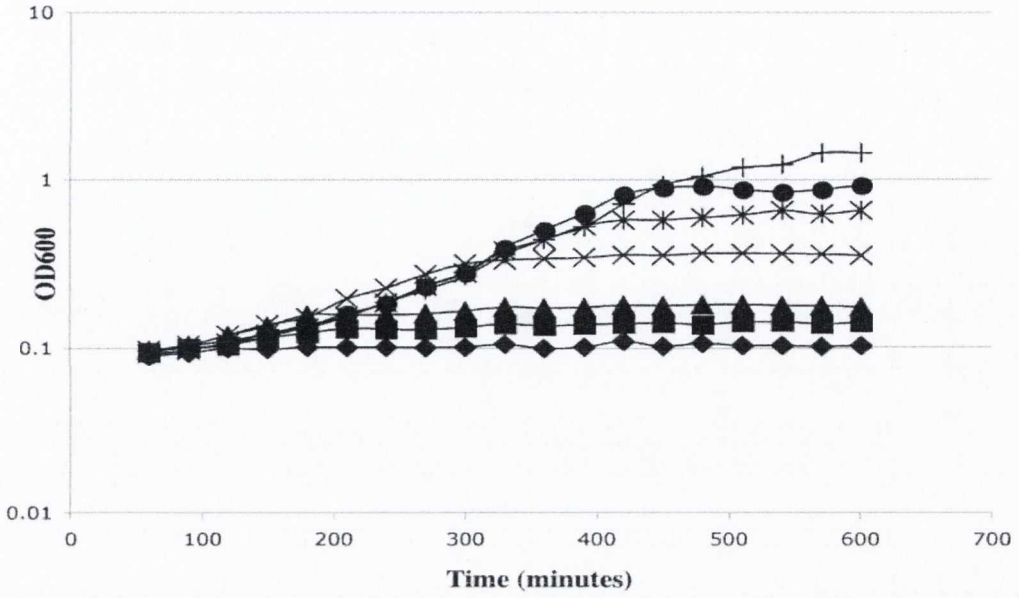


Figure 3.5 Growth and ß-galactosidase activity of BCJ363.1 (*trpC2 amyE*::pBCJ307.5 ($P_{T\text{-}box\ lysK}$-*lacZ* Cm^R) grown in Spizizen's minimal media. Open symbols represent growth. Closed symbols represent ß-galactosidase activity.

To assess whether the $P_{T\text{-}box\ lysK}$-*lacZ* fusion was sensitive to limitation for lysine, strain NF33.1 was grown in both lysine rich (100µg/ml lysine in Spizizen's media) and lysine depleted (20µg/ml lysine in Spizizen's media) conditions and the ß-galactosidase activity monitored (Figure 3.6B). Strain NF33.1 grown at 100µg/ml lysine produces 7-10 units of ß-galactosidase activity indicating a low level of expression. This level of activity remains constant throughout the time course. However, when the same strain is grown at 20µg/ml lysine we see an approximately 100 fold induction from 7-10 units to approximately 700 units of activity over 120 minutes from 240 to 360 minutes after inoculation. This induction of ß-galactosidase accumulation is coincident with cessation of growth as a result of lysine starvation. This indicates that the *B. cereus lysK* T-box regulatory element responds to a reduction in lysine concentration within the cell.

In order to further demonstrate that the *lysK* T-box mechanism responds to the level of $tRNA^{LYS}$ charging in the cell we created the strain BCJ367.1. In this strain the endogenous *lysS* gene of *B. subtilis* was put under the control of the IPTG inducible P_{spac} promoter (see strain construction). In order to reduce uninduced initiation of transcription by the P_{spac} promoter, the plasmid pMAP65 was transformed in. This plasmid is a multicopy plasmid containing the *lacI* gene which produces a protein that reduces transcription from the Pspac promoter in the absence of IPTG. This strain also contains the $P_{T\text{-}box\ lysK}$-*lacZ* fusion. Using the IPTG inducible system, the level of *lysS* expression can be altered, which should in turn affect the level of LysRS present in the cell and hence the amount of charged $tRNA^{LYS}$ in the cell. This strain was then grown at varying concentrations of IPTG and accumulation of ß-galactosidase was determined. Strain BCJ367.1 [*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R)] was grown at 1mM, 250µM and 100µM IPTG (Figure 3.7). In medium containing 1mM IPTG, growth of BCJ367.1 is similar to that of the wild type *B. subtilis* and the profile of ß-galactosidase activity is similar to that seen in BCJ363.1 (Figure 3.5). In medium containing 250µM IPTG the growth rate of BCJ367.1 was similar to that at 1mM IPTG (Figure 3.7). This IPTG concentration was experimentally evaluated to be the lowest IPTG concentration that would not affect the growth rate. Under these conditions ß-galactosidase activity was higher than at 1mM IPTG with approximately 300 units of *lacZ* expression signifying that the reduced expression of LysRS had in turn caused a reduction in $tRNA^{LYS}$ charging resulting in increased *lacZ* expression. BCJ367.1 grown at 100µM IPTG displayed a reduced growth rate thought to be due to reduced $tRNA^{LYS}$ charging. The level

(A)



(B)

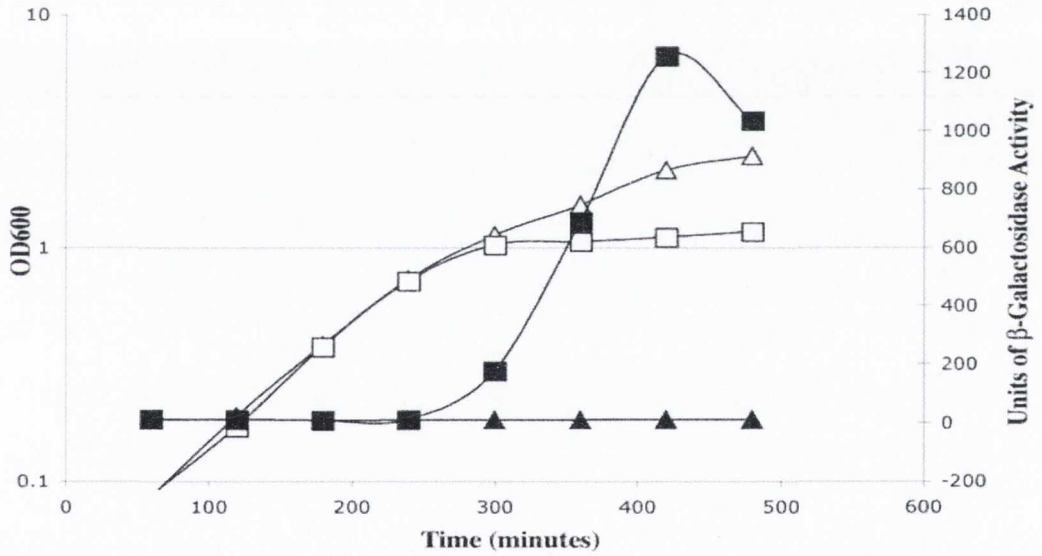
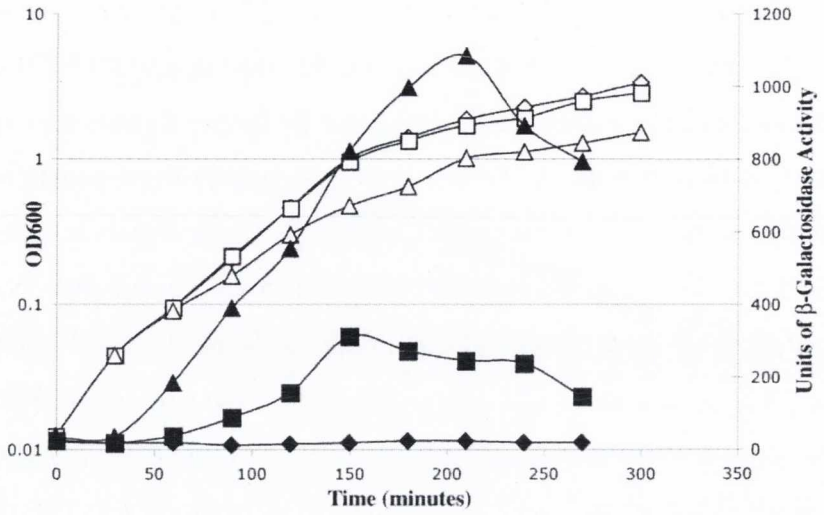


Figure 3.6 **(A)** Growth of the lysine auxotroph strain 1A765 in minimal media with varying lysine concentrations. Growth at different lysine concentrations is represented by the following symbols,100µg/ml lysine(+), 20µg/ml (●), 10µg/ml (✳), 5µg/ml (×), 1µg/ml (▲), 0.5. µg/ml (■), 0µg/ml (◆). **(B)** Growth and ß-galactosidase activity of strain NF33.1[*trpC2* lys *amyE*::pBCJ307.5($P_{T\text{-}box\ lysK\text{-}lacZ}$) Cm^R]. The symbols represent growth of NF33.1 at 100µg/ml lysine (△) and 20µg/ml lysine (□) and ß-galactosidase activity of NF33.1 at 100µg/ml (▲) and 20µg/ml lysine (■).

(A)



(B)

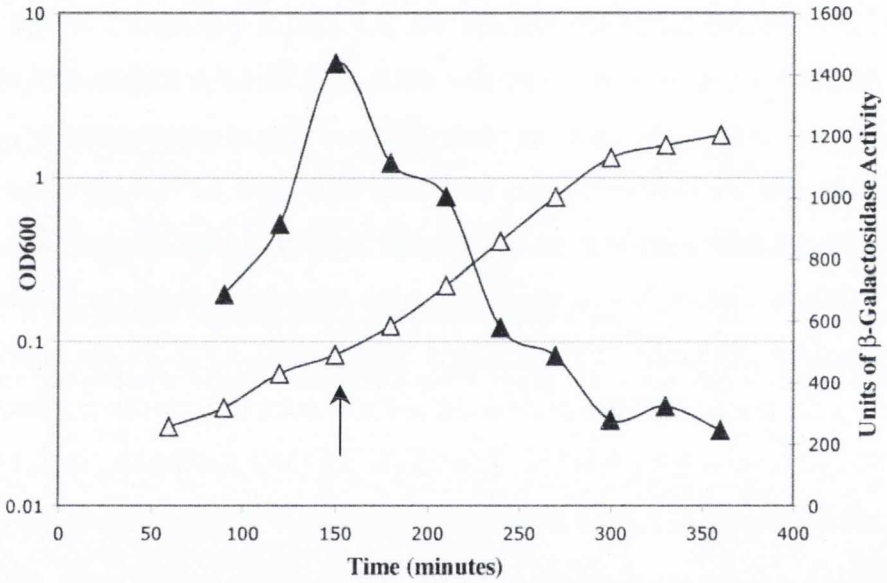


Figure 3.7 Growth and ß-galactosidase activity of BCJ367.1 [*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R)] in LB containing varying IPTG concentrations. (A) Growth and ß-galactosidase activity of BCJ367.1 in LB containing 1mM, 250µM IPTG and 100µM IPTG. Growth is represented by open symbols. ß-galactosidase activity is represented by closed symbols. Strains are as follows, BCJ367.1 1mM IPTG (◇, ◆), BCJ367.1 250µM IPTG (□, ■), BCJ367.1 100µM IPTG (△, ▲). (B) Growth and ß-galactosidase activity of BCJ367.1 in LB containing 100µM IPTG. The IPTG concentration was increased to 1mM at 150 minutes. (indicated by arrow). Growth is represented by open symbols. ß-galactosidase activity is represented by closed symbols.

of ß-galactosidase accumulation under these conditions reached approximately 1100 units at OD_{600} 1.0. These data indicate that reduced $tRNA^{LYS}$ charging results in a higher level of *lacZ* expression. This shows that the *B. cereus* 14579 *lysK* T-box regulatory element responds to depletion for charged $tRNA^{LYS}$.

We had observed that lowering the IPTG concentration in the media of cultures of BCJ367.1 resulted in a higher amount of ß-galactosidase accumulation. It was hypothesized that growth of BCJ367.1 at 250 µM IPTG would result in a reduced amount of LysRS in the cell than present in wild type *B. subtilis* cells, resulting in increased expression of the $P_{T\text{-box } lysK}$-*lacZ* fusion. Introduction of increased amounts of LysRS by addition of IPTG to the culture to a concentration of 1mM should then increase the level of LysRS and hence $tRNA^{LYS}$ charging thus reducing the level *lacZ* expression from the $P_{T\text{-box } lysK}$-*lacZ* fusion.

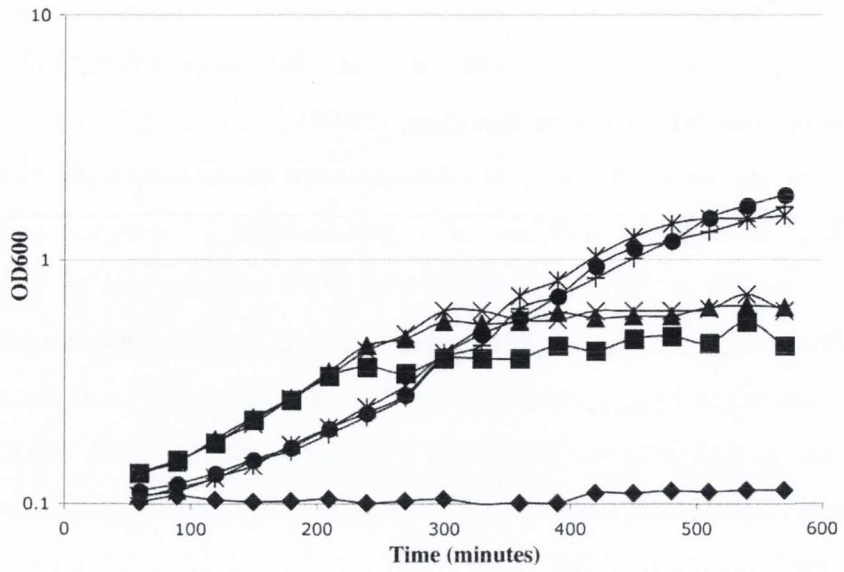
To test this we grew a culture of BCJ367.1 at 250µM IPTG and then spiked it by addition of further IPTG to a concentration of 1mM, at an OD_{600} of 0.1. Data for this experiment is presented in figure 3.7b. The results for this experiment show that in the initial time points of the growth curve when the IPTG concentration in the medium is at 250 µM there is an increase in ß-galactosidase activity. The IPTG concentration in the medium was then increased to 1mM. Following this there was an immediate decrease in the level of ß-galactosidase activity. This decrease was brought about by increased LysRS expression leading to increased charging of $tRNA^{LYS}$ (Figure 3.7). This data shows that the expression of the *B. cereus lysK* is sensitive to the level of $tRNA^{LYS}$ charging in the cell and that the degree of charging is likely to be sensed by a T-box anti-termination mechanism.

3.4 The *B. cereus lysK* T-box regulatory element does not respond to starvation for phenylalanine

An essential feature of the T-box antitermination mechanism is specificity i.e. each T-box element is sensitive only to their cognate tRNA. In order to ascertain whether this is the case for the *B. cereus lysK* T-box regulatory element, the plasmid pBCJ307.5 was integrated into the strain JH642 (*pheA2)* to generate strain NF37.1.

In order to establish the level of phenylalanine necessary to achieve phenylalanine limitation in this strain, NF37.1 was grown in minimal medium containing decreasing amounts of phenylalanine. Results are shown in figure 3.8. The data shows that as the level of phenylalanine in the culture is decreased, growth plateaus at a lower OD_{600}. We

(A)



(B)

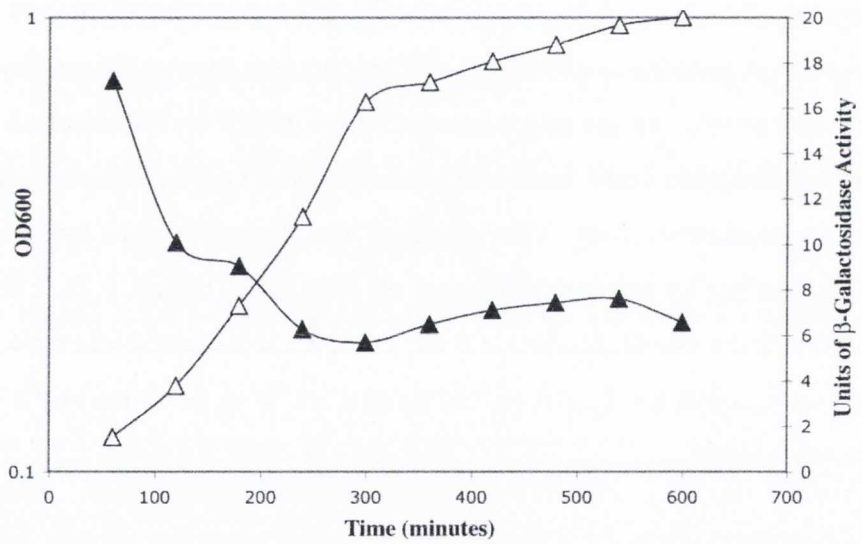


Figure 3.8 **(A)** Growth of the phenylalanine auxotroph strain NF37.1 in minimal media with varying phenylalanine concentrations. Growth at different phenylalanine concentrations is represented by the following symbols, 100µg/ml phenylalanine(┼), 20µg/ml (●), 10µg/ml (✳), 5µg/ml (✕), 3µg/ml (▲),2 µg/ml (■), 0µg/ml (◆). **(B)** Growth and ß-galactosidase activity of strain NF37.1[*pheA1 trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK\text{-}lacZ}$) Cm^R]. The open symbols represent growth of NF37.1 at 5µg/ml phenylalanine (△). The closed symbols represent ß-galactosidase activity of NF37.1 at 5µg/ml phenylalanine(▲).

concluded that a phenylalanine concentration of 5μg/ml yielding starvation at OD_{600} 1.0 was optimal for these experiments.

Strain NF37.1, containing the $P_{T\text{-}box\ lysK}$-*lacZ* fusion, grown under these phenylalanine limitation conditions showed no increase in ß-galactosidase activity (Figure 3.8). This shows that antitermination in the *B. cereus lysK* T-box regulatory element is not induced by starvation for phenylalanine.

3.5 Can the *B. cereus lysK* gene under the regulation of its own promoter and T-box regulatory element support growth of *B. subtilis*?

B. cereus encodes two *lysS* genes. One is a class II lysyl tRNA synthetase which charges $tRNA^{LYS}$ during exponential growth. The other is a class I synthetase which is expressed predominantly during stationary phase (Ataide., *et al,* 2005). The class I enzyme may act in concert with its class II counterpart to charge a novel tRNA species known as $tRNA^{Other}$ (Ataide., *et al,* 2005). The function of this tRNA is unknown. The *B. cereus* class I *lysK* was previously shown to support growth of *B. subtilis* when expressed from the strong ribosomal, P_{rpsD} promoter (Jester & Devine, unpublished work). However the level of enzyme produced in this strain is far greater than the endogenous expression signals (when used to express *Borrelia burgdorferi lysK* in *B. subtilis,* expression from the P_{rpsD} promoter produced amounts of enzyme 80 times greater than the endogenous *lysS* promoter of *B. subtilis*). Our analysis of T-box regulated AARS genes shows that transcriptional regulation of *lysS/K* genes by T-box regulatory mechanisms is a rare genetic arrangement. The data also shows that in strains containing a T-box regulated *lysS/K* gene, another *lysS/K* gene in encoded. We were curious as to whether a single *lysS/K* gene, whose expression is controlled by a T-box mechanism could support growth of *B. subtilis*. Here we sought to establish if the *lysK* promoter with its T-box regulatory element could produce enough *B. cereus* LysKRS to support growth of *B. subtilis*.

To achieve this the strain NF54.13 [*trpC2 amyE*::pNF48.44($P_{T\text{-}box}$-*lysK*) Δ*lysS* Kan^R] was constructed as outlined in the materials and methods. In this strain the *B. cereus lysK* promoter and T-box with the *lysK* structural gene were integrated in single copy at the amylase locus using plasmid pNF48.44. The endogenous *B. subtilis lysS* gene was then knocked out, again by double crossover, using plasmid pBCJ144.3.

Integration of the plasmid pNF48.44 into the amylase locus and integration of pBCJ144.3 at the endogenous *lysS* locus was confirmed by amplifying PCR products using primer pairs which bound the insert DNA, and also a region of the chromosome

just outside the inserted DNA. Primer pairings and predicted PCR products are shown in Figure 3.9. Confirmation of the integration of the pNF48.44 plasmid was achieved using the primer pairs AmyF and 36F yielding a 4.1kb PCR product and AmyR and 36R yielding a 3.1kb PCR product (Figure 3.9C). The AmyF and AmyR primers bind regions of *B. subtilis* DNA upstream and downstream respectively of the amylase gene. The 36F and 36R primers bind to the *B. cereus lysK* upstream leader region. Amplification of the AmyF-36F and AmyR-36R PCR products show that plasmid pNF48.44 successfully integrated into the amylase locus. Confirmation of the integration of the plasmid pBCJ144.3 into the *lysS* locus was achieved using the primer pairs Kan1 and 59F yielding a 1.1kb PCR product and Kan2 and 9R yielding a 980bp PCR product (Figure 3.9D). The Kan1 and Kan2 primers bind the 5' and 3' regions respectively of the kanamycin resistance gene on the pBCJ144.3 plasmid. Primers 59F and 9R bind regions upstream and downstream of the *lysS* gene respectively. Amplification of the Kan1-59F and Kan2-9R PCR products show that plasmid pBCJ144.3 successfully integrated into the *B. subtilis lysS* locus.

To further confirm that strain NF54.13 has the expected plasmid integration at the amylase locus we carried out a Southern blot. The probes and band predictions used are shown in Figure 3.10. The Southern blot showing the successful integration of pNF48.44 at the amylase locus was done by digesting chromosomal DNA from strain NF54.13 with ClaI and probing with a 610bp digoxygenin (DIG) labeled PCR product which binds the *lysK* gene. This probe would hybridize to both an 8.4kbp and a 1.1kb fragment if the *B. cereus lysK* gene was present. The data obtained from this analysis (Figure 3.10c) shows fragments corresponding to the band sizes described above in lanes containing DNA from strains NF52.9 and NF54.13. In addition to this bands corresponding to 2297bp and 1648bp were obtained for the positive control *B. cereus* 14579 representing the expected band sizes for a ClaI digest using the 610bp *lysK* probe (Figure 3.10c). The negative control lane containing *B. subtilis* 168 DNA did not produce a hybridization band (Figure 3.10c).

To confirm the successful integration of the pBCJ144.3 plasmid at the *lysS* locus we carried out another southern blot. Chromosomal DNA was digested with EcoRI and probed with a 346bp DIG labeled probe that would not produce a hybridization band if *lysS* was successfully removed but would produce a 2100bp band if *lysS* was present. The probes and band predictions for the Southern blots described above are illustrated in Figure 3.10d. The results show hybridization bands of approximately 2100bp in lanes 1 and 4 containing *B. subtilis* 168 and NF52.9 DNA respectively.

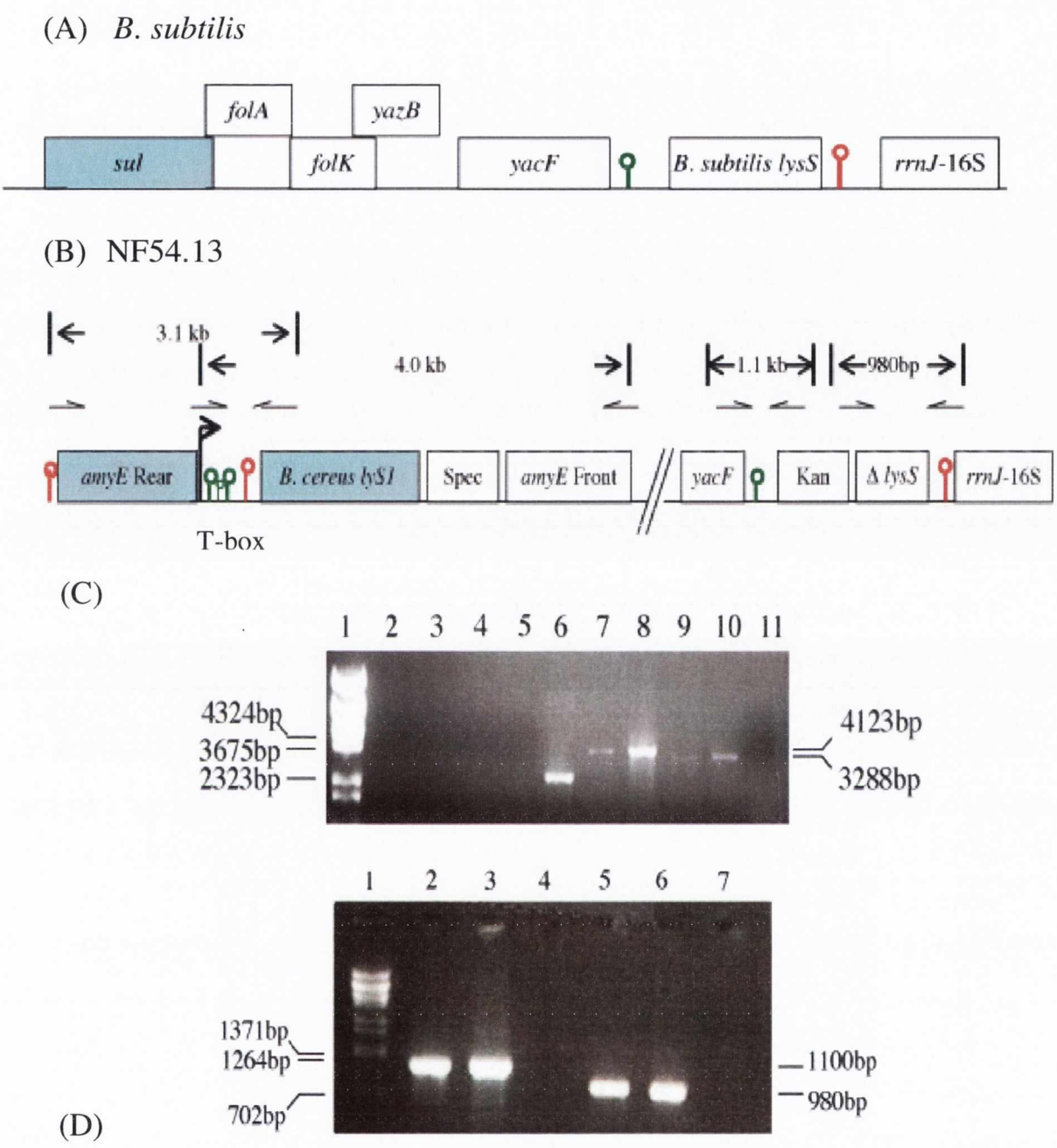
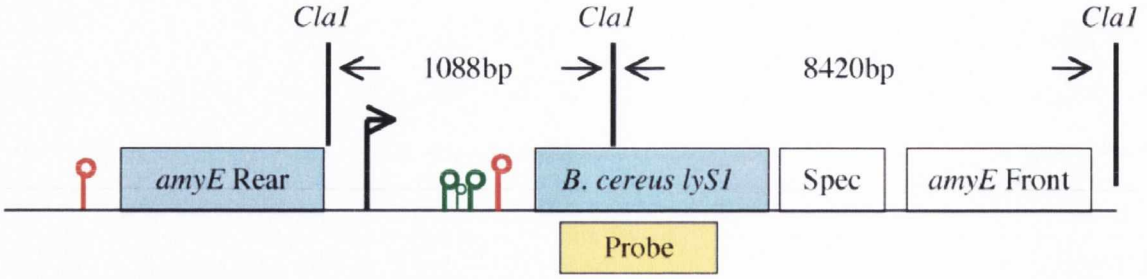
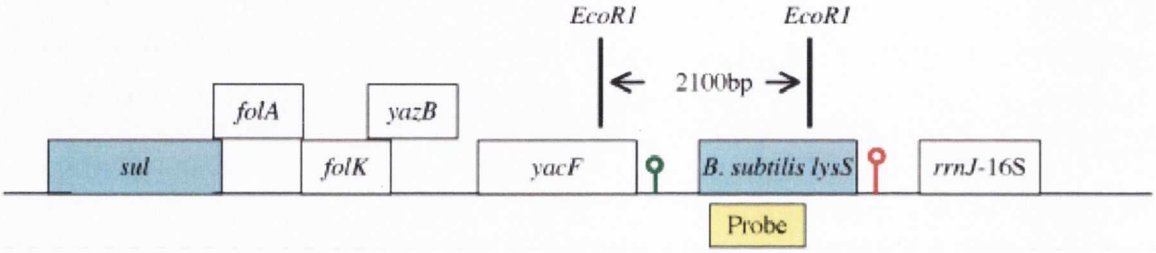


Figure 3.9 PCR mapping of NF54.13 [*trpC2 amyE*::pNF48.44($P_{T\text{-}box}$-*lysK*) Δ*lysS*Kan^R]. (A) wt. *B. subtilis lysS* locus map. (B) NF54.13 genetic map. (C) Agarose gel of PCR reactions using primers AmyF and 36F (4kb product), AmyR and 36R (3.1kb product). The samples loaded are lane (1) λ BstE2 ladder (8) NF54.13 (10) NF54.13 (11) Water negative control. (D) Agarose gel of PCR reactions using primers Kan1 and 59F (1.1kb product) and Kan2 and 9R (980bp product).). The samples loaded are lane (1) λ BstE2 ladder (2) NF54.13 (4) 168 (5) NF54.13 (7) 168. The green lollipops represent the stem loop structures of the T-box leader region. The red lollipops are transcription terminators

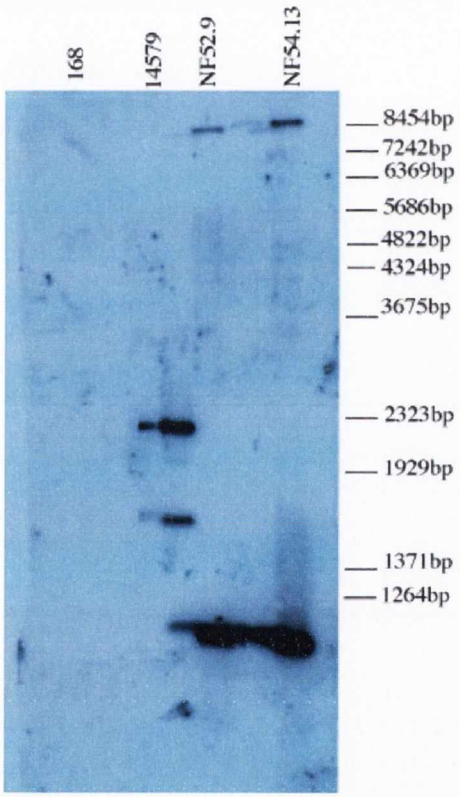
(A) NF54.13



(B) *B. subtilis* 168



(C)



(D)

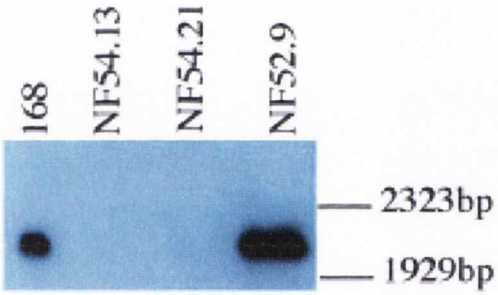


Figure 3.10 Southern blot analysis of NF54.13. (A) Genetic diagram of the NF54.13 amylase locus containing the *B. cereus lysK* insert. (B) Genetic diagram of the *B. subtilis lysS* locus. (C) Southern blot of ClaI digests probed with a *B. cereus lysK* specific probe. The strains loaded on the gel are from left to right *B. subtilis* 168, *B. cereus* 14579, NF52.9 and NF54.13. (D) Southern blot of EcoR1 digests probed with a *B. subtilis lysS2* specific probe. The strains loaded on the gel are indicated. The green lollipops represent the stem loop structures of the T-box leader region. The red lollipops are transcription terminators.

Lanes 2 and 3 containing DNA from strains NF54.13 and NF54.21 (this strain was another NF54 putative clone that was not used for further analysis) show no bands. This is consistent with a *lysS* knockout phenotype (Figure 3.8d).

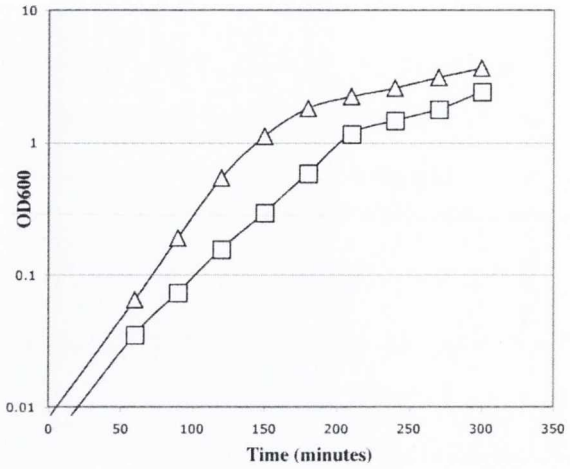
The PCR data combined with the Southern blot data show that the plasmid pNF48.44 was successfully integrated into the amylase locus of *B. subtilis*, thereby placing the *B. cereus lysK* gene under the transcriptional regulation of its own promoter and T-box regulatory element onto the chromosome of *B. subtilis*. The data also shows that the *B. subtilis lysS* gene was removed as no hybridization band was seen when probing for the *lysS* gene. As a functional lysyl tRNA synthetase is an essential requirement for growth this data indicates that the *B. cereus lysK* under the transcriptional regulation of its own promoter and T-box regulatory element can support growth of *B. subtilis*.

3.6 Phenotypic characterization of strain NF54.13

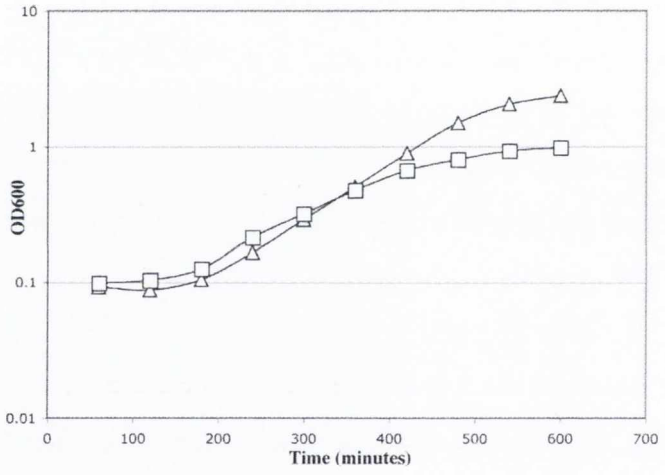
The *B. cereus lysK* under the regulation of its cognate transcriptional control mechanisms and T-box regulatory element can support growth of *B. subtilis*. We set out to assess if the strain NF54.13 had a phenotype. To establish this we compared growth of NF54.13 and wild-type *B. subtilis* in both LB and Spizizen's minimal media. Results are presented in figure 3.11.

Cultures were inoculated with strains NF54.13 and 168 to an OD_{600} of 0.01 for LB and 0.1 for minimal media. Generation times were calculated as described in the materials and methods. The data obtained from these experiments showed a reduction in the growth rate of NF54.13 relative to wild-type when grown in LB. The generation times were 31.2 minutes and 22 minutes respectively (Figure 3.11a). We were unable to culture NF54.13 in minimal media. Growth of strain NF54.13 in Spizizen's minimal media was only achieved by growing an overnight culture in LB, and then using this to inoculate minimal media. Under these conditions NF54.13 grew at a rate comparable to wild-type until an OD_{600} of approximately 0.5 (Figure 3.11b). After this point the growth rate decreased and eventually halted completely at an OD_{600} of approximately 1.0. We suspect that carryover of nutrients from the overnight culture grown in LB sustained the growth of NF54.13 until those nutrients were used up causing the cessation of growth. As stated earlier, *lysS* in the last gene in the folate operon in *B. subtilis*. It is possible that the integration of the plasmid pBCJ144.3 at the *lysS* locus results in instability of the folate operon mRNA transcript. This would confer a requirement for folate on the strain NF54.13 and could account for our inability to

(A)



(B)



(C)

Strain	LB	Minimal Media
B. subtilis	19.6 min.	76.1 min.
NF54.13	31.5 min.	109.8 min.*

* Strain NF54.13 grown in minimal media was supplemented with LB

Figure 3.11 Analysis of growth phenotypes of NF54.13. (A) Growth of NF54.13 [*trpC2 amyE*::pNF48.44($P_{T\text{-}box}$-*lysS*1) Δ*lysS*Kan^R] (□) vs 168 (△) in LB. (B) Growth of NF54.13 (□) in Spizizen's media inoculated from an LB overnight culture compared with *B. subtilis* 168 (△). (C) Comparison of generation times of NF54.13 and *B. subtilis* 168 in LB and minimal media.

culture this strain in minimal media. However, the plasmid pBCJ144.3 has previously been used in our lab to remove the class II *lysS* of *B. subtilis* in other gene replacement strains without resulting in a requirement for additional nutrients in minimal media. Thus, it is likely that our inability to culture NF54.13 in minimal media is due to the reduced charging efficiency of the *B. cereus* LysK enzyme compared to the B. subtilis LysS. These results show that while the *B. cereus lysK* gene under the transcriptional regulation of its own promoter and T-box element can sustain growth of *B. subtilis*, the fitness of this strain is negatively effected resulting in a slower growth rate.

3.7 The growth phenotype observed in strain NF54.13 may be due to reduced $tRNA^{LYS}$ charging

It has been reported that class I LysKRS enzymes are less catalytically efficient than their class II counterparts in terms of tRNA aminoacylation (Wang, *et al.*, 2006). We postulated that the reduction in catalytic efficiency in the class I LysKRS relative to the class II LysRS may be the cause of the reduced growth rate observed in strain NF54.13. A second explanation may be that the growth affect was due to a decreased amount of lysyl tRNA synthetase present in the cell in the NF54.13 strain due to a lower level of expression from the *lysK* promoter and T-box regulatory element. In order to see if the level of lysyl tRNA charging was affected we transformed the pBCJ307.5 plasmid into strain NF54.13 to create the strain NF206.1. This strain would then contain the $P_{T\text{-}box\ lysK}$-*lacZ* fusion. With this strain we could compare the ß-galactosidase activity with that of strain BCJ363.1, which has the same construct but in a wild-type *B. subtilis* background containing the class II LysRS. This comparison would provide an indication of differences in $tRNA^{LYS}$ charging through differences in *lacZ* expression. For example, a higher level of ß-galactosidase activity in strain NF206.1 relative to strain BCJ363.1 would indicate a reduction in the level of $tRNA^{LYS}$ charging due to the reduced level of catalytic efficiency of the class I LysKRS and/or reduced level of lysyl tRNA synthetase expression due to weaker expression signals from the *B. cereus lysK* promoter and T-box regulatory element. Experiments carried out on strain NF206.1 showed a level of ß-galactosidase activity of approximately 250-300 units during exponential growth (Figure 3.12). This represents a 20 fold increase in *lacZ* expression relative to the activity in strain BCJ363.1 grown under the same conditions. This indicates a reduction in $tRNA^{LYS}$ charging as a result of the replacement of the endogenous *lysS* gene with the *B. cereus lysK*.

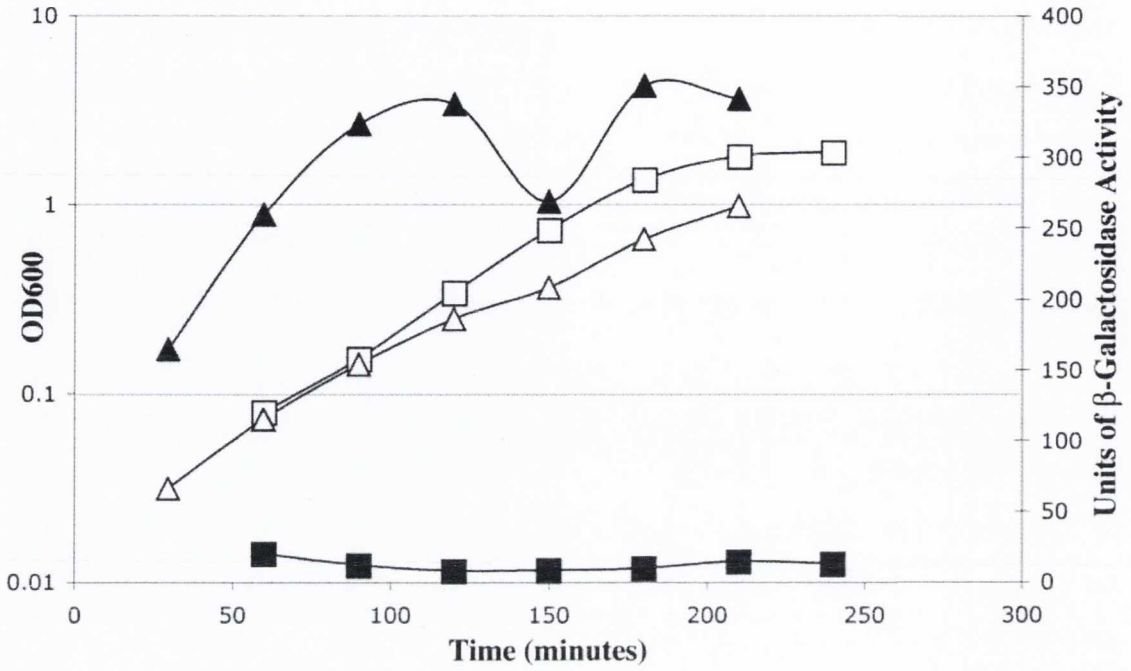


Figure 3.12 Comparison of growth and ß-galactosidase activity of strain NF206.1 and strain BCJ363.1. Growth and ß-galactosidase activity of NF206.1 [*trpC2 amyE*::pNF48.44($P_{T\text{-}box}$-*lysK*) Δ*lysS*Kan^R pBCJ307.5($P_{T\text{-box }lysK}$-*lacZ*) Cm^R] is represented by open and closed triangles respectively (△, ▲). Growth and ß-galactosidase activity of strain BCJ363.1 [*trpC2* pBCJ307.5($P_{T\text{-box }lysK}$-*lacZ*)] is represented by open and closed squares respectively (□, ■).

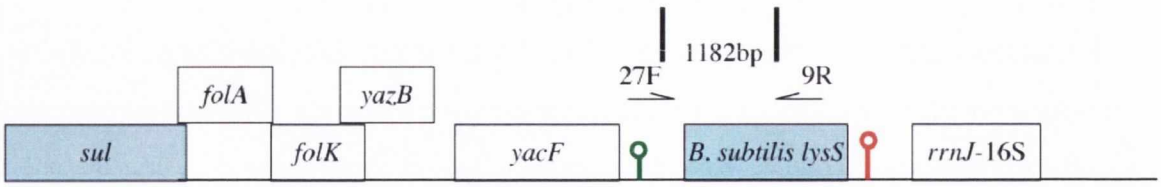
3.8 A *B. subtilis* strain expressing its endogenous *lysS* under the transcriptional regulation of the *B. cereus lysK* T-box mechanism is viable

The occurrence of a T-box regulated class II lysyl tRNA synthetase is an extremely rare phenomenon. This genetic arrangement does not exist in the *Bacilli*. Therefore we sought to establish whether the endogenous class II *lysS* from *B. subtilis* when expressed using the promoter and T-box regulatory element from the *B. cereus lysK* gene could support growth of *B. subtilis*.

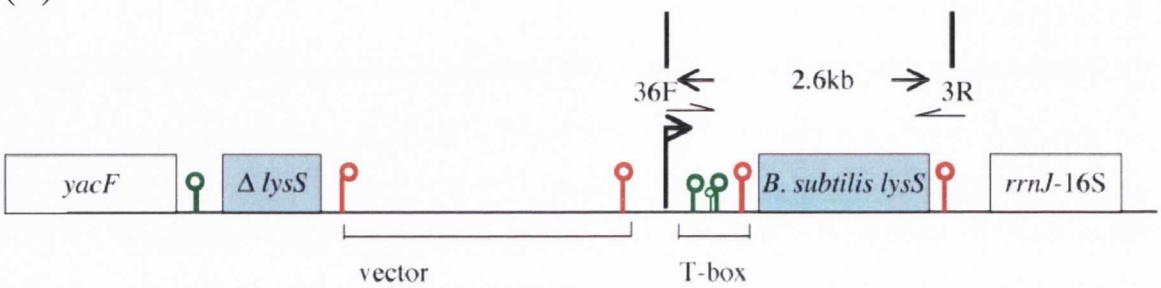
In order to investigate this the strain NF113.5 was constructed (*trpC2 lysS*::pNF112.5 ($P_{T\text{-}box\ lysK}$-*lysS* Cm^R). This strain was made by the integration of plasmid pNF112.5 by a single crossover into the *B. subtilis lysS* locus. This had the effect of separating the *lysS* gene from its endogenous promoter and placing it under the transcriptional control of the *B. cereus lysK* promoter and T-box mechanism. Confirmation of the integration of plasmid pNF112.5 at the *lysS* locus was achieved by amplifying PCR products using a primer pair which bound the *B. cereus lysK* T-box regulatory element sequence, and also a region of the chromosome just outside the inserted DNA (36F and 9R) (Figure 3.13). A negative PCR using primers which bind sites disrupted by pNF112.5 integration were also used (27F and 3R). Primer pairings and predicted PCR products are shown in Figures 3.13a and 3.13b. The data shows that a 2.6kb band was obtained from PCR reactions using NF113.5 chromosomal DNA and the 36F and 9R primers, consistent with successful integration of plasmid pNF112.5 at the *lysS* locus (Figure 3.13d). In addition, a PCR reaction carried out using the 27F and 3R primer pair which bind both the *lysS* gene and the region upstream of *lysS*, failed to produce a product from NF113.5 chromosomal DNA, but produced a 1182bp band when *B. subtilis* 168 DNA was used in the PCR reaction. This indicates a successful integration of plasmid pNF112.5 at the *lysS* locus resulting in regulation of expression of *B. subtilis lysS* by the *B. cereus lysK* promoter and T-box regulatory element (Figure 3.13c).

To further confirm the integration of plasmid pNF112.5 we carried out a Southern blot (Figure 3.14a and 3.14b). This was done by digesting chromosomal DNA with EcoR1 and probing with a 346bp probe specific for the *B. subtilis lysS* gene. If pNF112.5 had integrated successfully at the *lysS* locus, two bands would be obtained at 1512bp and 5714 bp. An uninterrupted *lysS* gene would produce a single band at 2100bp. The data obtained from this analysis shows fragments corresponding to the band sizes 1512bp and 5714bp in lanes containing NF113.5 and NF113.6 (NF113 putative clone not used for further analysis) (Figure 3.14c). Negative controls containing *B. cereus* 14579 and strain NF54.13

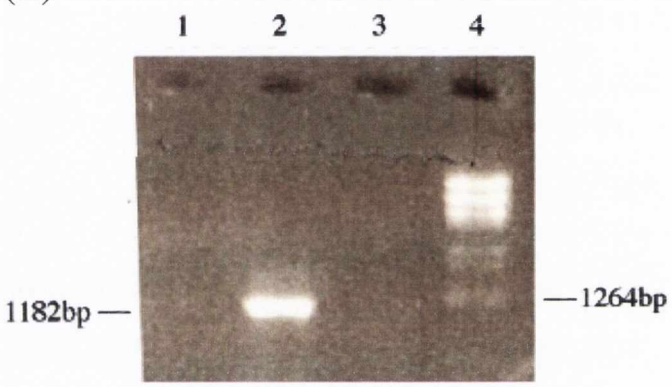
(A)



(B)



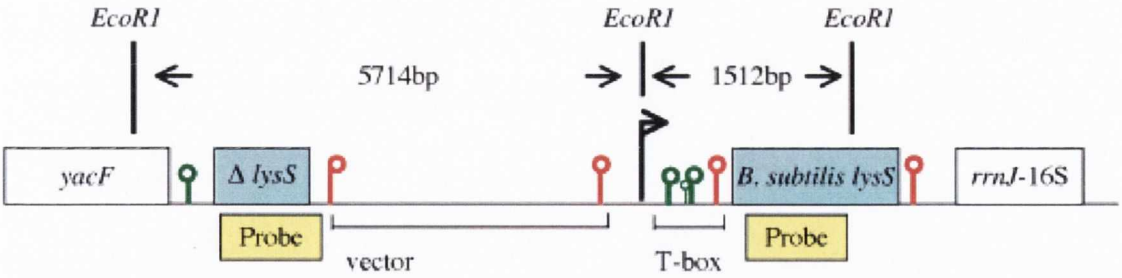
(C)



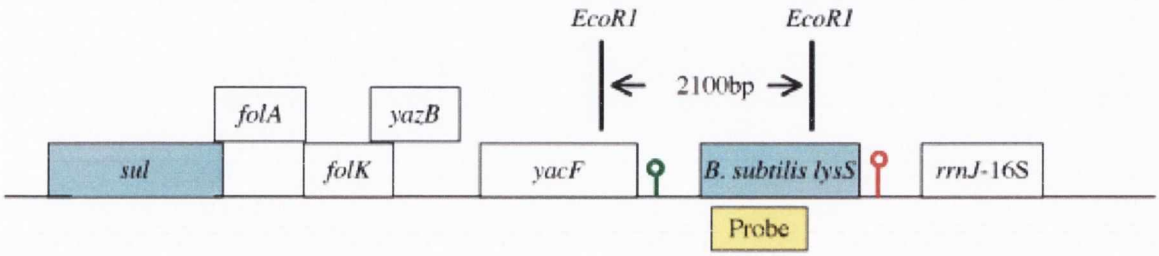
(D)

Figure 3.13 PCR mapping of NF113.5 [*trpC2 lysS*::pNF112.5 ($P_{T\text{-}box\text{-}lysS}$*lysS* Cm^R)]. (A) Genetic map of wt *B. subtilis lysS* locus. (B) Genetic map of NF113.5 *lysS* locus. (C) Agarose gel of PCR reactions using primers 27F and 3R. The samples loaded are (1) NF113.5 (2) 168 (3) water negative control (4) λ BstE2 ladder. (D) Agarose gel of PCR reactions using primers 36F and 9R. The samples loaded are (1) λ BstE2 ladder (2) NF113.5 (3) NF113.6 (4) water negative control

(A) NF113.5



(B) *B. subtilis* 168



(C)

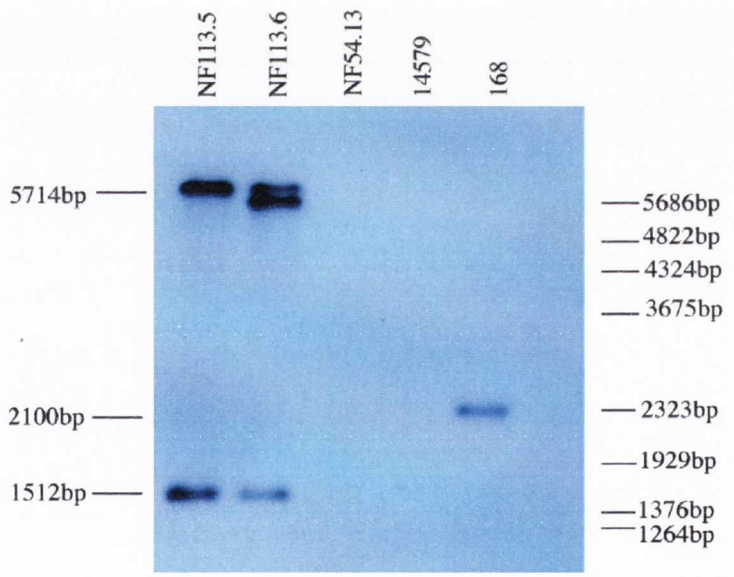


Figure 3.14 Southern blot analysis of NF113.5. (A) Genetic diagram of the NF113.5 *lysS* locus. (B) Genetic diagram of the wt. *B. subtilis lysS* locus. (C) Southern blot of EcoR1 digests probed with a *B. subtilis lysS* specific probe. The strains loaded on the gel are from left to right, NF113.5, NF113.6, NF54.13, *B. cereus* 14579 and *B. subtilis* 168. Band sizes of the molecular weight markers are shown on the right. The green lollipops represent the stem loop structures of the T-box leader region. The red lollipops are transcription terminators. Band sizes of the molecular weight markers λ BstEII are shown on the right

DNA show no hybridization bands consistent with the absence of the *lysS* gene in these strains. The lane containing *B. subtilis* 168 DNA shows a hybridization band at 2100bp.

The combined PCR and Southern blot analysis confirm that strain NF113.5 contains a *lysS* gene transcriptionally regulated by the promoter and T-box regulatory element from the *B. cereus lysK* gene. This results show that a *B. subtilis* strain with a T-box regulated class II lysyl tRNA synthetase is viable.

3.9 Phenotypic characterization of strain NF113.5

Having shown that the strain NF113.5 containing a T-box regulated class II lysyl tRNA synthetase was viable we sought to establish whether it had any unusual phenotypes. To do this we compared the growth rate of strain NF113.5 to *B. subtilis* 168 in both rich and minimal media. Results are presented in figure 3.15a and 3.15b.

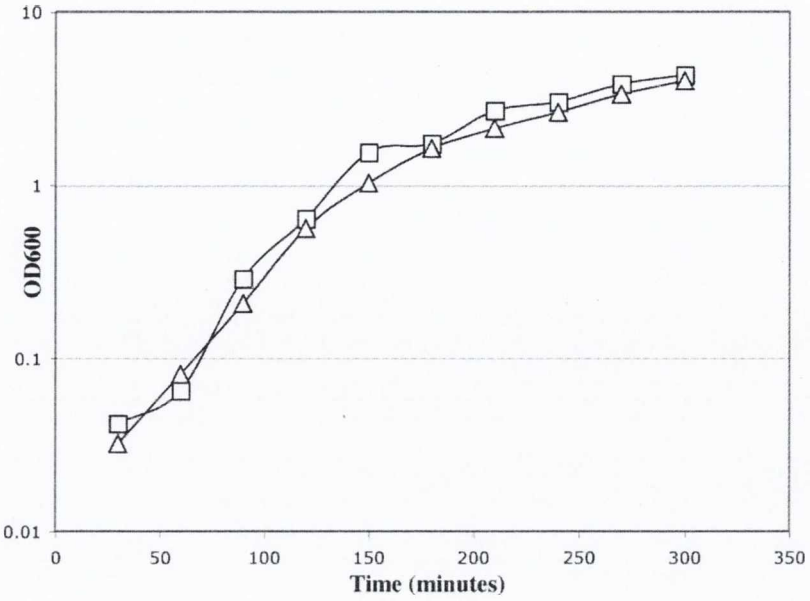
Cultures were set up as described for the phenotypic characterization of strain NF54.13. The growth rates of strain NF113.5 and *B. subtilis* 168 showed no difference in either LB or minimal media. These results show that the promoter and T-box regulatory element of the *B. cereus lysK* gene are sufficient to produce enough LysRS for wild-type growth of *B. subtilis*.

3.10 Does the $P_{T\text{-}box\ lysK}$-*lysS* genetic arrangement have an effect on $tRNA^{LYS}$ charging in *B. subtilis*

It was previously shown that in the strain NF54.13, containing *the B. cereus lysK* gene transcriptionally regulated by its own promoter and T-box regulatory element, there was a reduction in the level of $tRNA^{LYS}$ charging relative to wild-type *B. subtilis*. We sought to establish whether regulation of the endogenous *B. subtilis lysS* by a T-box mechanism would result in a change in the level of $tRNA^{LYS}$ charging.

In order to investigate this we made the strain NF205.1. This strain contains a *lysS* gene regulated by the promoter and T-box regulatory element from *B. cereus lysK* and the *B. cereus* ($P_{T\text{-}box\ lysK}$-*lacZ*) fusion at the amylase locus. Using this strain we could assess any changes in tRNA charging through differences in the ß-galactosidase activity relative to BCJ363.1 which contains the same *lacZ* fusion in a wild-type *B. subtilis* background. Data from this analysis is presented in figure 3.16. In both strains a low level of ß-galactosidase

(A)



(B)

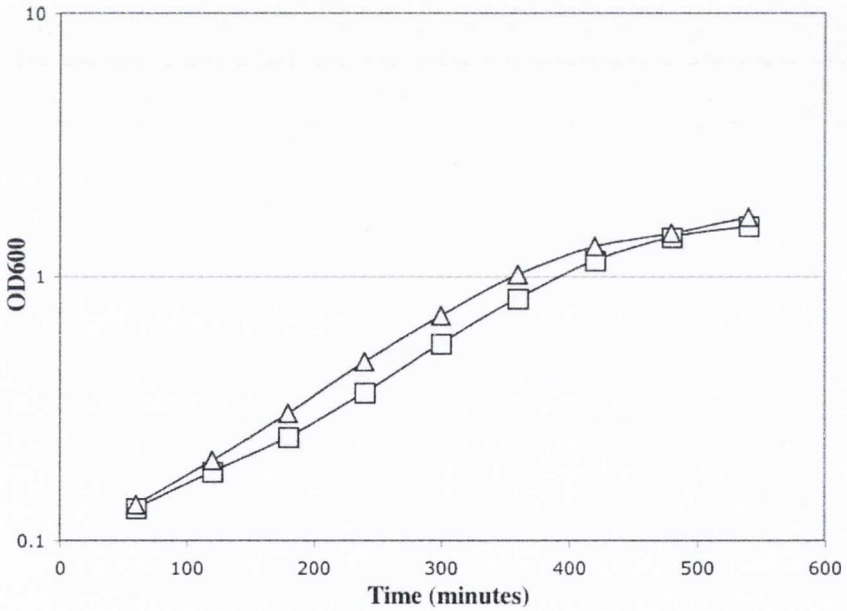


Figure 3.15 Analysis of growth phenotypes of NF113.5 in LB and minimal media. (A) Growth of NF113.5 [*trpC2 lysS*::pNF112.5 ($P_{T\text{-box } lysK}$*lysS* Cm^R)] (△) vs. *B. subtilis* 168 (□) in LB. (B) Growth of NF113.5 (△)vs. *B. subtilis* 168 (□) in Spizizen's minimal media.

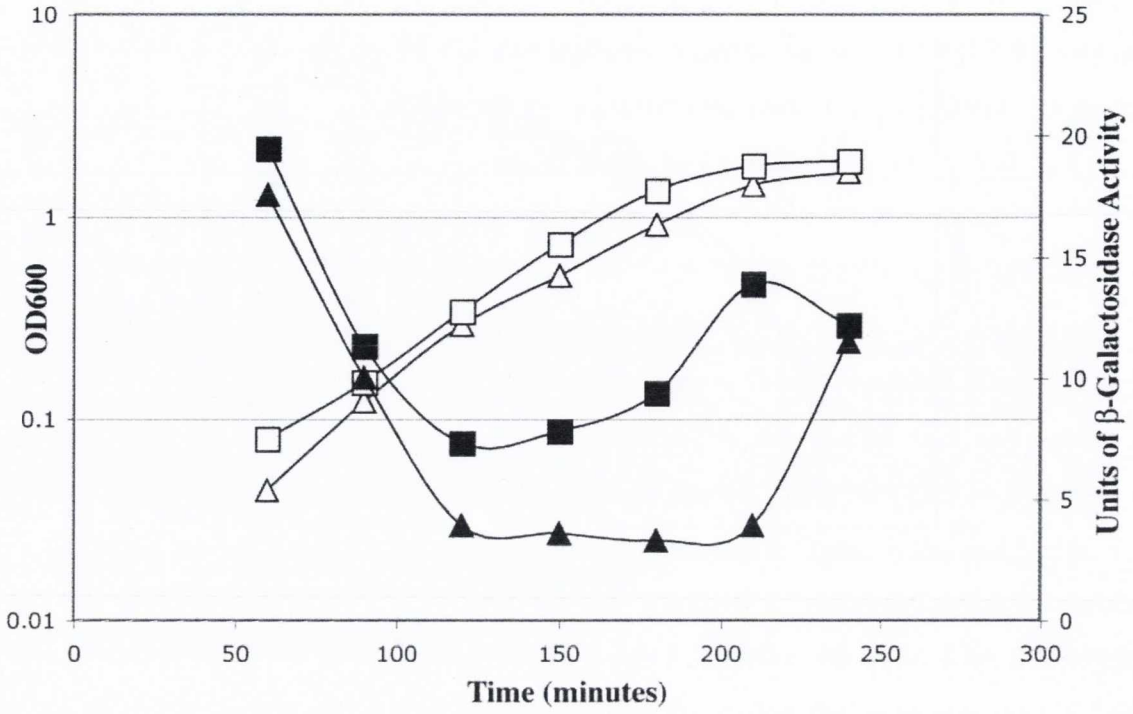


Figure 3.16 Analysis of growth and ß-galactosidase activity of strain NF205.1. Growth and ß-galactosidase activity of NF205.1 [*trpC2 lysS*::pNF112.5 ($P_{T\text{-}box\ lysK}$*lysS* Cm^R) pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R] is represented by open and closed triangles respectively (△,▲). Growth and ß-galactosidase activity of strain BCJ363.1 1 [*trpC2* pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*)] is represented by open and closed squares respectively (□, ■).

activity ranging from 5-10 units is observed during exponential growth (Figure 3.16). This result would appear to indicate that both methods of regulation result in a similar amount of lysyl tRNA synthetase being produced and therefore similar levels of $tRNA^{LYS}$ charging. This in turn indicates that the reduced growth rate seen in strain NF54.13 (Figure 3.11) is due to the reduced catalytic efficiency of the LysKRS enzyme relative to the LysRS enzyme and not due to lower expression signals from the *B. cereus lysK* promoter and T-box regulatory element relative to those of *B. subtilis lysS.*

3.11 Why is lysyl tRNA synthetase not T-box regulated in *B. subtilis*?

Our previous data shows that a *B. subilis* strain with a class II *lysS* regulated by a T-box anti-termination mechanism is viable and grows normally. Why then in *B. subtilis*, where the T-box mechanism is so commonly used for tRNA synthetases, is this method of regulation not used to control *lysS* expression?

An examination of T-box regulated tRNA synthetase genes in *B. subtilis* showed that those synthetases not regulated in this fashion tend to charge tRNA's with amino acids that are found in mixed codon boxes (Table 3.2). The tRNA synthetases in *B. subtilis* that are not T-box regulated are methionyl, prolyl, glutamyl, asparaginyl and lysyl tRNA synthetase. All these tRNA synthetases charge tRNAs with amino acids found in mixed codon boxes with the exception of prolyl tRNA synthetase. Therefore the correlation is a good one but has that exception.

The genes encoding lysyl and asparaginyl tRNA synthetase are not T-box regulated in *B. subtilis*. Given that both lysine and asparagine occupy the same mixed codon box we hypothesized that the specifier codon of the *lysK* T-box mechanism might be unable to discriminate between the anticodons of asparaginyl (UUG) and lysyl tRNA (UUU) as they differ only at the third or "wobble" position. The different wobble bases of $tRNA^{LYS}$ and $tRNA^{ASN}$ are important for discrimination of these two species on the ribosome during translation. However, in a non-ribosomal context, as in tRNA interactions with T-box regulatory elements, these differences may not be sufficient to allow successful discrimination between the two species by the specifier codon. A prediction of this model is that the *lysK* T-box mechanism of *B. cereus* would be sensitive to a reduction in charging of both $tRNA^{LYS}$ and $tRNA^{ASN}$. We also noted the possibility that the *lysK* T-box regulatory element may have evolved from a T-box element specific for asparagine. In theory a single mutation at the specifier codon may be enough to switch specificity from asparagine to lysine. An analysis of sequences upstream of genes related to asparagine biosynthesis and $tRNA^{ASN}$ charging in *B.*

cereus showed that there are some regions of sequence homology between the T-box regulatory element of the *asnA* gene and *lysK* (Figure 3.17). The sequence homology is not extensive enough to suggest with great confidence that the *lysK* T-box regulatory element evolved that of *asnA*. Nonetheless it remains a possibility. A further indication of possible ability of $tRNA^{ASN}$ to interact with the *lysK* T-box regulatory element comes from the observation that $tRNA^{LYS}$ and $tRNA^{ASN}$ share significant sequence homology in the single stranded

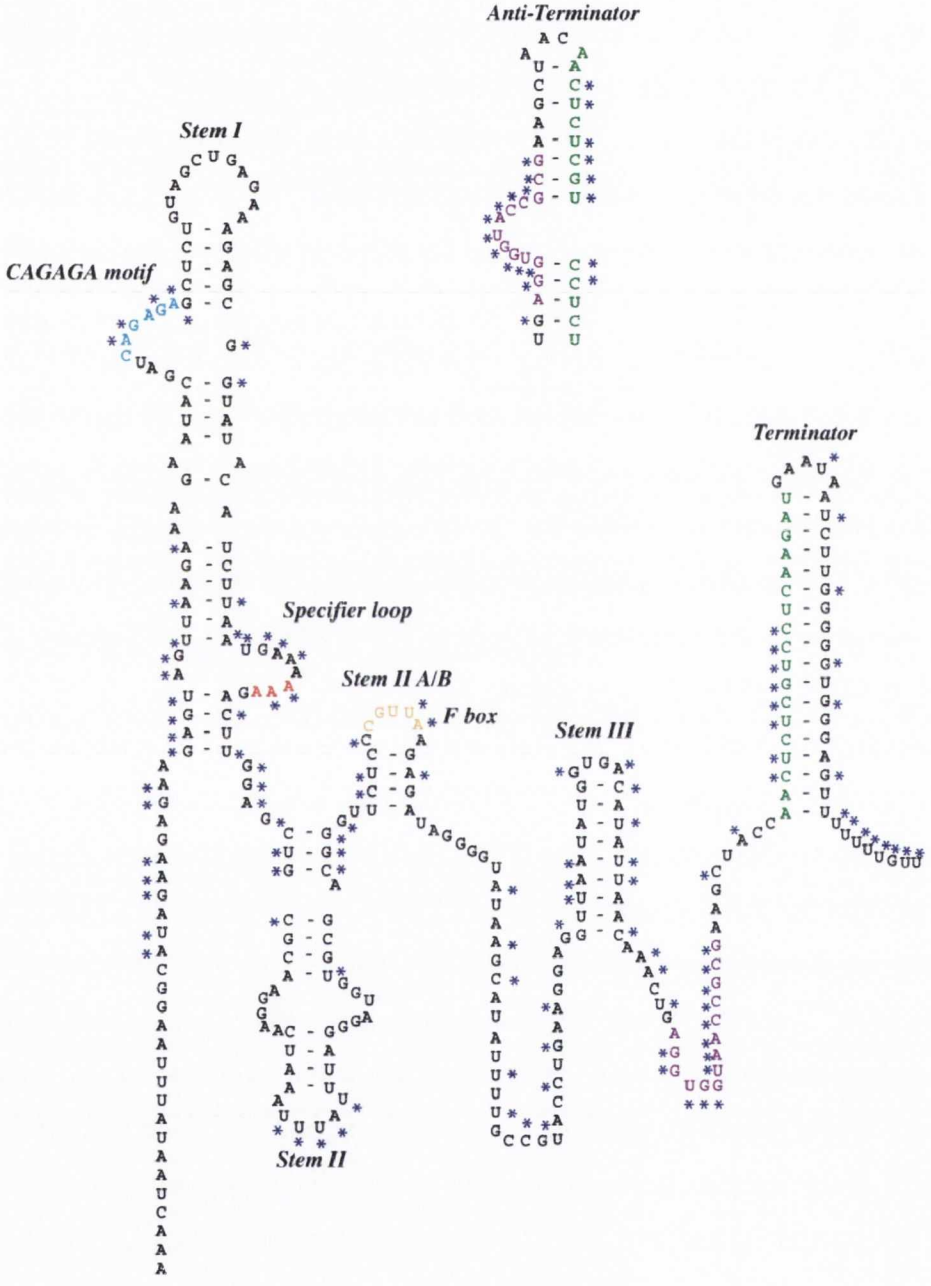


Figure 3.17 Proposed secondary strucuture of the *lysK* T-box regulatory element of *B. cereus*with possible homology with the T-box regulatory element of *asnA* from *B. cereus* highlighted. The specifier codon (AAA) is indicated in red. Important motifs and structural regions are indicated. The T-box sequence is indicated in purple. Nucleotides flanked by asterisks indicate regions of sequence homology with the T-box regulatory element upstream of the *asnA* gene in *B. cereus*. Adapted from Ataide., *et al,* (2005).

1st Position	2nd Position: T		2nd Position: C		2nd Position: A		2nd Position: G	
T	TTT	Phe	TCT	Ser	TAT	Tyr	TGT	Cys
	TTC	Phe	TCC	Ser	TAC	Tyr	TGC	Cys
	TTA	Leu	TCA	Ser	TAA	Stop	TGA	Stop
	TTG	Leu	TCG	Ser	TAG	Stop	TGG	Trp
C	CTT	Leu	**CCT**	**Pro**	CAT	His	**CGT**	**Arg**
	CTC	Leu	**CCC**	**Pro**	CAC	His	**CGC**	**Arg**
	CTA	Leu	**CCA**	**Pro**	**CAA**	**Gln**	**CGA**	**Arg**
	CTG	Leu	**CCG**	**Pro**	**CAG**	**Gln**	**CGG**	**Arg**
A	ATT	Ile	ACT	Thr	**AAT**	**Asn**	AGT	Ser
	ATC	Ile	ACC	Thr	**AAC**	**Asn**	AGC	Ser
	ATA	Ile	ACA	Thr	**AAA**	**Lys**	**AGA**	**Arg**
	ATG	**Met**	ACG	Thr	**AAG**	**Lys**	**AGG**	**Arg**
G	GTT	Val	GCT	Ala	**GAT**	**Asp**	GGT	Gly
	GTC	Val	GCC	Ala	**GAC**	**Asp**	GGC	Gly
	GTA	Val	GCA	Ala	**GAA**	**Glu**	GGA	Gly
	GTG	Val	GCG	Ala	**GAT**	**Glu**	GGG	Gly

Table 3.2 Codon table highlighting codons for amino acids whose cognate tRNA synthetases are T-box regulated in *B. subtilis* (blue). Codons for amino acids whose cognate tRNA synthetases are not T-box regulated are in black.

loop regions of the D loop and TψC loop (figure 3.18). These regions have been shown to be important in tRNA-T-box leader region interactions [Van de Guchte, *et al.*, (2001), Yousef, *et al.*, (2003)].

In order to test this hypothesis it was required that we deplete *B. subtilis* for charged $tRNA^{ASN}$ and assess its effect on the ß-galactosidase activity of the $P_{T\text{-}box\ lysK}$-*lacZ* fusion. An analysis of strains from the Bacillus Genetic Stock Centre (www.bgsc.org) showed no asparagine auxotroph was available. Limitation for asparagine in the medium was therefore ruled out as a strategy to deplete for charged $tRNA^{ASN}$. Instead we set out to make a strain with an IPTG inducible *asnS* gene. Using such a strain we could lower the level of *asnS* expression which is consistent with lower AsnRS levels and hence lower $tRNA^{ASN}$ charging. To this end we created the strain NF53.4 (*trpC2* (P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R). In this strain expression of the *asnS* gene is IPTG inducible. In order to improve the tightness of regulation by the P_{spac} promoter the non-integrating plasmid pMAP65 was transformed in. This plasmid is a multicopy plasmid that contains the *lacI* gene that reduces transcription from the P_{spac} promoter in the absence of IPTG. The plasmid pBCJ307.5 was then integrated by double crossover into the amylase locus to create the strain NF60.6. To confirm that strain NF60.6 was IPTG dependant we grew a culture of NF60.6 in LB until the culture was turbid. We then made serial dilutions of the culture and plated these out on LB plates with and without IPTG. The results are presented in figure 3.19. The data shows that in the absence of IPTG little or no growth of NF60.6 is observed. This shows that the *asnS* gene of strain NF60.6 is IPTG dependant.

3.12 The *B. cereus lysK* T-box mechanism responds to a reduction in the level of *asnS* expression

To test the hypothesis that $tRNA^{ASN}$ can interact with the specifier codon of the *B. cereus lysK* T-box regulatory element due to the specifier codons inability to discriminate between the anticodons of $tRNA^{ASN}$ and $tRNA^{LYS}$, we set out to lower the levels of charged $tRNA^{ASN}$ and observe the effect on the $P_{T\text{-}box\ lysK}$-*lacZ* fusion.

The strain NF60.6 containing an IPTG dependant *asnS* gene and the $P_{T\text{-}box\ lysK}$-*lacZ* fusion was grown in LB at IPTG concentrations of 1mM, 250µM and 100µM, and the ß-galactosidase activity examined (Figure 3.20a). At 1mM IPTG the ß-galactosidase activity increased from 20 to 80 units through the growth cycle. As the level of IPTG in the cultures was decreased, the level of *lacZ* expression increased. At 250µM IPTG the level of expression increased from 50 to 500 units of activity through the timecourse.

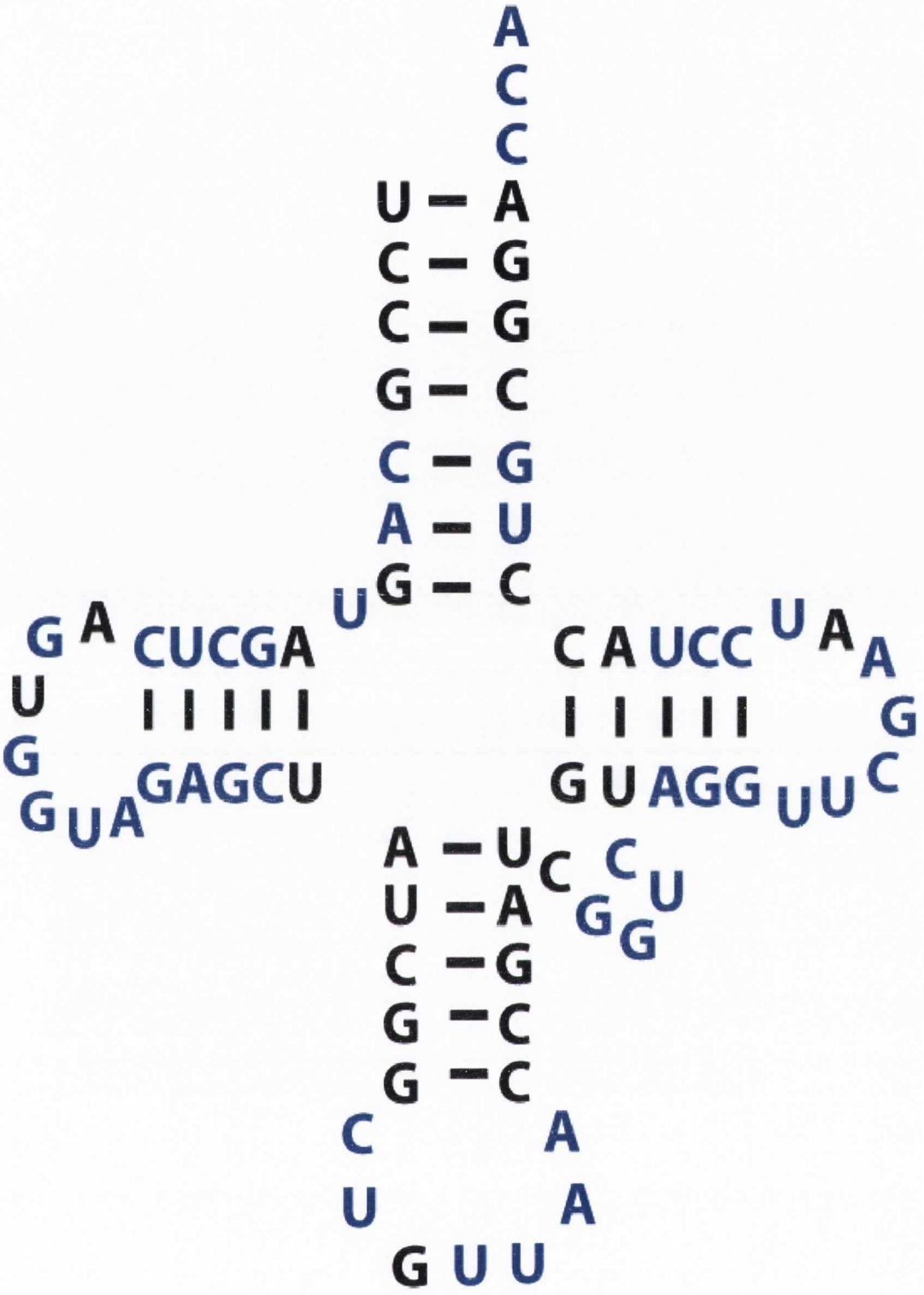


Figure 3.18 Secondary structure of $tRNA^{ASN}$ from *B. subtilis*. Regions of homology between $tRNA^{ASN}$ and $tRNA^{LYS}$ are highlighted in blue.

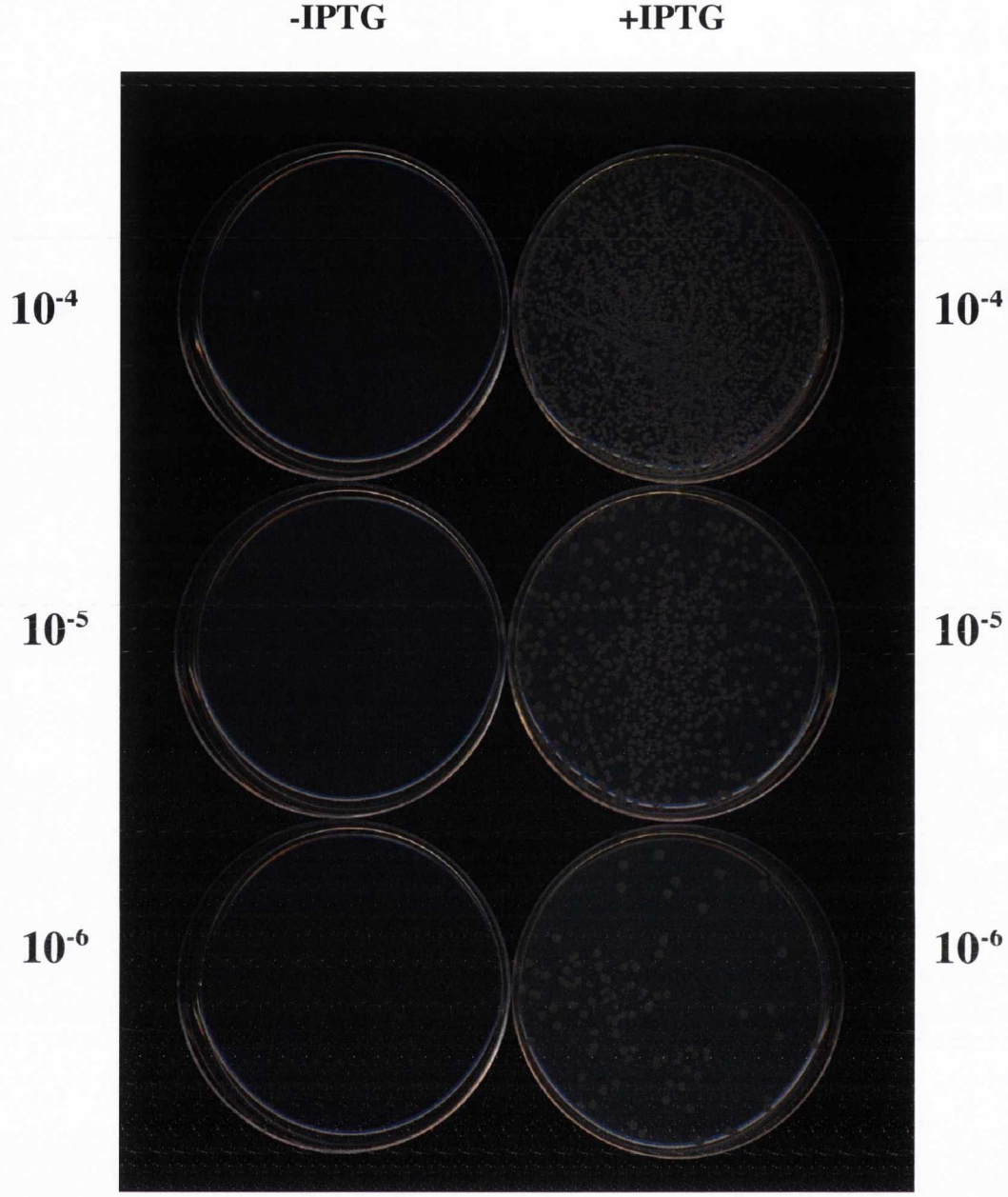


Figure 3.19 Demonstration of the IPTG dependance of the *asnS* gene of NF60.6 [*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R]. NF60.6 was grown in LB until the culture was turbid and 10 fold dilutions of this culture were lawned on LB plates with and without IPTG. Plates containing IPTG are shown on the right hand side. Plates without IPTG are shown on the left hand side. The dilution for each plate is indicated.

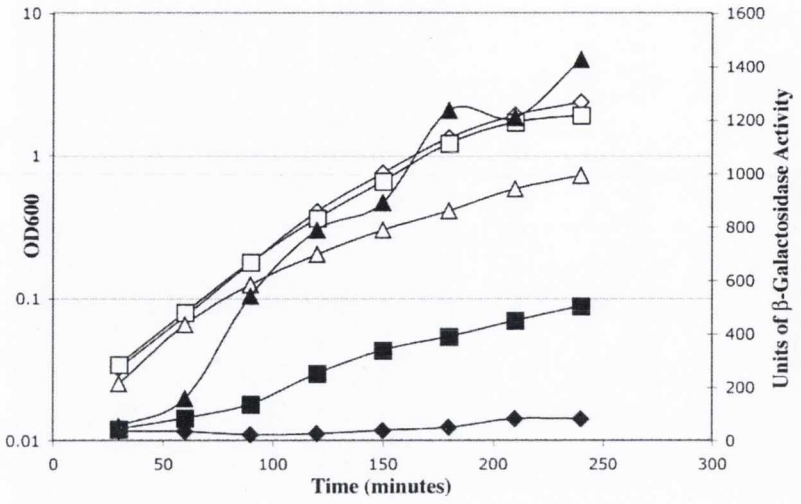
At 100μM IPTG the level of ß-galactosidase activity increased further from 50 to 1500 units. This result indicates that the *B. cereus lysK* T-box responds to depletion for charged $tRNA^{ASN}$. In addition to the experiments described above we carried out an IPTG spiking experiment (Figure 3.20b). In this experiment the strain NF60.6 was grown at an initial IPTG concentration of 100μM to an OD_{600} of 0.1 at which point the concentration was increased to 1mM IPTG. Growth of NF60.6 at 100μM IPTG causes a reduced growth rate relative to that observed in NF60.6 cultures containing 250μM IPTG or 1mM IPTG. It can be seen in figure 3.20b that following addition of IPTG to a concentration of 1mM, at 120 minutes, there is an increase in the growth rate. The data shows a steady increase in ß-galactosidase activity until an OD_{600} of 0.15. At this point we see the growth rate increase and the ß-galactosidase activity begins to decrease consistent with an increase in $tRNA^{ASN}$ charging. These results indicate that the *B. cereus lysK* T-box regulatory element is sensitive to the level of charging of $tRNA^{ASN}$.

3.13 The *B. cereus lysK* T-box leader region requires a purine at the wobble position of the specifer codon for function

The data obtained so far suggests that the *lysK* T-box regulatory element may be unable to distinguish between uncharged $tRNA^{ASN}$ and uncharged $tRNA^{LYS}$. Analysis of the anticodons of $tRNA^{ASN}$ and $tRNA^{LYS}$ show that they only differ at the wobble base. This is a U carrying a 5-methylaminomethyl-2-thiouridine modification in the case of $tRNA^{LYS}$ and a G carrying a 7-deazoguanosine (queuosine) modification in the case of $tRNA^{ASN}$ (Reader, *et al.*, 2004; Agris, 2004). We hypothesized that the third position of the specifier codon was non-discriminating in this particular T-box system suggesting that recognition of tRNA depended on recognition of the first two bases of the tRNA anticodon. In order to examine this further we created a series of $P_{T\text{-}box\ lysK}$-*lacZ* fusions where the third position of the specifier codon was mutated from A to either a G, C or U (Figure 3.21a) and transformed them into *B. subtilis* strains with either IPTG dependant *lysS* genes or IPTG dependant *asnS* genes. Their respective ß-galactosidase activities were then established (Figure 3.22).

Results presented in figure 3.22 show ß-galactosidase activity data at 1mM IPTG, 250μM IPTG and 100μM IPTG for strains containing the $P_{T\text{-}box\ lysK}$-*lacZ* fusion with the wild type lysine specifier codon (AAA) and also codons with a mutation at the wobble base (AAG, AAC and AAU). The effect of these mutations at the specifier codon was examined in both an IPTG inducible *lysS* background and an IPTG inducible *asnS* background. The results show that mutation of the

(A)



(B)

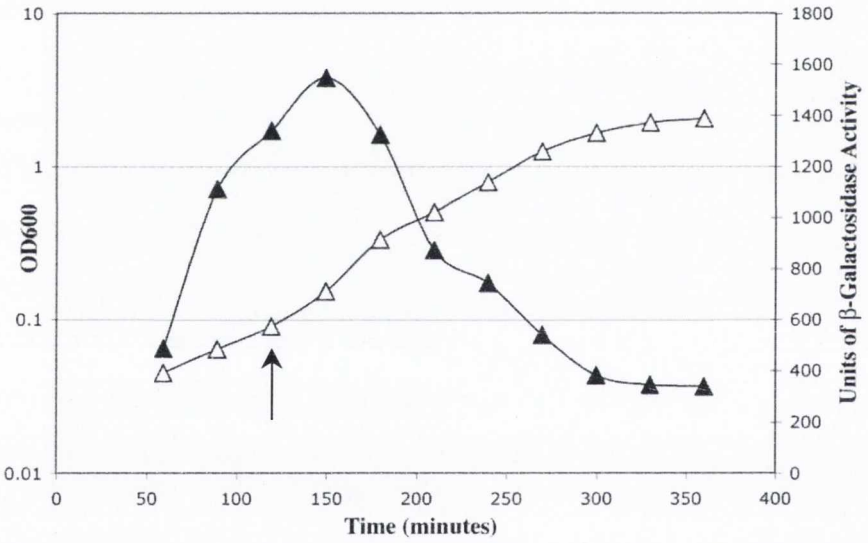
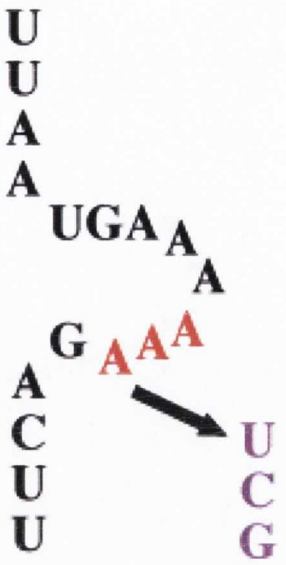


Figure 3.20 Growth and ß-galactosidase activity of NF60.6 [*trpC2 amyE*::pBCJ307.5($P_{T\text{-box } lysK}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R)] in LB containing varying IPTG concentrations. (A) Growth and ß-galactosidase activity of NF60.6 in LB containing 1mM, 250µM IPTG and 100µM IPTG. Growth is represented by open symbols. ß-galactosidase activity is represented by closed symbols. Strains and IPTG concentrations are as follows, NF60.6 1mM IPTG (◇, ◆), NF60.6 250µM IPTG (□, ■), NF60.6 100µM IPTG (△, ▲). (B) Growth and ß-galactosidase activity of NF60.6 in LB containing 100µM IPTG. The IPTG concentration was increased to 1mM at 120 minutes (indicated by arrow). Growth is represented by open symbols. ß-galactosidase activity is represented by closed symbols.

(A)



(B)

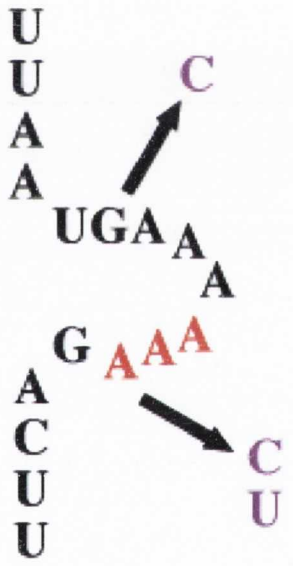


Figure 3.21 Diagrams showing the mutations made to the specifier loop of the *B. cereus lysK* T-box regulatory element. (A) Diagram showing the specifier loop of the *B. cereus lysK* T-box regulatory element. Mutations at the third position of the specifier codon are indicated. (B) Diagram of the *B. cereus lysK* T-box regulatory element specifier loop showing the G-C mutations and corresponding mutations at the third position of the specifier codon

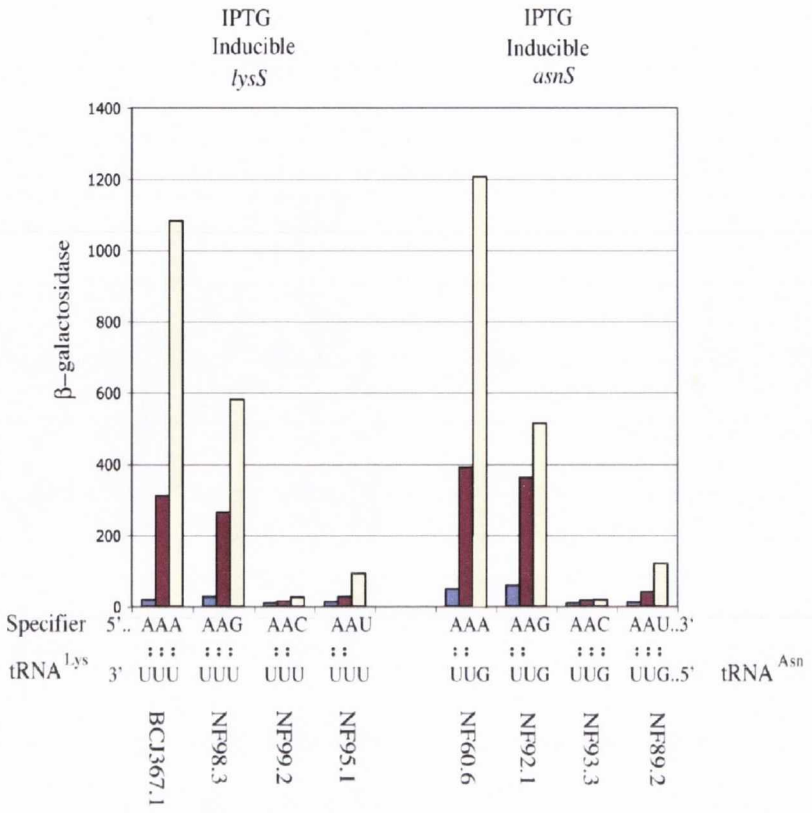


Figure 3.22 Analysis of the effect of mutations of the specifier codon of the *B. cereus lysK* T-box regulatory element on induction of ß-galactosidase activity by depletion for charged $tRNA^{ASN}$ and $tRNA^{LYS}$. The bar chart illustrates the ß-galactosidase activity of IPTG inducible *lysS* and *asnS* strains containing the $P_{T\text{-}box\ lysK}$-*lacZ* fusion with the wild-type specifier codon and also specifier codons mutated at the third position. The different coloured bars represent ß-galactosidase activity at 1mM (■), 250μM (■) and 100μM IPTG (□). The values indicated are taken from growth curves of the above strains in LB at an OD_{600} of 1.0. The specifier codons contained in each strain and their proposed tRNA interaction are indicated. The strains tested are from left to right BCJ367.1, NF98.3, NF99.2, NF95.1, NF60.6, NF92.1, NF98.3 and NF89.2

wobble base to a G, C or U results in lowered ß-galactosidase activity in both IPTG inducible *lysS* and IPTG inducible *asnS* backgrounds. Mutation of the wobble base of the specifier codon from an A to a G results in another lysine codon (AAG). In the strain NF98.3 (AAG) with an IPTG inducible *lysS* gene there is approximately 600 units of ß-galactosidase activity at 100μM IPTG compared with approximately 1100 units of activity at the same IPTG concentration in the strain BCJ367.1 containing the wild-type AAA specifier codon. A similar reduction in ß-galactosidase activity is observed with the AAG specifier codon in the IPTG inducible *asnS* background. Strain NF92.1 containing the AAG specifier codon shows a reduced level of ß-galactosidase activity of approximately 500 units at 100μM IPTG relative to the 1200 units of ß-galactosidase activity seen in the strain NF60.6 containing the wild type AAA specifier codon. These results show that both $tRNA^{LYS}$ and $tRNA^{ASN}$ are capable of interacting with the mutated AAG specifier codon albeit with less efficiency than observed with the wild-type AAA specifier codon (Figure 3.22).

Mutation of the specifier codon from AAA to AAC results in the switching of the specifier codon from a lysine codon to an asparagine codon. We posited that this mutation would increase the affinity of binding of $tRNA^{ASN}$ to the *lysK* leader specifier loop. The effect of this mutation on ß-galactosidase activity was examined in both an IPTG inducible *lysS* background (NF99.2) and an IPTG inducible *asnS* background (NF93.3). In contrast to the expected result, the interaction of $tRNA^{ASN}$ with the specifier codon was not improved by the AAA to AAC mutation. The results from these analyses show that this mutation results in severely reduced ß-galactosidase activity in both IPTG inducible *lysS* and IPTG inducible *asnS* backgrounds. Mutation of the specifer to the alternative asparagine codon AAU also results in a severe drop in ß-galactosidase activity. This reduction is not as great as that seen for the AAC specifier and some response to reduction in both $tRNA^{LYS}$ charging (NF95.1) and $tRNA^{ASN}$ charging (NF89.2) is seen, however the level of activity is ten fold less than that observed for the wild-type AAA specifier codon.

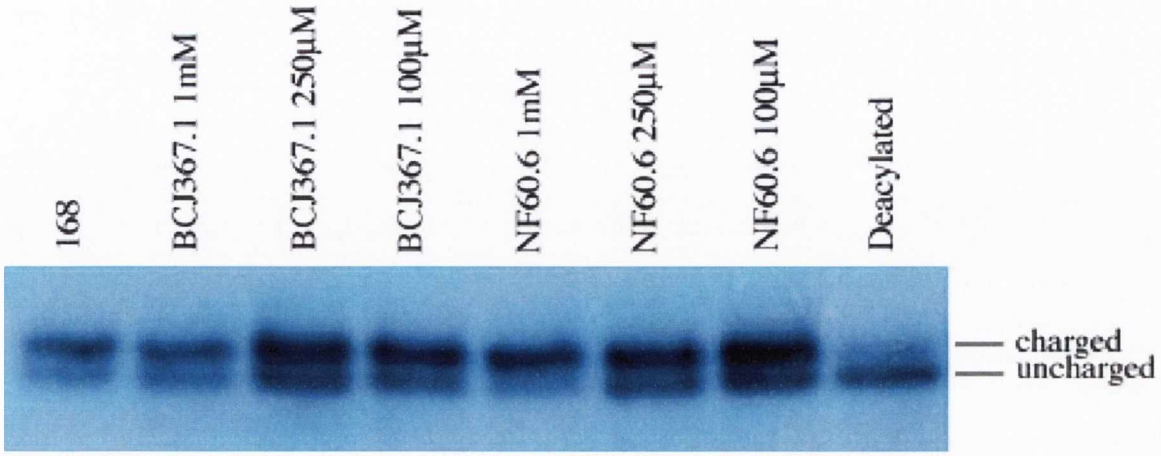
The results from these analyses indicate that mutation of the third position of the *B. cereus lysK* T-box leader region specifier codon from an A to any other nucleotide results in a reduction in the ability of $tRNA^{LYS}$ or $tRNA^{ASN}$ to interact with it. Mutation of the wobble base from A to another purine G results in an approximately 50% reduction in ß-galactosidase activity. However mutation of the wobble base to either of the pyrimidines C or U results in an almost complete removal of ß-galactosidase activity in the $P_{T\text{-}box\ lysK}$-*lacZ* fusion. This data would suggest therefore that a purine at

the wobble position is an essential requirement for function of the *B. cereus lysK* T-box regulatory element.

3.14 Analysis of the effect of a reduction in *asnS* expression on the level of charging of $tRNA^{LYS}$

Results to this point indicate that $tRNA^{ASN}$ may be capable of interacting with the *B. cereus lysK* T-box leader region. This has been demonstrated through a reduction in the level of $tRNA^{ASN}$ charging brought about through manipulation of an IPTG inducible *asnS* gene. In order to be certain that the results generated as a result of reduction in *asnS* expression were not brought about through a concomitant reduction in $tRNA^{LYS}$ charging we carried out a tRNA charging Northern blot experiment to assess the level of $tRNA^{LYS}$ charging during reduced *asnS* expression as described in materials and methods. *B. subtilis* strains BCJ367.1 (*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R)and NF60.6 [*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R) were grown in LB containing 1mM, 250μM and 100μM IPTG. Cells were harvested at an OD_{600} of approximately 0.5 and the RNA extracted. This RNA was then probed with a single stranded oligonucleotide specific for $tRNA^{LYS}$. The results for the IPTG inducible *lysS* strain BCJ367.1 showed a level of $tRNA^{LYS}$ charging of 83.7% in LB containing 1mM IPTG (Figure 3.23). This was reduced to 71.73% in medium containing 250μM IPTG and 71.82% when 100μM IPTG was present in the medium. The data for the IPTG inducible *asnS* strain NF60.6 showed 81.36% $tRNA^{LYS}$ charging in LB containing 1mM IPTG. However this level also fell upon reduction of IPTG concentration, to approximately 73% and 71% charging in medium containing 250μM and 100μM IPTG respectively (Figure 3.23). The observed reduction in $tRNA^{LYS}$ charging at 250μM IPTG was a surprise as there was no effect on growth of either BCJ367.1 of NF60.6 at this level of IPTG. This result also indicated that the increase in *lacZ* expression observed in NF60.6 upon reduction of IPTG concentration could be due to a reduction in $tRNA^{LYS}$ charging rather than $tRNA^{ASN}$ charging. In order to ascertain whether or not $tRNA^{ASN}$ was indeed capable of interacting with the *B. cereus lysK* T-box regulatory element we sought to establish the highest level of IPTG that would give us induction of ß-galactosidase activity in the $P_{T\text{-}box\ lysK}$-*lacZ* fusion. We observed that at 700μM and 600μM IPTG there was an induction of ß-galactosidase expression in both BCJ367.1 and NF60.6 (Figure 3.24). We

(A)

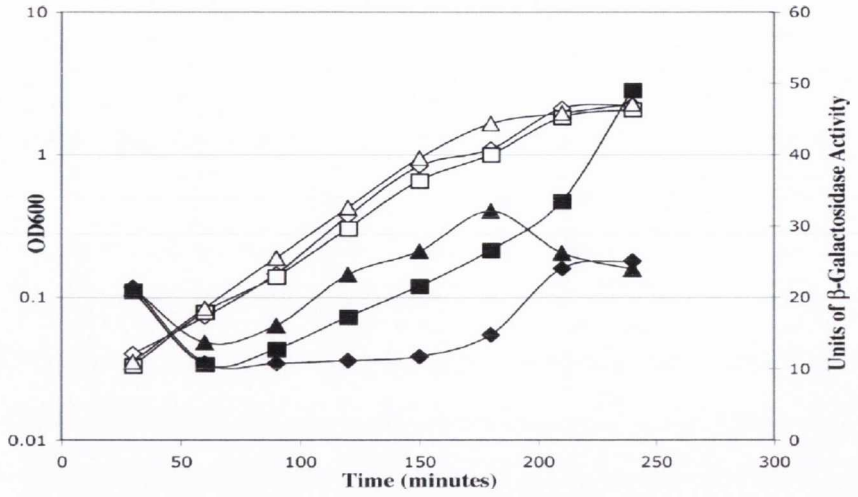


(B)

Strain	168	BCJ367.1 1mM	BCJ367.1 250µM	BCJ367.1 100µM	NF60.6 1mM	NF60.6 250µM	NF60.6 100µM
Percentage of charged tRNA	82.78	83.7	71.73	71.82	81.36	73.8	71.26

Figure 3.23 tRNA charging Northern blot analysis of *B. subtilis* strains containing IPTG inducible *lysS* and *asnS* genes grown in LB containing 1mM IPTG, 250μM IPTG and 100μM IPTG. (A) tRNA charging Northern analysis of the wt. *B. subtilis* 168 strain and strains BCJ367.1 [*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R)] and NF60.6 [*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R)] at different IPTG concentrations. (B) Table illustrating the percentage of charged $tRNA^{LYS}$ for the strains 168, BCJ367.1 and NF60.6.

(A)



(B)

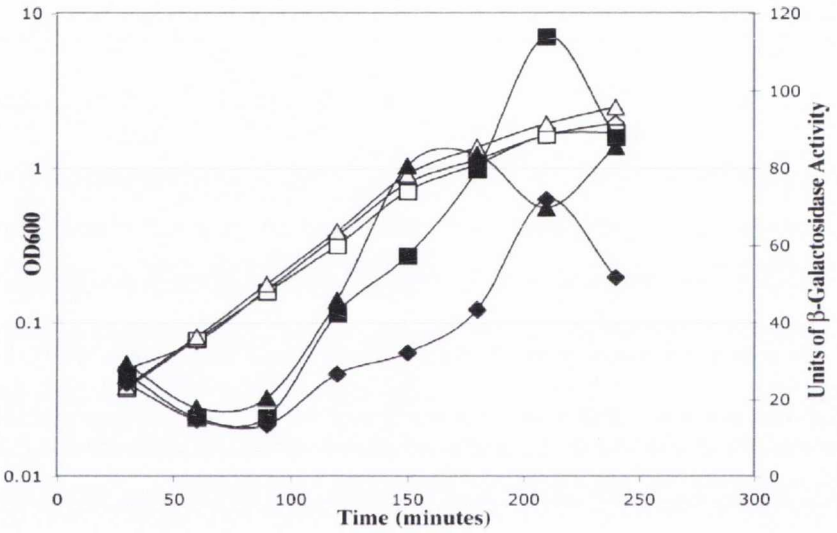
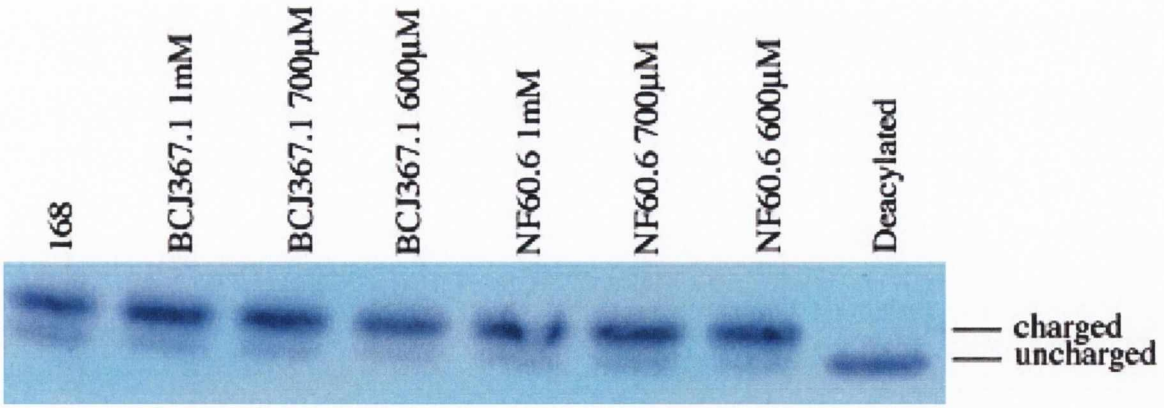


Figure 3.24 Growth and ß-galactosidase activity of BCJ367.1 [*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R)] and NF60.6 [*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R] at the highest concentrations of IPTG that induce an increase in ß-galactosidase activity in the $P_{T\text{-}box\ lysK}$-*lacZ* fusion. (A) Growth and ß-galactosidase activity of BCJ367.1 in LB containing 1mM, 700µM IPTG and 600µM IPTG. Growth is represented by open symbols. ß-galactosidase activity is represented by closed symbols. Strains and IPTG concentrations are as follows, BCJ367.1 1mM IPTG (◇, ◆), BCJ367.1 700µM IPTG (□, ■), BCJ367.1 600µM IPTG (△, ▲). (B) (A) Growth and ß-galactosidase activity of NF60.6 in LB containing 1mM, 700µM IPTG and 600µM IPTG. Growth is represented by open symbols. ß-galactosidase activity is represented by closed symbols. Strains are as follows, NF60.6 1mM IPTG (◇, ◆), NF60.6 700µM IPTG (□, ■), NF60.6 600µM IPTG (△, ▲).

(A)



(B)

Strain	168	BCJ367.1 1mM	BCJ367.1 700µM	BCJ367.1 600µM	NF60.6 1mM	NF60.6 700µM	NF60.6 600µM
Percentage of charged tRNA	74.2	83.4	83.94	83.16	80	80.3	81.09

Figure 3.25 tRNA charging Northern blot analysis of *B. subtilis* strains containing IPTG inducible *lysS* and *asnS* genes. (A) tRNA charging Northern analysis of the wt. *B. subtilis* 168 strain and strain BCJ367.1 [*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (*penP-lacI* Phl^R)] and NF60.6 [*trpC2 amyE*::pBCJ307.5($P_{T\text{-}box\ lysK}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R)] at different IPTG concentrations. (B) Table illustrating the percentage of charged $tRNA^{LYS}$ for the strains 168, BCJ367.1 and NF60.6 at different IPTG concentrations.

then repeated the tRNA charging Northern blot protocol to establish the percentage of charged $tRNA^{LYS}$ in strain BCJ367.1 and strain NF60.6 grown at 1mM, 700μM and 600μM IPTG (Figure 3.25). Under these growth conditions the level of charging of $tRNA^{LYS}$ in strain NF60.6 was approximately 80%. Upon reduction of IPTG concentration to 700μM and 600μM IPTG, no reduction in the percentage of charged $tRNA^{LYS}$ was observed. Under the same growth conditions, BCJ367.1 showed approximately 83% charged $tRNA^{LYS}$ at 1mM IPTG. Reduction of the IPTG concentration to 700μM and 600μM IPTG did not produce a concomitant reduction in $tRNA^{LYS}$ charging. As we were able to observe an increase in lacZ expression as a result of reduced $tRNA^{LYS}$ charging in strain BCJ367.1, the tRNA charging data infers that the densitometry carried out was not sensitive enough to pick up small changes in $tRNA^{LYS}$ charging. However, the *lacZ* expression data for strain NF60.6 indicates that the *B. cereus lysK* T-box regulatory element is sensitive to the level of $tRNA^{ASN}$ charging in the cell .

3.15 Mutation of the specifier codon from AAA to AAC or AAU does not produce base pairing within the specifier loop

We observed that the mutations made at the third position of the specifier codon from AAA to both AAC and AAU caused a significant reduction in the ß-galactosidase activity of the $P_{T\text{-}box\ lysK}$-*lacZ* fusion. We noted that it was possible that base pairing between the G at the second position of the specifier loop and the C or U at the wobble base of the mutated specifier codons could potentially alter the structure of the loop, making it unable to interact with the anticodons of lysyl or asparaginyl tRNA. In order to test this hypothesis we made $P_{T\text{-}box\ lysK}$-*lacZ* fusions containing the AAU and AAC specifier codons with the second position G nucleotide of the specifier loop mutated to a C (Figure 3.21b). These fusions were then transformed into *B. subtilis* strains containing either IPTG inducible *lysS* or *asnS* genes and inserted into the amylase locus by double crossover thereby placing this fusion in single copy into the chromosome. Strains containing the $P_{T\text{-}box\ lysK}$-*lacZ* fusion with the G-C mutation at the second position of the specifier loop with the AAC and AAU specifier codons were then grown at 250μM IPTG and their ß-galactosidase activity established to assess the affect of the G-C mutation. A fusion containing the G-C mutation with the AAA wild-type specifier was also created to assess the G-C mutation on an otherwise unaltered *lysK* T-box regulatory element. Data from this analysis is presented in figure 3.26.

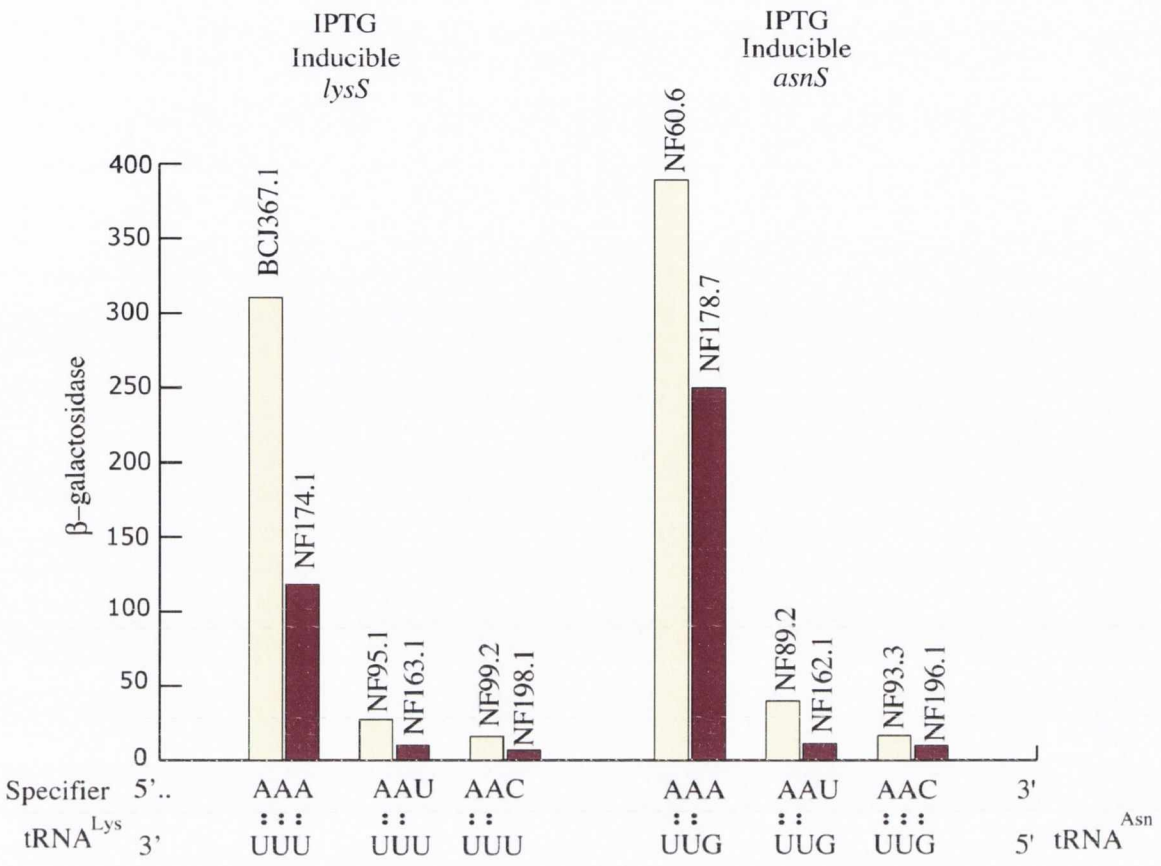


Figure 3.26 Bar chart illustrating the ß-galactosidase activity of IPTG inducible *lysS* and *asnS* strains containing the the $P_{T\text{-}box\ lysK}$-*lacZ* fusion with the wild-type specifier loop and also specifier loops with a mutation of the G at the second position of the specifier loop to a C. Strains containing the G-C mutation are represented by the wine coloured bars. Strains containing the wild-type G at the second position of the loop are represented by the pale yellow bars. The different specifier codons and their proposed tRNA interactions are indicated. The strains used in this experiment are listed over the corresponding ß-galactosidase data. The strains employed in this analysis from left to right are BCJ367.1, NF174.1, NF95.1, NF163.1, NF99.2, NF198.1, NF60.6, NF178.7, NF89.2, NF162.1, NF93.3 and NF196.1.

The results from this analysis show that mutation of the G at the second position of the specifier loop to a C causes a reduction in ß-galactosidase activity in all the $P_{T\text{-}box\ lysK}$-*lacZ* fusions tested regardless of the nucleotide at the wobble base. In strain NF174.1 containing the G-C mutation with the wild-type AAA specifier codon in an IPTG inducible *lysS* background and NF178.7 containing the G-C mutation with the wild-type AAA specifier codon in an IPTG inducible *asnS* background a 1.5 fold and 3 fold reduction respectively, was observed relative to the wild-type $P_{T\text{-}box\ lysK}$-*lacZ* fusion (Figure 3.26). The strains NF162.1 and NF163.1 containing the AAU specifier and the G-C mutation in an IPTG inducible *asnS* and IPTG inducible *lysS* background respectively, also showed a reduction in ß-galactosidase activity relative to NF89.2 (IPTG inducible *asnS* with AAU specifier codon) and NF95.1 (IPTG inducible *lysS* with AAU specifier codon) which have the wild-type G at the second position of the specifier loop. A similar result was observed for strains NF196.1 (IPTG inducible *asnS*) and NF198.1 (IPTG inducible *lysS*) which have the G-C mutation and the AAC specifier codon. These data infer that the mooted G-U and G-C interactions between the G at the second position of the specifier loop and the third base of the specifier codon do not occur. Therefore the reductions in ß-galactosidase activity observed in strains containing AAU and AAC specifier codons cannot be attributed to this. The reduction in *lacZ* expression observed in strains NF174.1 and NF178.7 containing the wild-type AAA specifier codon indicates that the G nucleotide at the second position of the specifier loop plays a role in insuring a successful specifier codon-tRNA anticodon interaction. A guanosine in this position may be important for the overall structure of the specifier loop and as a result, in presentation of the specifier codon for tRNA.

Chapter 4

Investigation into the T-box transcriptional regulatory element of the asparaginyl tRNA synthetase of *Bacillus cereus*

4.1 *B. cereus* has a predicted T-box regulatory element upstream of the *asnS* gene

We presented data in the previous chapter that the class 1 *lysK* of *B. cereus* is sensitive to a reduction in the level of charged $tRNA^{ASN}$ in addition to $tRNA^{LYS}$. We hypothesized that this promiscuity in tRNA recognition in this system could be a reason for the relative paucity of T-box regulated *lysS* genes relative to other AARS genes among the sequenced bacterial genomes. As asparagine and lysine occupy the same mixed codon box we were curious as to whether T-box regulatory mechanisms controlling transcription of *asnS* genes also showed a lack of specificity in terms of tRNA recognition i.e. do they respond to depletion for charged $tRNA^{LYS}$? We carried out a search of the sequenced bacterial genomes in order to discover the number of bacterial species containing a T-box regulated *asnS* gene. There are 24 bacterial species that contain a T-box regulated *asnS* gene. Thus it is one of the AARS genes least regulated by a T-box mechanism. The bacterial species containing a T-box regulated *asnS* gene are listed in table 4.1 and come from three families, the *bacillaceae*, the *lactobacillaceae* and the *clostridaceae.*

Amongst the bacterial strains listed in table 4.1 are *B. cereus* 14579 and *Lactobacillus delbrueckii bulgaricus* 11842. In *B. cereus* there is one copy of the *asnS* gene that has a putative T-box regulatory element. In *L. delbrueckii bulgaricus* two copies of the *asnS* gene are present. The two copies of *asnS* are highly homologous differing at only two nucleotide positions. One of these *asnS* genes appears to be co-regulated in conjunction with an asparagine synthase gene by a T-box mechanism. The predicted secondary structures of both the *B. cereus* and the *L. delbrueckii bulgaricus* T-box regulatory elements are presented in figures 4.1 and 4.2.

Table 4.1 Distribution of T-box regulated *asnS* genes among the sequenced bacterial genomes

List of species containing T-box regulated *asnS* genes
Alkaliphilus_metalliredigens
Bacillus anthracis Ames NC_00753
Bacillus anthracis Ames NC 003997
Bacillus anthracis str Sterne
Bacillus cereus 14579
Bacillus cereus ATCC 10987
Bacillus cereus ZK NC 006274
Bacillus cereus cytotoxis
Bacillus thuringiensis Al Hakam
Bacillus thuringiensis konkukian
Clostridium botulinum ATCC 19397
Clostridium botulinum Hall
Clostridium botulinum NC 009495
Clostridium botulinum F Langeland
Clostridium difficile 630 NC 00908
Clostridium kluyveri DSM 555 NC 00970
Clostridium novyi NT NC 008593
Clostridium perfringens ATCC 13124 NC 008261
Clostridium perfringens NC 003366
Clostridium perfringens SM101 NC 008262
Clostridium tetani E88 NC 004557
Lactobacillus brevis ATCC 367 NC 008497
Lactobacillus delbrueckii bulgaricus NC 008529
Pediococcus pentosaceus NC 008525

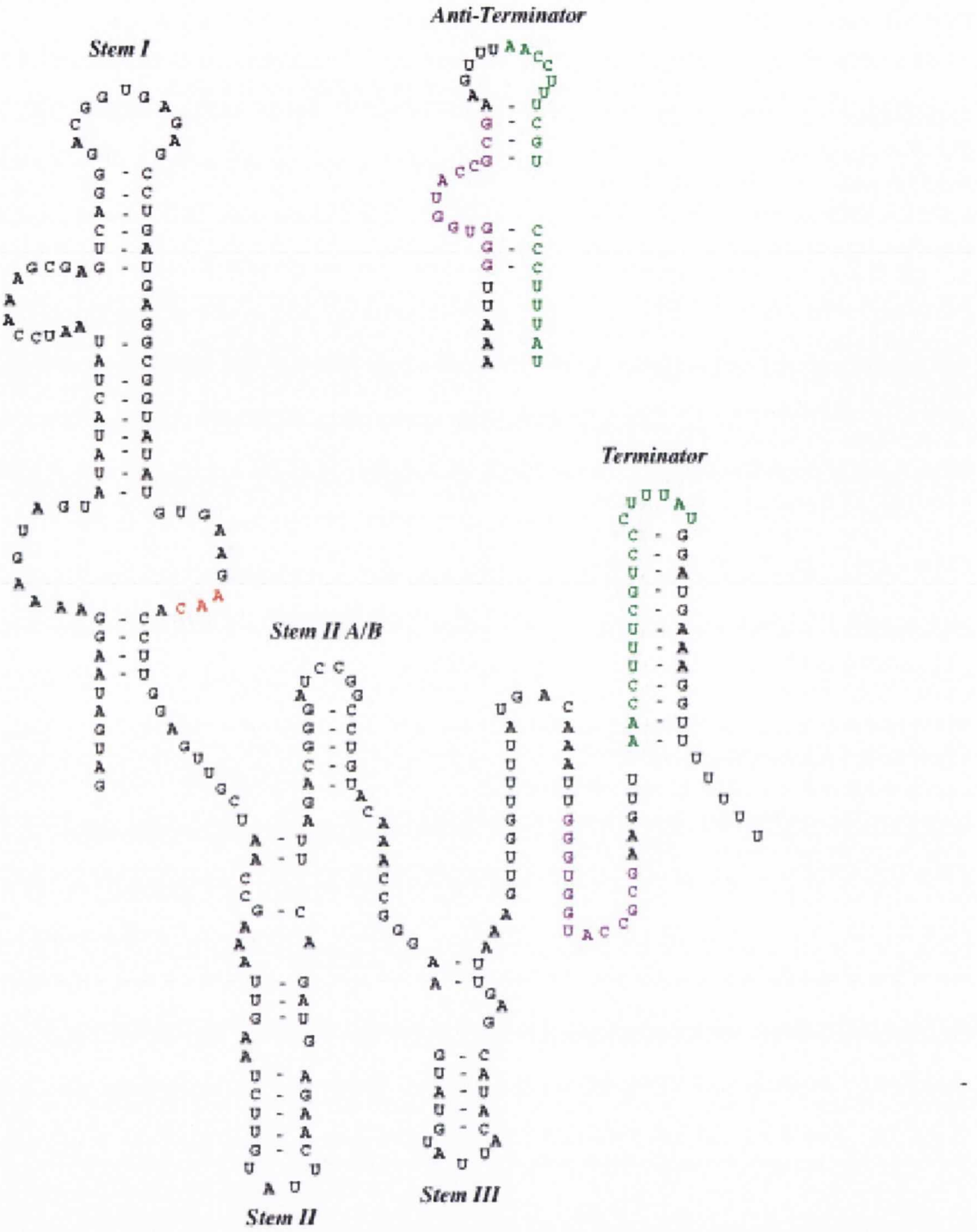


Figure 4.1 Proposed secondary structure of the upstream leader region of the asparaginyl tRNA synthetase gene of *B. cereus* 14579. The T-box region is highlighted in purple. Sequences that can form part of either the terminator or antiterminator are highlighted in green. The specifier codon is indicated in red. Other important structural features are indicated.

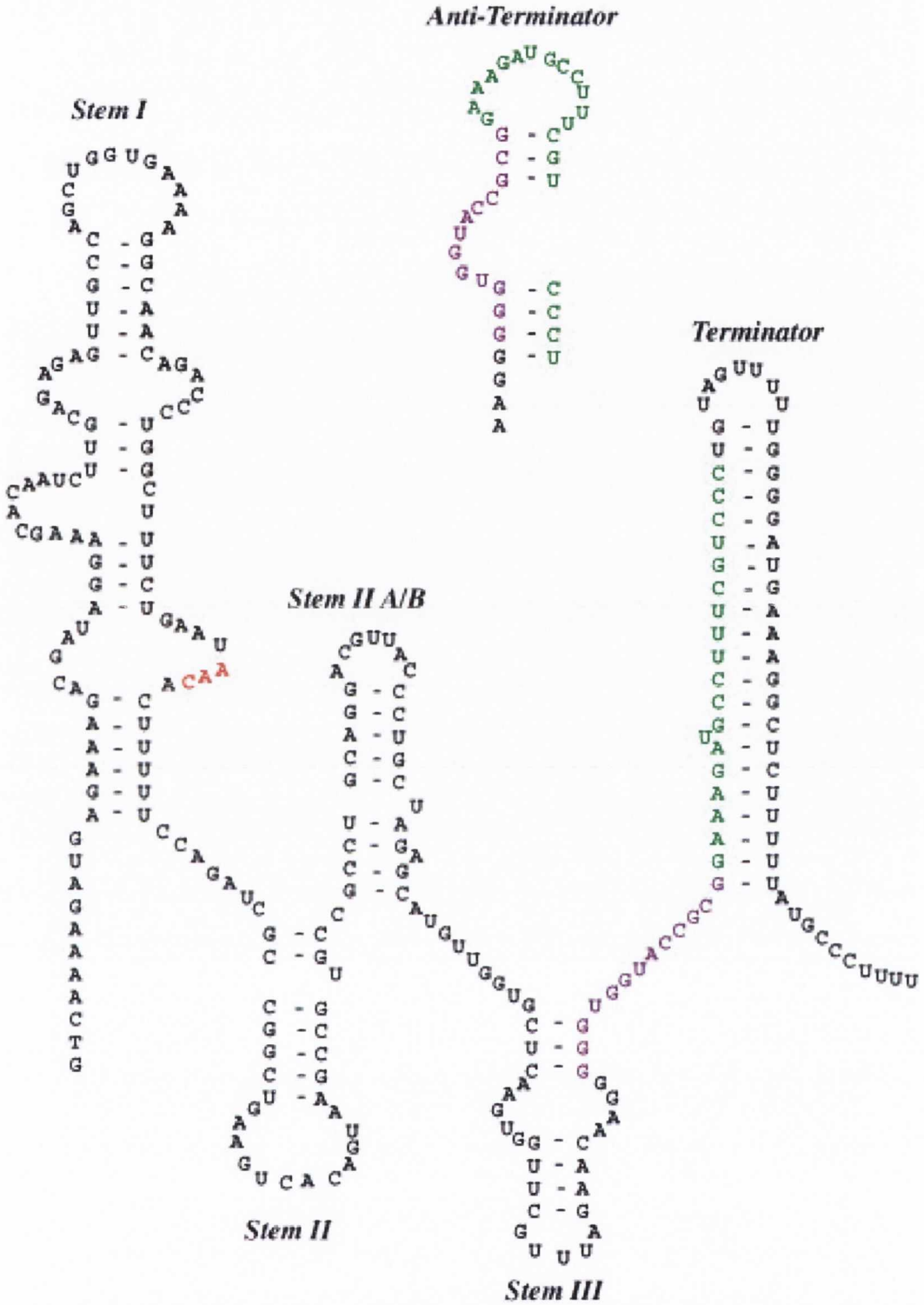


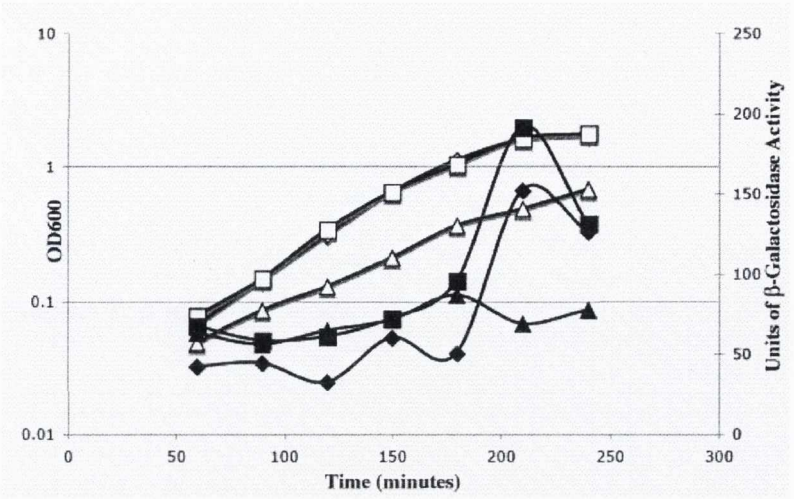
Figure 4.2 Proposed secondary structure of the upstream leader region of the asparaginyl tRNA synthetase gene of *L. delbrueckii bulgaricus*. The T-box region is highlighted in purple. Sequences that can form part of either the terminator or antiterminator are highlighted in green. The specifier codon is indicated in red. Other important structural features are indicated.

4.2 The *asnS* T-box regulatory elements of both *B. cereus* and *L. delbrueckii bulgaricus* do not respond to depletion for charged tRNA^{ASN} in *B. subtilis*

In order to assess whether or not the *asnS* T-box regulatory elements for *B. cereus* and *L. delbrueckii bulgaricus* were functional we created a transcriptional fusion of the promoter and upstream leader regions of the *B. cereus asnS* gene to the *lacZ* reporter gene ($P_{B.\ cereus}$ T-box $_{asnS}$-*lacZ*) and also a transcriptional fusion of the promoter and upstream leader region of the *L. delbrueckii bulgaricus asnS* gene to the *lacZ* reporter gene ($P_{L.\ d.\ bulgaricus}$ T-box $_{asnS}$-*lacZ*). These transcriptional fusions were then inserted in single copy into the chromosome of *B. subtilis* strains containing IPTG inducible *asnS* genes to create the strains NF175.7 containing the ($P_{B.\ cereus}$ T-box $_{asnS}$-*lacZ*) fusion and NF179.2 containing the ($P_{L.\ d.\ bulgaricus}$ T-box $_{asnS}$-*lacZ*) fusion. In these strains, the amount of *asnS* expression can be altered consistent with changes in the level of AsnRS in the cell and hence changes in the level of charged tRNA^{ASN}. These strains were grown in LB containing different concentrations of IPTG and the ß-galactosidase activity in these strains was determined.

Both NF175.7 ($P_{B.\ cereus}$ T-box $_{asnS}$-*lacZ*) and NF179.2 ($P_{L.\ d.\ bulgaricus}$ T-box $_{asnS}$-*lacZ*) were grown in LB containing 1mM, 600µM and 100µM IPTG. It was shown in the previous chapter that a reduction in the concentration of IPTG in the medium containing a strain with an IPTG inducible *asnS* gene resulted in increased ß-galactosidase activity in strains containing ($P_{T\text{-box }lysK}$-*lacZ*) fusions indicating that the level of asn-tRNA^{ASN} had been reduced. We expected that as the IPTG concentration in the medium was reduced from 1mM IPTG to 600µM IPTG to 100µM IPTG, that there would be a corresponding increase in the level of ß-galactosidase activity in strains NF175.7 and NF179.2 as the level of charged tRNA^{ASN} was reduced. However, the results from these experiments, presented in figure 4.3A and B show that for both NF175.7 and NF179.2, as the concentration of IPTG in the medium is reduced, no difference in ß-galactosidase activity is observed. In the case of NF175.7 ($P_{B.\ cereus}$ T-box $_{asnS}$-*lacZ*) the level of ß-galactosidase activity ranged from 40 units to approximately 200 units during the growth cycle at 1mM and 600µM IPTG. In media containing 100µM IPTG, expression was slightly lower ranging from 40 units to ~ 100 units (figure 4.3A). For NF179.2 ($P_{L.\ d.\ bulgaricus}$ T-box $_{asnS}$-*lacZ*) the level of ß-galactosidase activity was much lower ranging

(A)



(B)

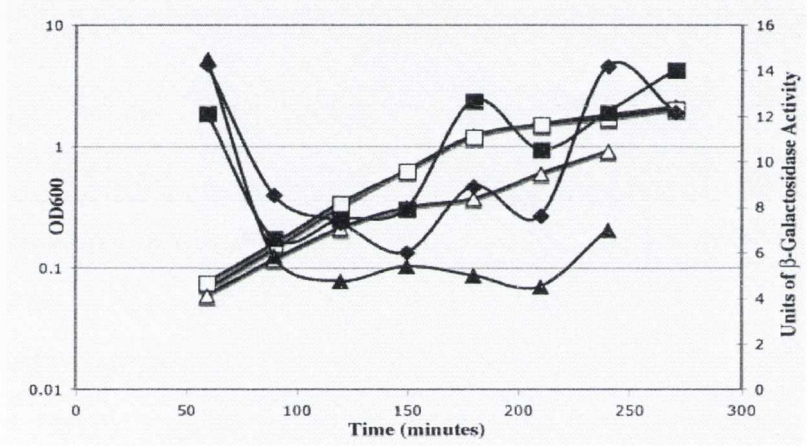


Figure 4.3 Analysis of the response of the *B. cereus* and *L. delbrueckii bulgaricus asnS* T-box regulatory elements to starvation for charged $tRNA^{ASN}$. (A) Growth and ß-galactosidase activity of NF175.7 [*trpC2 amyE*::pNF161.11($P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R)] in LB containing 1mM, 600µM IPTG and 100µM IPTG. Growth is represented by open symbols. ß-galactosidase activity is represented by closed symbols. Strains and IPTG concentrations are as follows, NF175.7 1mM IPTG (◇, ◆), NF175.7 600µM IPTG (□, ■), NF175.7 100µM IPTG (△, ▲). (B) Growth and ß-galactosidase activity of NF179.2 [*trpC2 amyE*::pNF165.2 ($P_{L.\ d\ bulgaricus\ \text{T-box}\ asnS}$-*lacZ*) Cm^R(P_{spac}-*asnS* Em^R) pMAP65 (*penP-lacI* Phl^R)] in LB containing 1mM, 600µM IPTG and 100µM IPTG. Growth is represented by open symbols. ß-galactosidase activity is represented by closed symbols. Strains and IPTG concentrations are as follows, NF179.2 1mM IPTG (◇, ◆), NF179.2 600µM IPTG (□, ■), NF179.2 100µM IPTG (△, ▲).

from 6-14 units. The level of ß-galactosidase activity was not affected by a reduction in the concentration of IPTG (figure 4.3B) From these data we conclude that neither the putative *asnS* T-box regulatory element from *B. cereus* or the putative *asnS* T-box regulatory element from *L. delbrueckii bulgaricus* is sensitive to the level of charged $tRNA^{ASN}$ in a *B. subtilis* background.

4.3 The *asnS* T-box regulatory element of *B. cereus* does not respond to depletion for charged $tRNA^{LYS}$ in *B. subtilis*

We observed no increase in *lacZ* expression in either the $P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ* fusion or the $P_{L.\ d\ bulgaricus\ \text{T-box}\ asnS}$-*lacZ* fusion as a result of reduced charging of $tRNA^{ASN}$. Following this observation we decided to investigate only the *B. cereus asnS* T-box regulatory element. In the *B. cereus lysK* T-box regulatory element we had previously observed promiscuity in tRNA recognition between $tRNA^{LYS}$ and $tRNA^{ASN}$. We decided to investigate the possibility of a similar effect in the *B. cereus asnS* T-box regulatory element. In order to investigate this we integrated the ($P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ*) fusion in single copy at the amylase locus of a *B. subtilis* strain containing an IPTG inducible *lysS* gene. We could then assess the response of the *B. cereus asnS* T-box regulatory element to depletion for charged $tRNA^{LYS}$ by lowering the concentration of IPTG in the medium and thus *lysS* expression. The resultant strain was NF177.1.

It was shown in the previous chapter that a reduction in the concentration of IPTG in the medium of a culture containing a strain with an IPTG inducible *lysS* gene and a ($P_{T\text{-}box\ lysK}$-*lacZ*) fusion resulted in an increase in ß-galactosidase activity indicating a reduction in the level LysRS in the cell and thus a reduced level of charged $tRNA^{LYS}$. Strain NF177.1 ($P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ*::*amyE* P_{spac}*lysS*) was grown in LB containing 1mM, 600µM and 250µM IPTG. Results are presented in figure 4.4. The data shows no difference in ß-galactosidase activity as a result of a reduction in the concentration of IPTG in the medium. Similarly to the profile of ß-galactosidase activity observed in strain NF175.7 ($P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ*), the units of activity range from 40 units to approximately 180 units during the growth cycle regardless of the concentration of IPTG in the medium (figure 4.4) The results from this data indicate that the putative *asnS* T-box regulatory element from *B. cereus* does not respond to changes in the level of charged $tRNA^{LYS}$ in the cell.

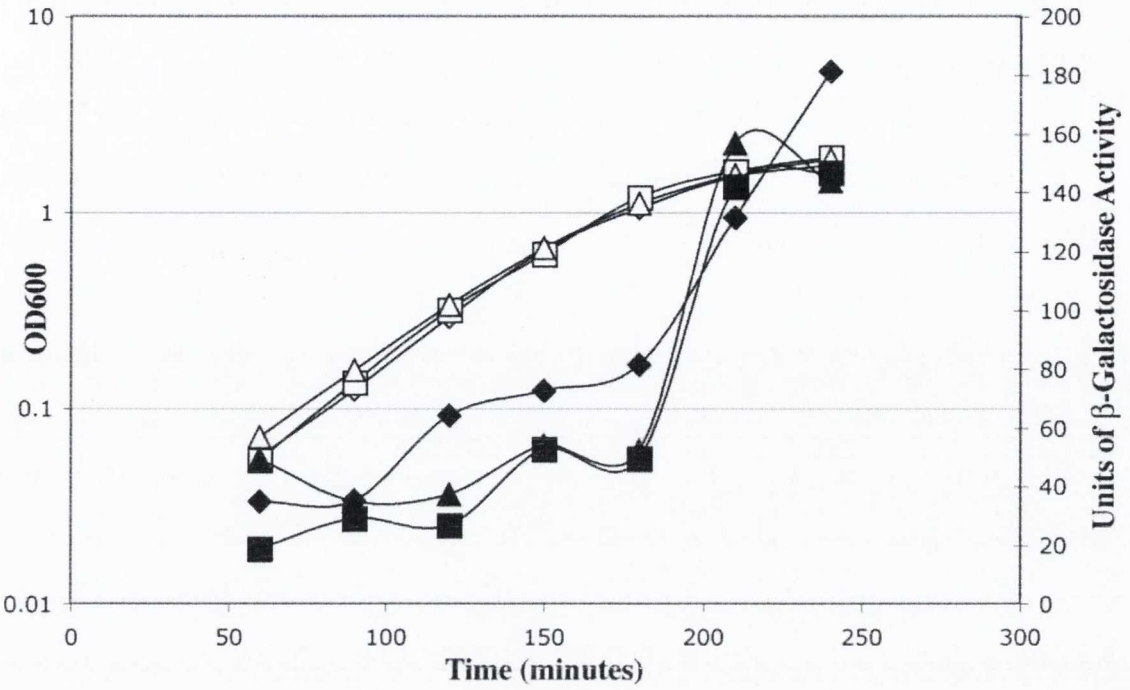


Figure 4.4 Analysis of the response of the *B. cereus asnS* T-box regulatory element to starvation for charged $tRNA^{LYS}$. Growth and ß-galactosidase activity of NF177.1 1 [trpC2 amyE::pNF161.11($P_{B.\ cereus\ T\text{-}box\ asnS}$-*lacZ*) Cm^R(P_{spac}-*lysS* Em^R) pMAP65 (penP-lacI Phl^R)] in LB containing 1mM, 600µM IPTG and 250µM IPTG. Growth is represented by open symbols. ß-galactosidase activity is represented by closed symbols. Strains and IPTG concentrations are as follows, NF177.1 1mM IPTG (◇, ◆), NF177.1 250µM IPTG (□, ■), NF177.1 100µM IPTG (△, ▲).

4.4 The T-box regulatory element of *asnS* from *B. cereus* does not respond to depletion for tryptophan

To assess the affect of general amino acid starvation on the $P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ* fusion we analyzed the response of the *B. cereus asnS* T-box regulatory element to a reduction in the level of charged tryptophanyl tRNA. In order to address this question we constructed the strain NF187.1 containing the $P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ* fusion integrated in single copy at the amylase locus of the *B. subtilis* tryptophan auxotroph strain 168. Strain NF187.1 was grown in Spizizens minimal media containing 50 μg/ml (tryptophan replete) and 1 μg/ml tryptophan (tryptophan limiting) and the effect on ß-galactosidase activity was analyzed (Figure 4.5).

In both the tryptophan replete and tryptophan limited culture the ß-galactosidase activity ranges from 40 units to approximately 160 units during the growth cycle indicating no change in ß-galactosidase activity as a result of limitation for tryptophan (Figure 4.5). This result shows that the *B. cereus asnS* T-box regulatory element is not sensitive to starvation for tryptophan.

4.5 Could differences in $tRNA^{ASN}$ structure between *B. subtilis* and *B. cereus* account for the lack of response of $P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ* fusions to charged $tRNA^{ASN}$ depletion in *B. subtilis*

Our results show that the T-box regulatory element of the *asnS* gene from *B. cereus* is not responsive to charging levels of $tRNA^{ASN}$ in *B. subtilis*. To assess the possibility that differences in $tRNA^{ASN}$ structure/sequence between *B. subtilis* and *B. cereus* may be the reason for the lack of antitermination of the *B. cereus asnS* T-box regulatory element in response to reduced *B. subtilis* $tRNA^{ASN}$ charging, we examined the sequence of $tRNA^{ASN}$ in both strains.

We compared the secondary structures of the five asparaginyl tRNAs of *B. cereus* and the four asparaginyl tRNAs of *B. subtilis* (figure 4.6). We found that two out of the four *B. subtilis* $tRNA^{ASN}$ species were identical to those of *B. cereus* except for the presence of an adenine at position 58 instead of the guanine found in *B. cereus* $tRNA^{ASN}$. The other $tRNA^{ASN}$ species found in *B. subtilis* also had an adenine at position 58. In addition to this

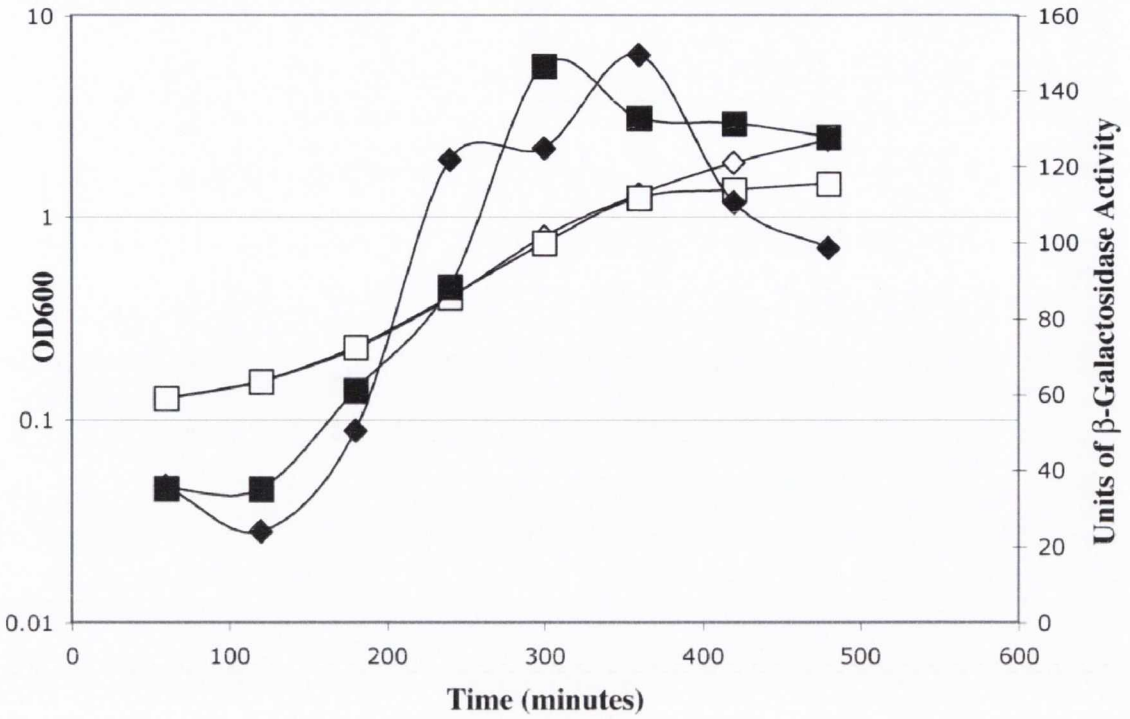


Figure 4.5 Analysis of the response of the *B. cereus asnS* T-box regulatory element to starvation for tryptophan. The data shows growth and ß-galactosidase activity of NF187.1 [trpC2 amyE::pNF161.11($P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ*) Cm^R] in minimal media containing 50µg/ml tryptophan and 1µg/ml tryptophan. Growth is represented by open symbols. ß-galactosidase activity is represented by closed symbols. Strains and tryptophan concentrations are as follows, NF187.1 50µg/ml tryptophan (◇, ◆), NF187.1 1µg/ml tryptophan (□, ■).

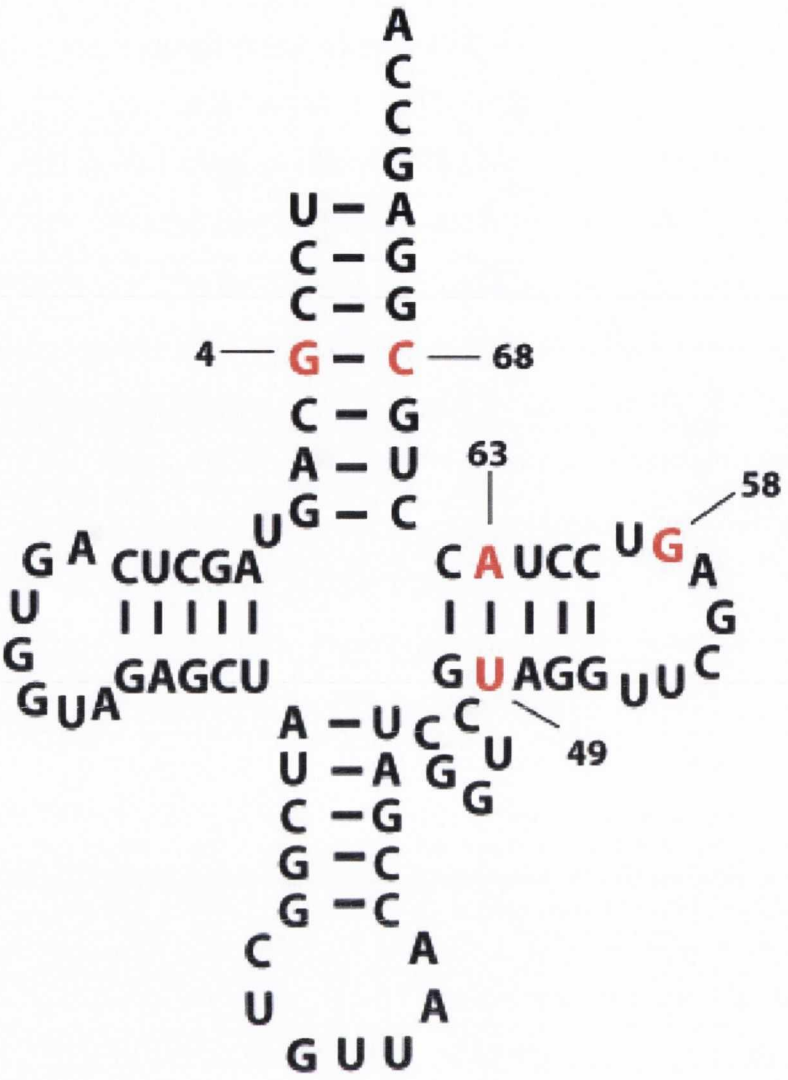


Figure 4.6 Secondary structure of the asparaginyl tRNA of *B. cereus*. Bases highlighted in red indicate the points at which *B. subtilis* asparaginyl tRNA's differ. The guanine at position 58 of *B. cereus* $tRNA^{ASN}$ is an adenine residue in all *B. subtilis* asparaginyl tRNA's. In two out of the four *B. subtilis* asparaginyl tRNA's there are two further mismatches. These are an adenine at position 4 with a complementary uracil at position 68 and a cytosine at position 40 with a complementary guanine at position 63.

they possessed an adenine at position 4 with a complementary uracil at position 68 and a cytosine at position 49 with a complementary guanine at position 63. In *B. cereus* $tRNA^{ASN}$ there is a guanine at position 4 with complementary cytosine at position 68 and uracil at position 49 with complementary adenine at position 63.

The $tRNA^{ASN}$ sequences of *B. subtilis* and *B. cereus* are very similar although not identical. It is possible that the minor differences in the tRNA, possibly in combination with other differences in the internal environment of the two bacterial species could account for the lack of induction of expression of the $P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ* fusion to depletion for charged $tRNA^{ASN}$ in *B. subtilis*.

To further investigate this we decided to attempt a transcriptional fusion analysis of the *B. cereus asnS* T-box regulatory element in *B. cereus*. Additionally, as we would be carrying out a transcriptional fusion analysis in B. cereus, we decided to repeat our investigation of the *B. cereus lysK* T-box regulatory element. This would allow us to confirm the interaction of $tRNA^{ASN}$ with the *lysK* T-box regulatory element in its natural environment.

4.6 Analysis of the *asnS* T-box regulatory element in *B. cereus*

To analyze the expression of both the $P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ* and the $P_{\text{T-box}\ lysK}$-*lacZ* fusions in *B. cereus* we had to create new transcriptional fusion constructs. This was because *B. cereus* requires approximately 1kb of nucleotide homology for successful integration onto the chromosome. By comparison with *B. subtilis*, *B. cereus* is a relatively less well studied organism. As a result of this less genetic manipulation techniques have been developed for working in this organism. For example, no vectors for integration in single copy at a neutral locus are available for *B. cereus*. We constructed new transcriptional fusion plasmids by inserting approximately 1kb of DNA containing the promoter and T-box regulatory elements of the *asnS* and *lysS* genes into plasmid pDG268. The resultant plasmids were pNF223.8 containing the new $P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ* fusion and pNF222.1 containing the new $P_{\text{T-box}\ lysK}$-*lacZ* fusion. These plasmids were integrated onto the genome of *B. cereus* using electroporation. The resultant strains were NF232.1 containing the $P_{\text{T-box}\ lysK}$-*lacZ* fusion and NF259.1 containing the $P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ* fusion. These strains were then grown in LB at

30°C and their ß-galactosidase activity was analyzed. The data is presented in figure 4.7.

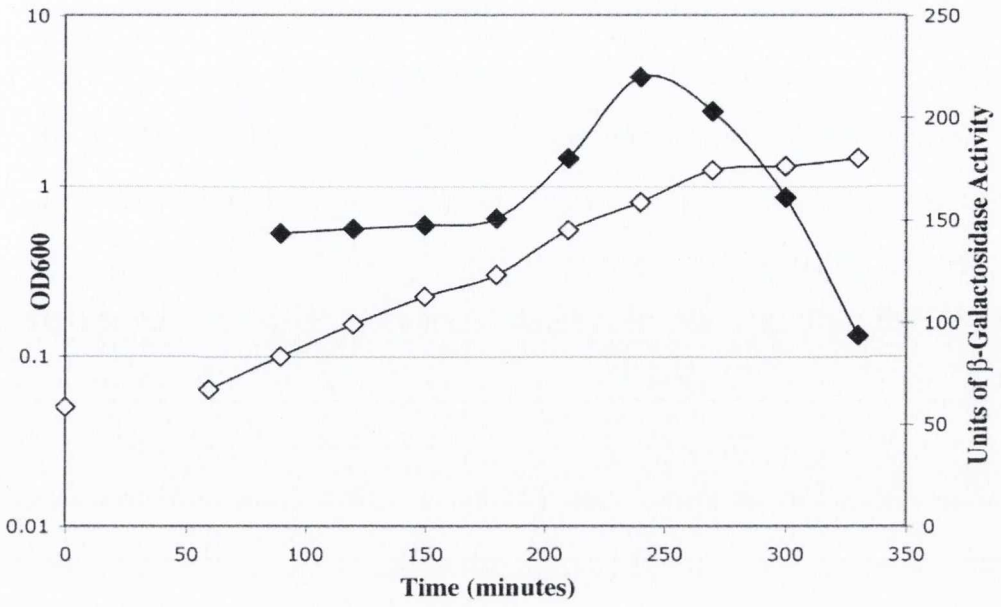
For strain NF232.1 approximately 150 units of ß-galactosidase activity was produced until an OD_{600} of around 0.4. This activity then increased to approximately 220 units at OD_{600} 1.0 before decreasing again in stationary phase. In strain NF259.1 ß-galactosidase activity increased from 200 units to approximately 800 units during the exponential growth phase. This activity then declined during stationary phase (Figure 4.7).

In order to rule out differences in tRNA structure and possible differences in the internal environments of *B. subtilis* and *B. cereus* as the reason for the lack of induction of expression in the $P_{B.\ cereus}$ T-box *asnS*-*lacZ* fusion when depleting for charged $tRNA^{ASN}$ in *B. subtilis*, we attempted to make a strain of *B. cereus* with an IPTG inducible *asnS* gene. Using such a strain we would be able to lower the amount of *asnS* expression and hence AsnRS levels in the cell resulting in a reduced amount of charged $tRNA^{ASN}$. We could then observe the effect of this on the $P_{B.\ cereus}$ T-box *asnS*-*lacZ* fusion. A similar strategy could also be used to analyze the effect of lowered asn-$tRNA^{ASN}$ levels on the *lysK* T-box regulatory element.

To place the *B. cereus asnS* gene under IPTG inducible control we cloned a 1036bp fragment of a region upstream and overlapping the start of the *B. cereus asnS* gene into the plasmid pMUTINXZ containing the P_{spac} promoter. The resultant plasmid was pNF223.8. This construct was then electroporated into *B. cereus*. Integration of plasmid pNF223.8 into the chromosome of *B. cereus* failed to produce strains whose expression of *asnS* was IPTG dependant. All transformants recovered had IPTG independent expression of *asnS*. Efforts to tighten the regulation of the P_{spac} promoter through the use of the pMAP65 plasmid which contains multiple *lacI* sites were prevented due to the natural resistance of *B. cereus* to the two antibiotic drug cassettes (Kanamycin and Phleomycin) present on the plasmid. Efforts to create a *B. cereus* strain with an IPTG dependant *lysS* gene were also unsuccessful.

As an alternative approach to reducing the level of $tRNA^{ASN}$ charging in the cell we searched for an asparagine analogue that would competitively inhibit the function of AsnRS and as a result reduce the level of asn-$tRNA^{ASN}$ in the cell. The effect of reduced

(A)



(B)

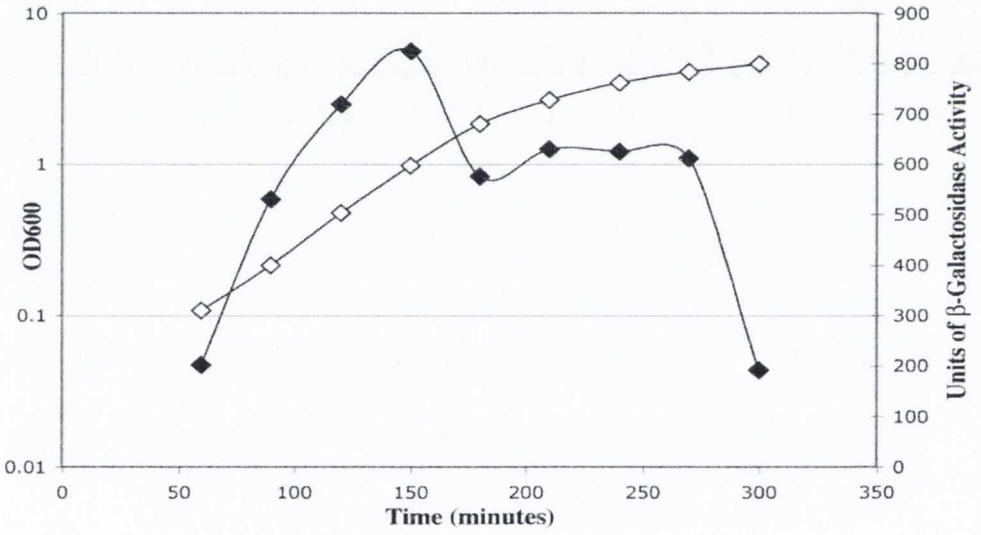


Figure 4.7 Analysis of the expression levels of the $P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ* and $P_{\text{T-box}\ lysK}$-*lacZ* fusions in a *B. cereus* background in LB. (A) Graph displaying growth and ß-galactosidase activity of NF232.1 [*lysK*::pNF222.1($P_{\text{T-box}\ lysK}$-lacZ) Cm^R] in LB. (B) Graph displaying growth and ß-galactosidase activity of NF259.1 [*asnS*::pNF223.8($P_{\text{T-box}\ asnS}$-lacZ) Cm^R] in LB. For both graphs growth is represented by (◇) and ß-galactosidase activity by (◆).

$tRNA^{ASN}$ charging levels in the cell, on the $P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ* fusion could then be observed. Following a literature search we found the chemical albizziin (L-2-Amino-ureidopropionic acid), which was found to inhibit asparaginyl tRNA synthetase activity and could also be misacylated to $tRNA^{ASN}$ in Chinese hamster ovary cells (Andrulis, *et al.*, 1985). Use of this asparagine analogue to induce expression of the $P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ* fusion was also unsuccessful.

Growth of *B. cereus* in LB containing millimolar concentrations of albizziin was no different to that of *B. cereus* in LB only. Thus, we were unsure if albizzin has an effect on $tRNA^{ASN}$ charging in *B. cereus*. To rule out the possibility that an abundance of asparagine in the LB medium could compete with albizziin and in doing so prevent it from inhibiting AsnRS. we attempted to analyze the effects of albizziin on growth of *B. cereus* in minimal media containing no asparagine. Efforts to grow *B. cereus* in basal limitation medium, M9 medium and Spizizen's minimal medium all failed. In addition to this, *B. subtilis* grown in minimal medium containing various concentrations of albizziin were all unaffected by its addition. *B. subtilis* strains containing the $P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ* and $P_{\text{T-box}\ lysK}$-*lacZ* fuions showed little or no response to the presence of albizziin (data not shown).

The lack of induction of expression of the $P_{B.\ cereus\ \text{T-box}\ asnS}$-*lacZ* fusion to a reduction in $tRNA^{ASN}$ charging in *B. subtilis* indicated that the *B. cereus asnS* T-box regulatory element either may not be functional, or may not be able to interact successfully with $tRNA^{ASN}$ from *B. subtilis*. Attempts to investigate these hypotheses in *B. cereus* were unsuccessful.

Chapter 5

Investigation into the T-box transcriptional regulatory elements controlling aminoacyl tRNA synthetases for amino acids in mixed codon boxes in *B. subtilis*

5.1 Introduction

Our work with the *lysK* T-box regulatory element of *B. cereus* indicates that it is possible (at least in one case) for a T-box regulatory element to recognize more than one tRNA i.e. its cognate tRNA and that tRNA which is charged with the other amino acid in the same mixed codon box. We therefore decided to test the possibility that other T-box regulatory elements present in *B. subtilis* would be unable to discriminate between their cognate tRNA and tRNAs whose anticodons only differ at the third (or wobble) position. There are six T-box regulatory elements in *B. subtilis* that are found upstream of AARS genes whose products charge tRNAs whose cognate amino acids are found in mixed codon boxes (Table 3.2). The T-box regulatory elements tested were those upstream of the *pheS, leuS, ileS, hisS* and *aspS, trpS* and *cysS*.

The first aim was to establish that the putative T-box elements were functional by integrating the relevant T-box leader region-*gfp* transcriptional fusion into a *B. subtilis* strain auxotrophic for that AARS gene's cognate amino acid. In some cases this experiment served to confirm work that had already been carried out. The second aim was to assess the response of the same T-box leader-*gfp* transcriptional fusion to starvation for the other amino acid in that mixed codon box i.e. we analyzed the response of the T-box leader-*gfp* transcriptional fusion to uncharged tRNA whose anticodon matched the T–box regulatory elements cognate tRNA anticodon at two out of three positions.

In order to analyze each T-box regulatory element and its response to uncharged tRNAs we must first define what a T-box response is, i.e. what does induction of expression of a T-box leader-*gfp* transcriptional fusion look like? The criterion that must be fulfilled in order to state that a T-box regulatory element responds to a reduction in the level of charging of a tRNA is a rapid increase in *gfp* expression that is coincident with growth cessation brought about by amino acid limitation. A T-box regulatory element that satisfies this criterion in response to starvation for a particular amino acid and therefore a reduction in charging of the relevant tRNA, can be described as responding to the charging level of that particular tRNA. The units presented in each graph showing *gfp* expression are calculated by dividing the GFP values with background subtracted by the OD_{600} at each time point.

The putative secondary structure of each T-box element analysed is shown in figures 5.1-5.7.

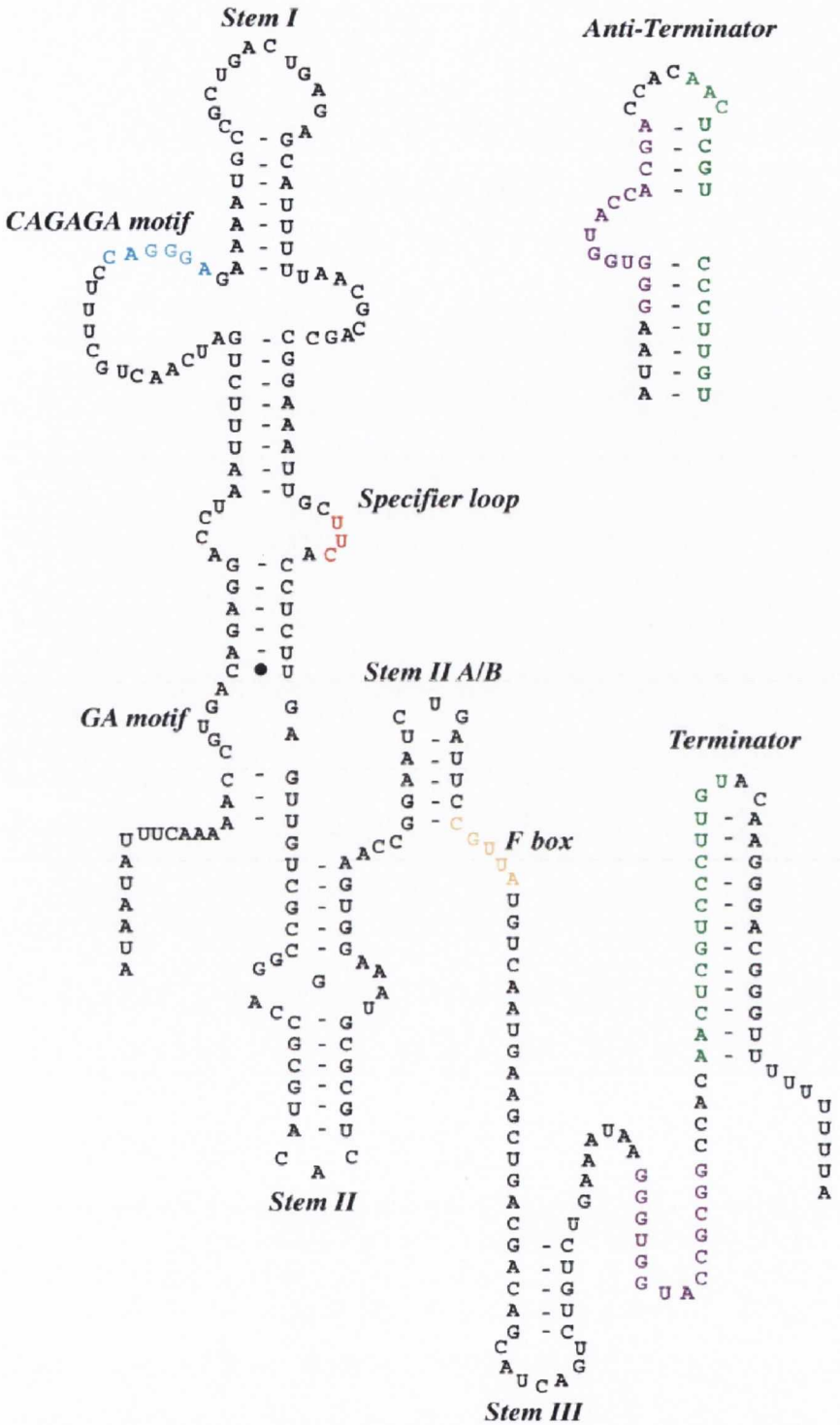


Figure 5.1 Proposed secondary structure of the upstream leader region of the phenylalanyl tRNA synthetase gene of *B. subtilis*. The T-box region is highlighted in purple. Sequences which can form part of either the terminator or antiterminator are indicated in green. The specifier codon is indicated in red. Other important structural features are indicated.

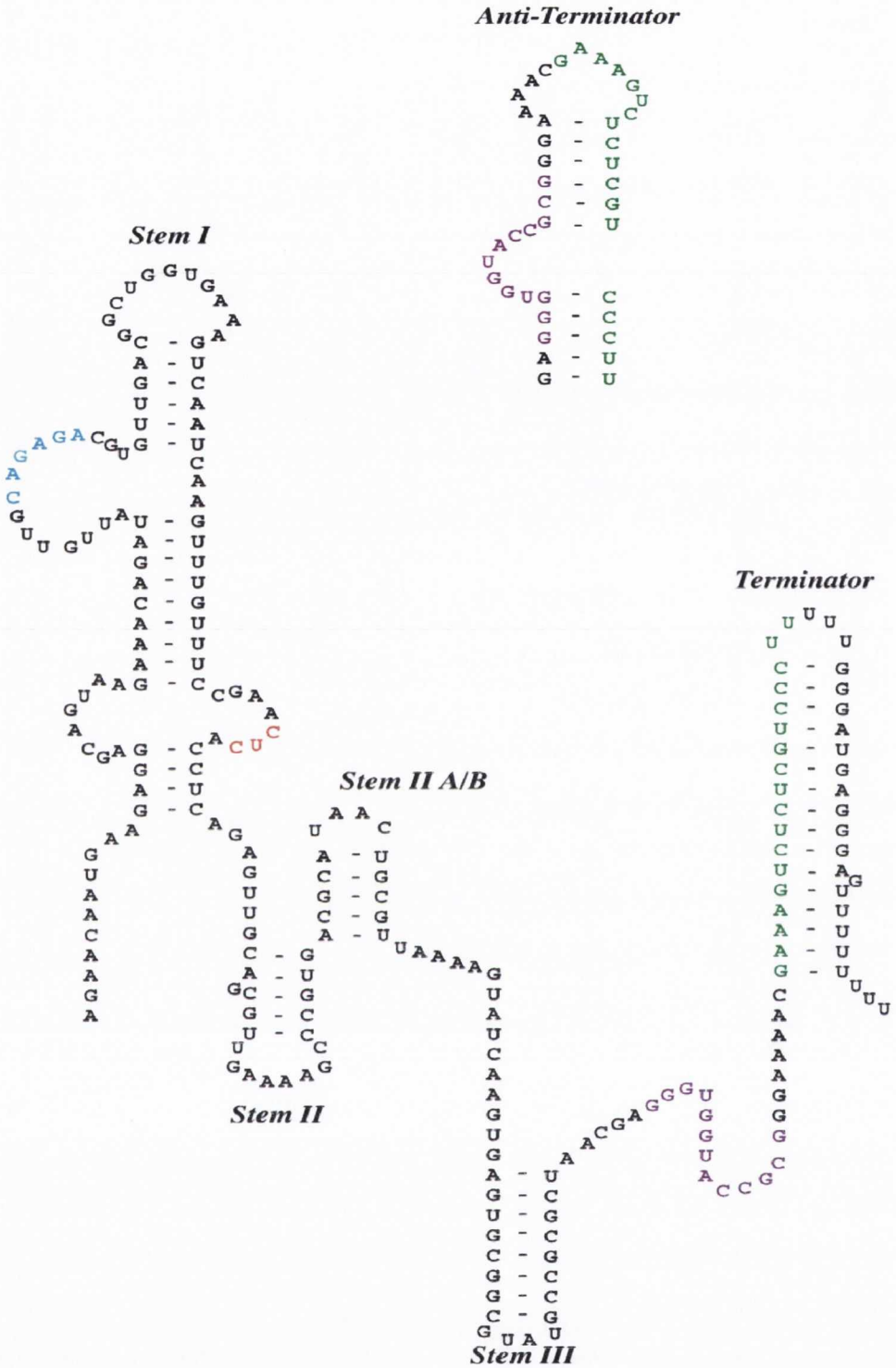


Figure 5.2 Proposed secondary structure of the upstream leader region of the leucyl tRNA synthetase gene of *B. subtilis*. The T-box region is highlighted in purple. Sequences which can form part of either the terminator or antiterminator are indicated in green. The specifier codon is indicated in red. Other important structural features are indicated.

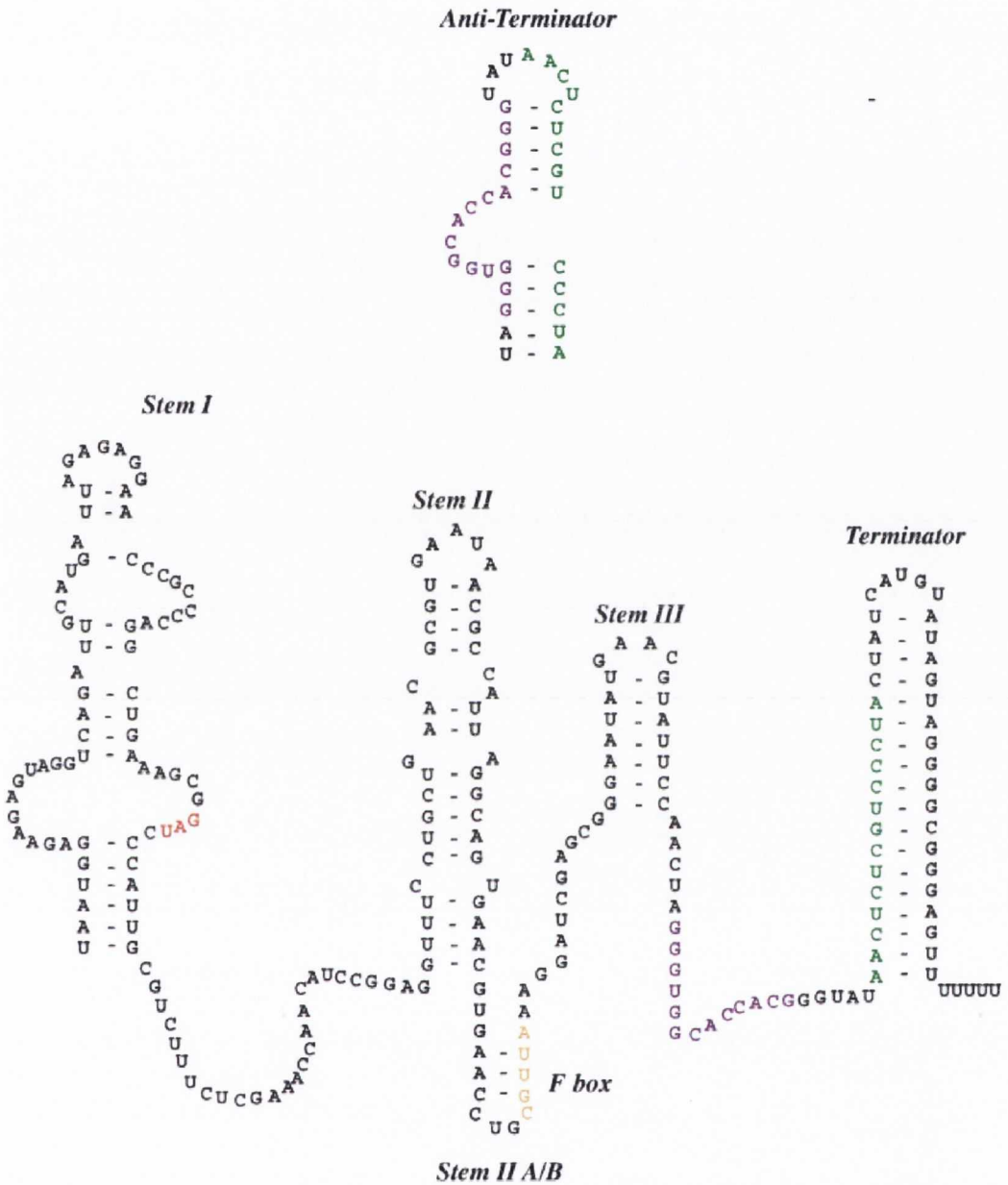


Figure 5.3 Proposed secondary structure of the upstream leader region of the histidyl tRNA synthetase and aspartyl tRNA synthetase genes of *B. subtilis*. The T-box region is highlighted in purple. Sequences which can form either pat of the terminator or antiterminator are indicated in green. The specifier codon is indicated in red. Other important structural features are indicated.

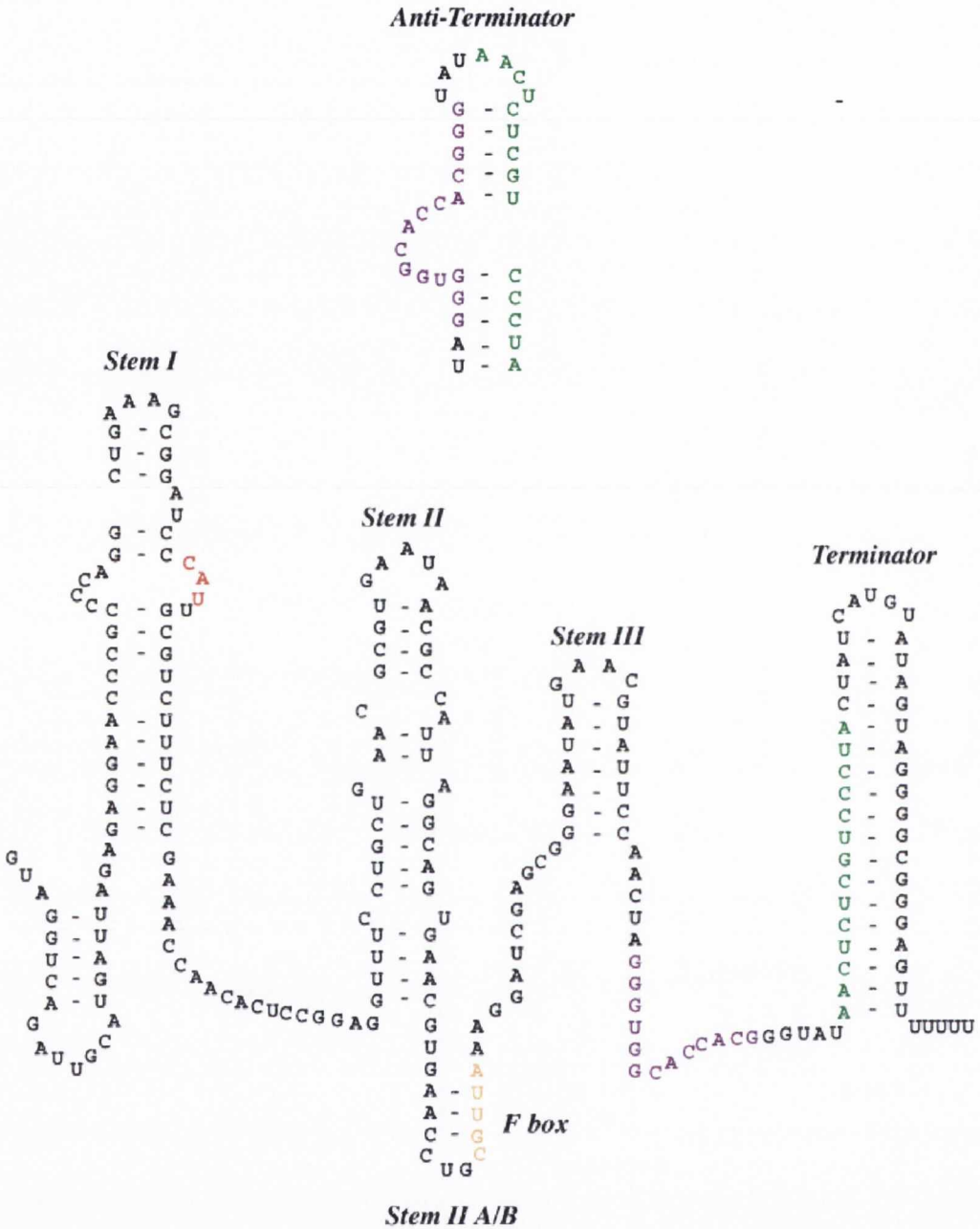


Figure 5.4 Alternative proposed secondary structure of the upstream leader region of the histidyl tRNA synthetase and aspartyl tRNA synthetase genes of *B. subtilis*. The T-box region is highlighted in purple. Sequences which can form either pat of the terminator or antiterminator are indicated in green. The specifier codon is indicated in red. Other important structural features are indicated.

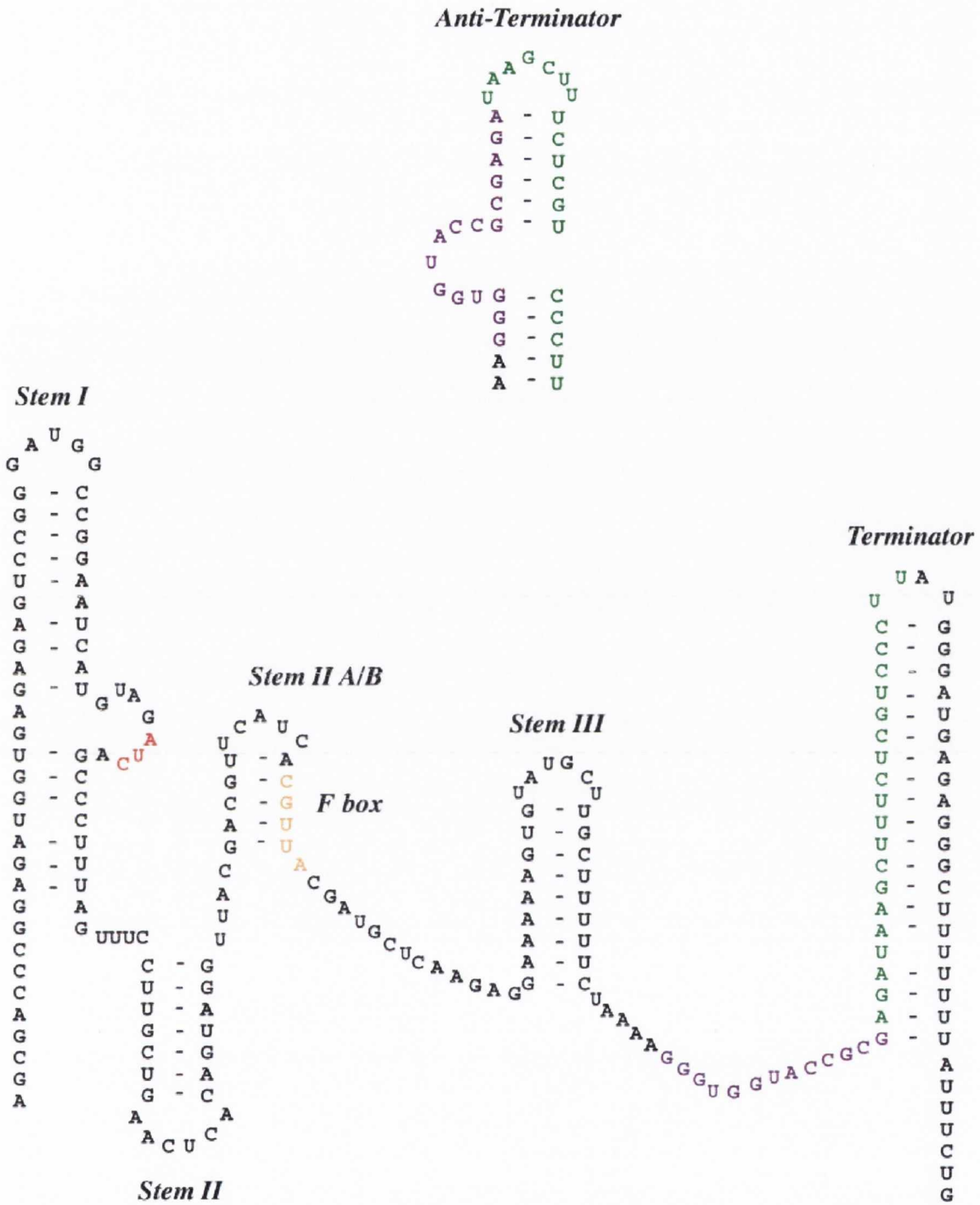


Figure 5.5 Proposed secondary structure of the upstream leader region of the isoleucyl tRNA synthetase gene of *B. subtilis*. The T-box region is highlighted in purple. Sequences which can form either part of the terminator or antiterminator are indicated in green. The specifier codon is indicated in red. Other important structural feaures are indicated.

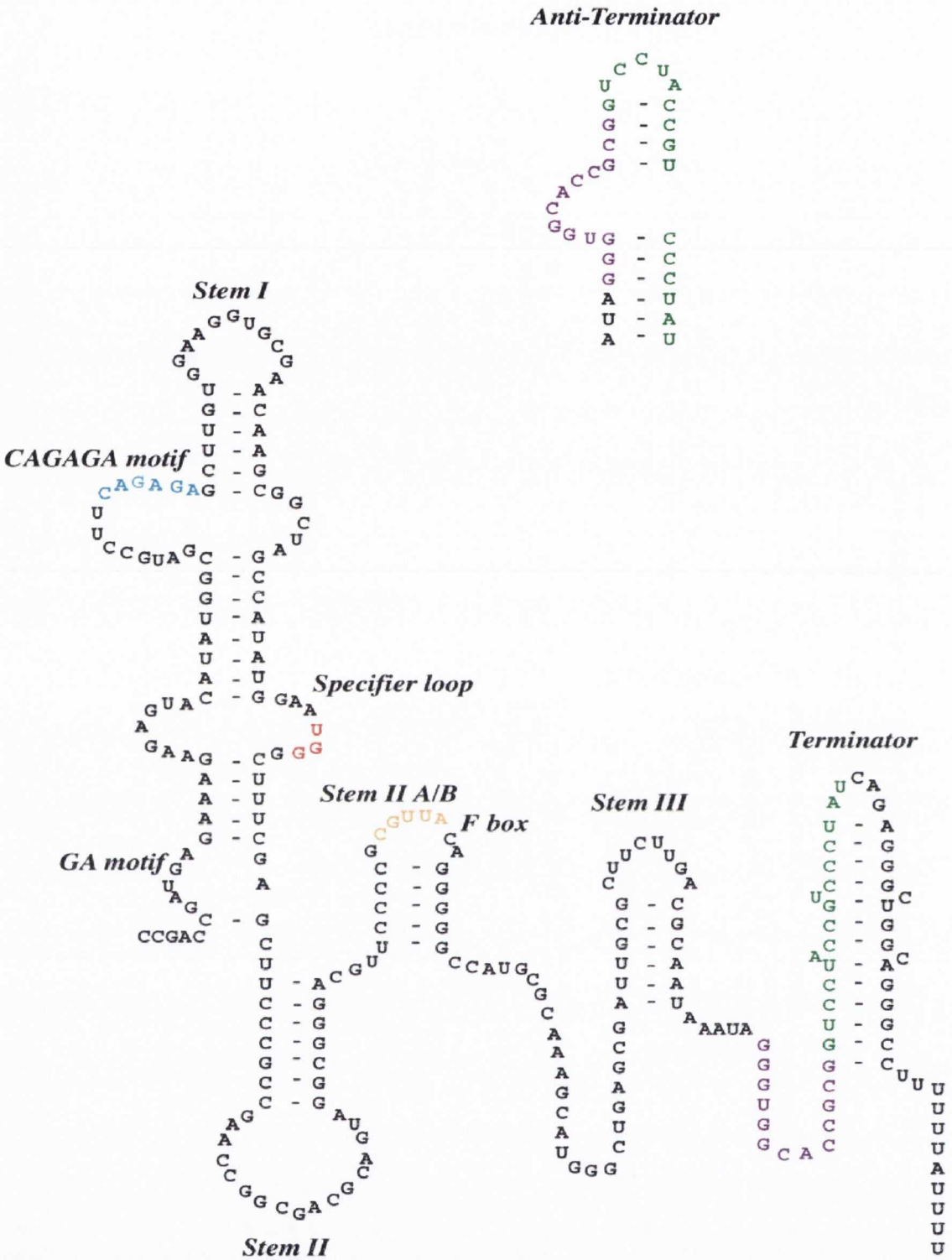


Figure 5.6 Proposed secondary structure of the upstream leader region of the tryptophanyl tRNA synthetase gene of *B. subtilis*. The T-box region is highlighted in purple. Sequences which can form either part of the terminator or antiterminator are indicated in green. The specifier codon is indicated in red. Other important structural features are indicated.

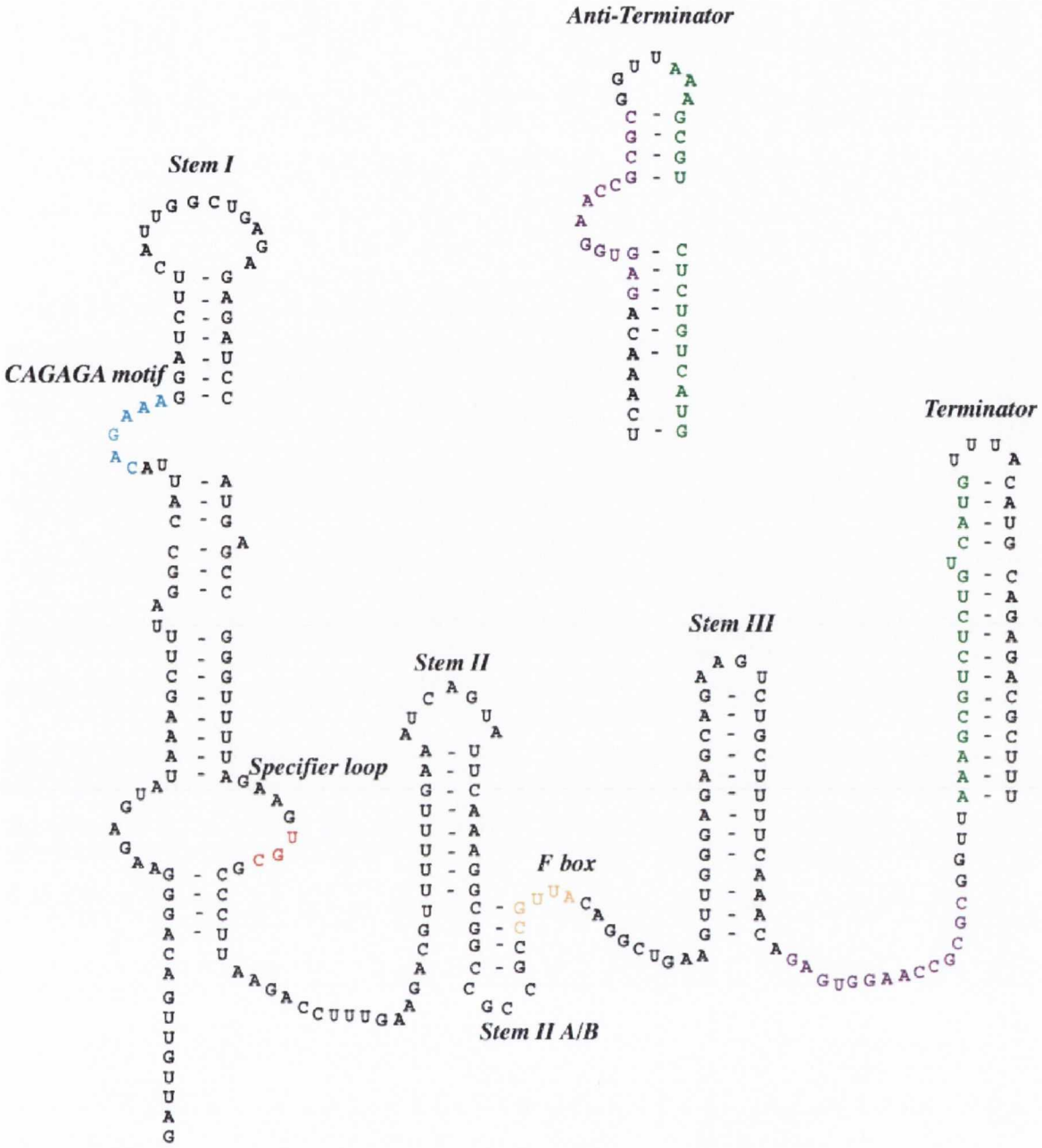


Figure 5.7 Proposed secondary structure of the T-box regulatory element of the serine transacetlase and cysteinyl tRNA synthetase genes of *B. subtilis*. The T-box region is highlighted in purple. Sequences that can form either part of the terminator or antiterminator are indicated in green. The specifier codon is indicated in red. Other important structural features are indicated.

5.2 The *pheS* T-box regulatory element of *B. subtilis*

In order to test the expression of the *pheS* T-box regulatory element of *B. subtilis* we constructed a transcriptional fusion of the promoter and upstream leader region of *pheS* containing the T-box regulatory element to the *gfp* gene ($P_{T\text{-box }pheS}$-*gfp*) and integrated this construct by Campbell type mechanism into the chromosome of *B. subtilis* strains JH642, 1A75 and 1A10 that are auxotrophic for phenylalanine, leucine and threonine respectively. These auxotroph strains were selected to (i) test the response of the *pheS* T-box regulatory element to starvation for its cognate amino acid (ii) test the response of the *pheS* T-box regulatory element to reduced charging of leucyl tRNA (there are 8 $tRNA^{LEU}$ isoacceptors in *B. subtilis,* 4 of which have anticodons that are complementary to the *pheS* T-box regulatory element's specifier codon at two out of three positions) and (iii) act as a control for induction. The $tRNA^{THR}$ anticodon is only complementary to the *pheS* T-box regulatory element's specifier codon at one out of three positions and its discriminator base differs from that of $tRNA^{PHE}$ and should inhibits its potential interaction with the T-box. Experiments to analyze the response of these strains to starvation for these amino acids were carried out as described below.

5.2.1 The T-box regulatory element upstream of the *pheS* gene in *B. subtilis* is functional and sensitive to phenylalanine starvation

It has previously been shown that the *pheS* T-box regulatory element responds to starvation for phenylalanine (Putzer, *et al.*, 1995). To assess the response of the *pheS* T-box regulatory element to starvation for other amino acids we first had to establish what induction of the *pheS* T-box regulatory element looks like using a $P_{T\text{-box }pheS}$-*gfp* fusion. We integrated the $P_{T\text{-box }pheS}$-*gfp* fusion onto the *B. subtilis* chromosome at the *pheS* locus. The resultant strain, NF265.1 was then grown in Basal Limitation medium (BLM) in both phenylalanine replete (50µg/ml) and phenylalanine limiting (4µg/ml) conditions and expression of the $P_{T\text{-box }pheS}$-*gfp* fusion was monitored. The results are shown in figure 5.8.

Growth of strain NF265.1 in 4µg/ml phenylalanine begins to slow and plateau at approximately OD_{600} 0.8 whereas growth in the phenylalanine replete media continues to OD_{600} 4. At the point that growth begins to slow there is a concomitant and rapid increase in expression of the $P_{T\text{-box }pheS}$-*gfp* fusion in the phenylalanine limited NF265.1 culture as the cells become starved for phenylalanine. In the phenylalanine replete culture, no increase in *gfp* expression is observed in response to the slow down in

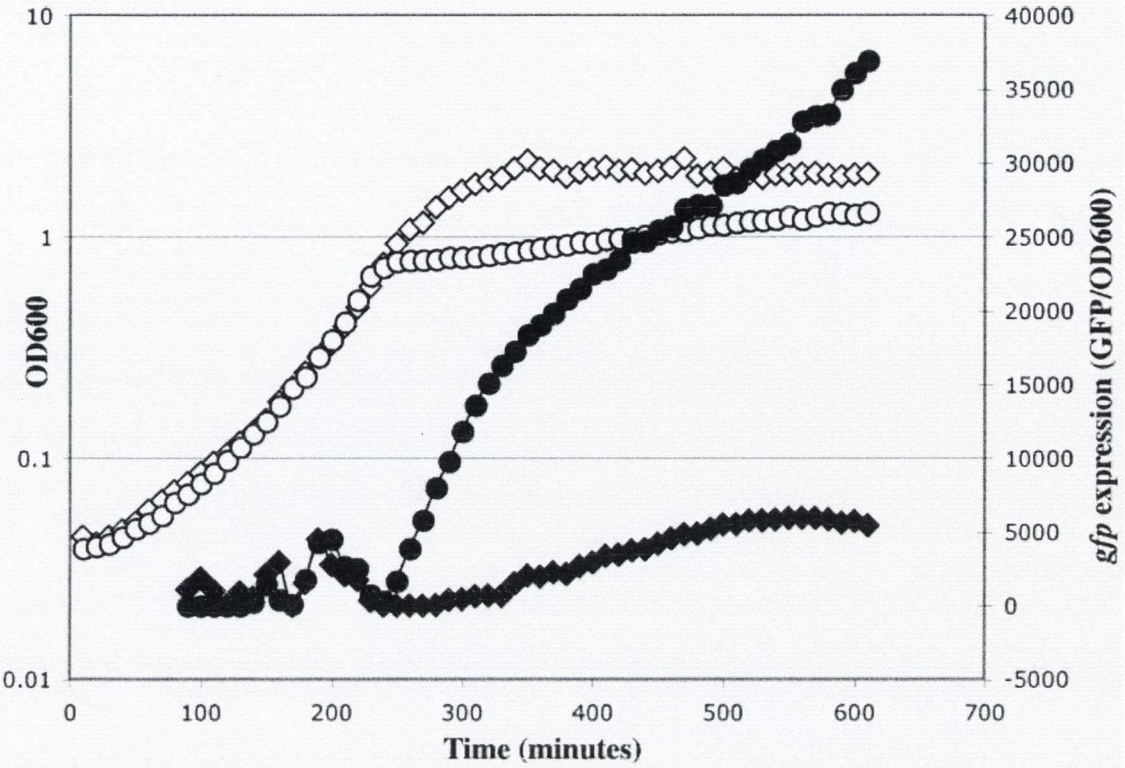


Figure 5.8 Analysis of the response of the *B. subtilis pheS* T-box regulatory element to limitation for phenylalanine. The graph displays growth and *gfp* expression data for NF265.1 [*pheA1 trpC2 pheS*::pNF255.1 ($P_{T\text{-}box\ pheS}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml phenylalanine and 4µg/ml phenylalanine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF265.1 in BLM containing 50µg/ml phenylalanine (◇, ◆). NF265.1 in BLM containing 4µg/ml phenylalanine (○, ●).

growth in stationary phase. This result confirms that the *pheS* T-box regulatory element is responsive to a reduction in $tRNA^{PHE}$ charging (Putzer, *et al.*, 1995).

5.2.2 The T-box regulatory element upstream of the *pheS* gene in *B. subtilis* does not respond to starvation for leucine

As phenylalanine shares a mixed codon box with leucine we wanted to assess whether or not the $P_{T\text{-}box\ pheS}$-*gfp* fusion responded to depletion for leucine. To test this strain NF266.1 was constructed in which the $P_{T\text{-}box\ pheS}$-*gfp* fusion was inserted into the strain 1A75 that is auxotrophic for leucine. This strain was grown in BLM in leucine replete (50μg/ml) and leucine limiting conditions (10μg/ml). The results are presented in figure 5.9.

Growth of strain NF266.1 in 10μg/ml leucine begins to slow and plateau at ~OD_{600} 1.3 whereas growth in the leucine replete media continues to approximately OD_{600} 3. As growth in the leucine starved NF266.1 culture slows there is a slight increase in expression of the $P_{T\text{-}box\ pheS}$-*gfp* fusion from around 1000 units to approximately 5000 units. However, a greater increase in *gfp* expression is seen in the leucine replete culture. Expression levels reach a value of approximately 10000 units in stationary phase. We conclude that the *pheS* T-box regulatory element does not respond to a reduction in the concentration of leucine in the cell.

5.2.3 Mutation of the specifier codon of the *pheS* T-box regulatory element to a stop codon inhibits the response of the system to phenylalanine starvation

In order to further demonstrate that the *pheS* T-box mechanism is sensitive to the level of uncharged $tRNA^{PHE}$ in the cell we mutated the specifier codon on Stem I from UUC to the Stop codon UAA. This mutation should remove or reduce the ability of the $tRNA^{PHE}$ anticodon to interact with the specifier loop. The mutated $P_{T\text{-}box\ pheS}$-*gfp* fusion was then integrated into the *pheS* locus of the phenylalanine auxotroph strain JH642 by a single crossover to create the strain NF305.1. This strain was then grown in BLM containing either 50μg/ml phenylalanine (phenylalanine replete) or 4μg/ml phenylalanine (phenylalanine limiting) and the *gfp* expression was analyzed (Figure 5.10).

Growth of strain NF305.1 in 4μg/ml phenylalanine begins to slow and plateau at approximately OD_{600} 0.8 whereas growth in the phenylalanine replete media continues to OD_{600} 2. At the point that growth begins to slow there is a concomitant increase in

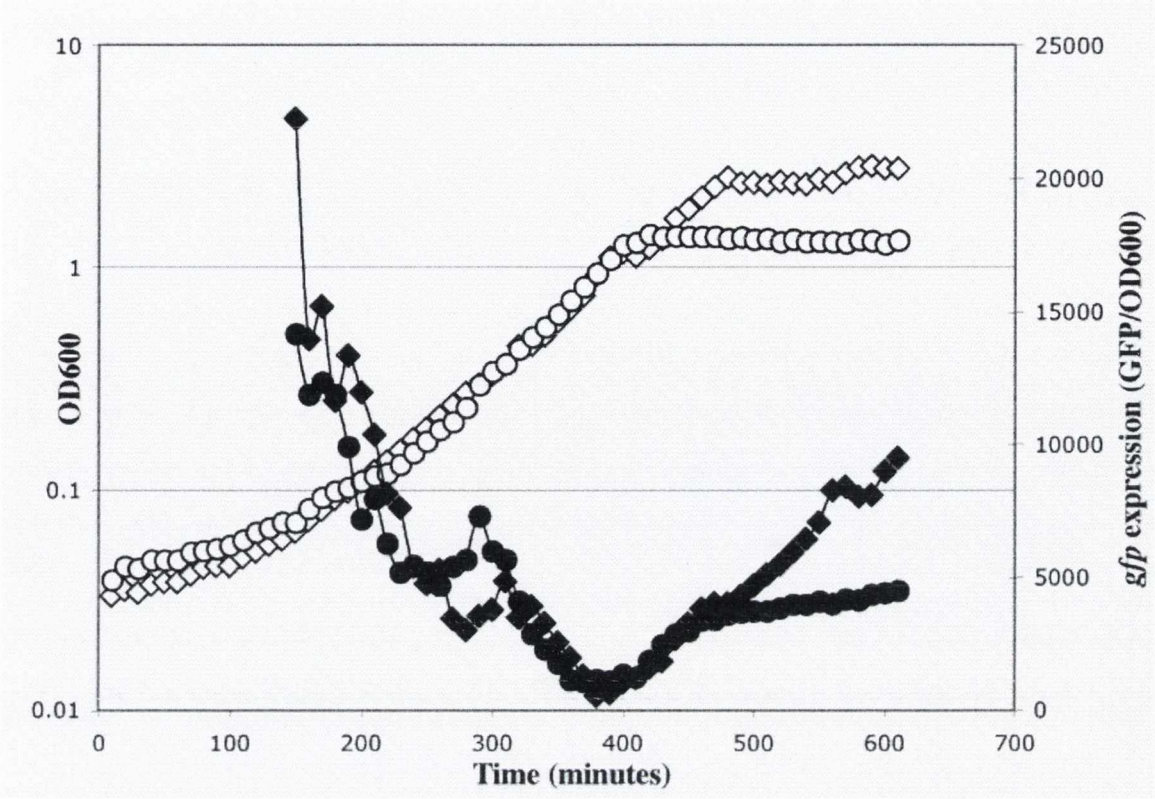


Figure 5.9 Analysis of the response of the *B. subtilis pheS* T-box regulatory element to starvation for leucine. The graph displays growth and *gfp* expression data for NF266.1 [*ilvA1 leuB8 metB5*::pNF255.1 ($P_{T\text{-box }pheS}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml leucine and 10µg/ml leucine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF266.1 in BLM containing 50µg/ml leucine (◇, ◆). NF266.1 in BLM containing 10µg/ml leucine (○, ●).

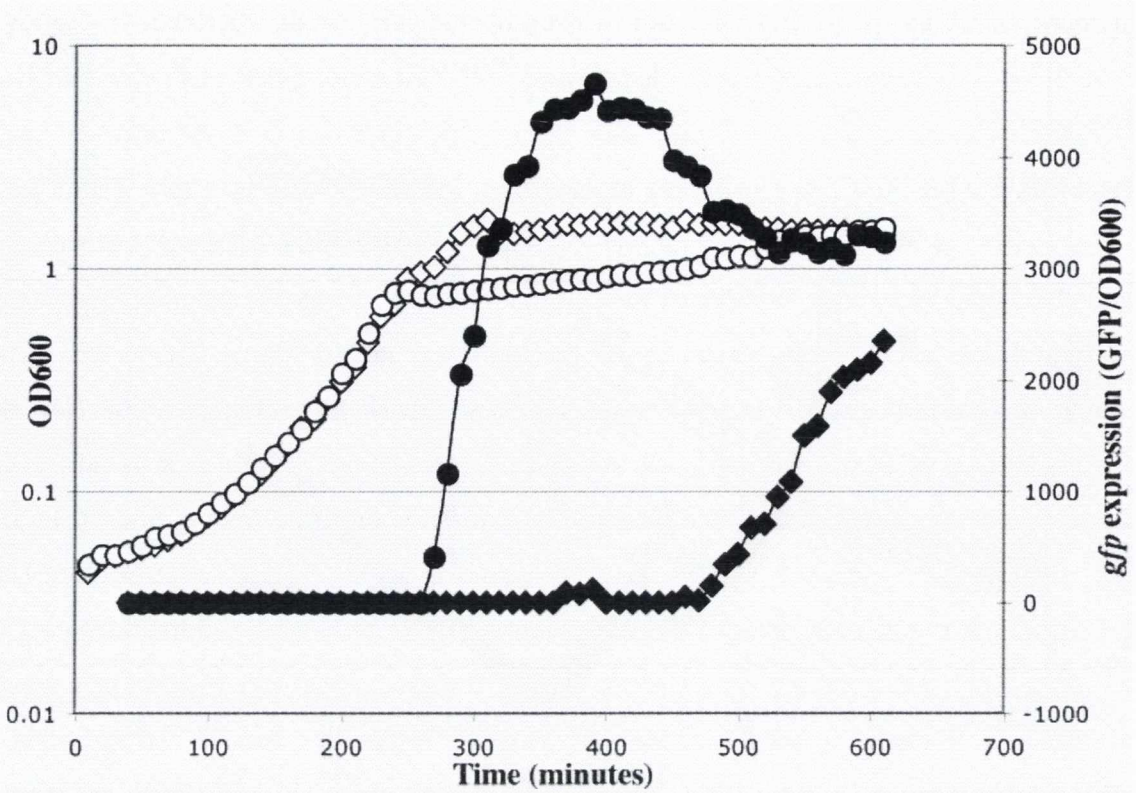


Figure 5.10 Analysis of the response of the *B. subtilis pheS* T-box regulatory element containing a UAA specifier codon to limitation for phenylalanine. The graph displays growth and *gfp* expression data for NF305.1 [*pheA1 trpC2 pheS*::pNF289.1 ($P_{T\text{-}box\ pheS\ STOP}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml phenylalanine and 4µg/ml phenylalanine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF305.1 in BLM containing 50µg/ml phenylalanine (◇, ◆). NF305.1 in BLM containing 4µg/ml phenylalanine (○, ●).

expression of the $P_{T\text{-}box\ pheS}$-*gfp* fusion in the phenylalanine limited NF265.1 culture as the cells become starved for phenylalanine. The increase in expression is 7.5 fold less than that seen in the wild-type $P_{T\text{-}box\ pheS}$-*gfp* fusion as a result of phenylalanine starvation (Figure 5.8). In the phenylalanine replete culture, no increase in *gfp* expression is observed in response to the slow down in growth in stationary phase.

The results show that $tRNA^{PHE}$ can still interact with the specifier codon in spite of the mutation to a UAA stop codon, albeit much less efficiently than with the wild-type UUC specifier codon. This result shows that while the specifier codon-anticodon interaction is important for antitermination, other interactions of the tRNA with the T-box regulatory element can support a successful interaction of $tRNA^{PHE}$ with the T-box regulatory element despite a suboptimal interaction between the tRNA's anticodon and the specifier codon.

5.2.4 Expression of *pheS* is induced by starvation for threonine

In order to assess whether or not the *pheS* T-box regulatory mechanism was sensitive to more general amino acid starvation we integrated the $P_{T\text{-}box\ pheS}$-*gfp* fusion by single crossover into the chromosome of the threonine auxotroph strain 1A10 to create the strain NF295.1. This strain was then grown in BLM containing 100μg/ml threonine (threonine replete) and 5μg/ml threonine (threonine limiting) and the expression of the $P_{T\text{-}box\ pheS}$-*gfp* fusion was analyzed. The results for this experiment are shown in figure 5.11.

Growth of strain NF295.1 in 5μg/ml threonine begins to slow and plateau at approximately OD_{600} 0.8 whereas growth in the threonine replete media continues to OD_{600} 4. As strain NF295.1 becomes starved for threonine at approximately 250 minutes there is a five fold increase in the amount of *gfp* expression from 2000 units to approximately 10,000 units over a period of 100 minutes. This contrasts with cells growing in high concentrations of threonine whose *gfp* expression levels are constantly between 1000 and 3000 units throughout the time course. It is unlikely that uncharged threonyl tRNA can interact with the *pheS* T-box regulatory element to successfully stabilize formation of the anti-terminator. It is possible that a reduction in threonine concentration can bring about an indirect reduction in the level of $tRNA^{PHE}$ charging in the cell or that limitation for threonine results in increased transcription initiation from the *pheS* promoter. The cause of the increase in expression of the $P_{T\text{-}box\ pheS}$-*gfp* fusion as a result of limitation for threonine requires further analysis.

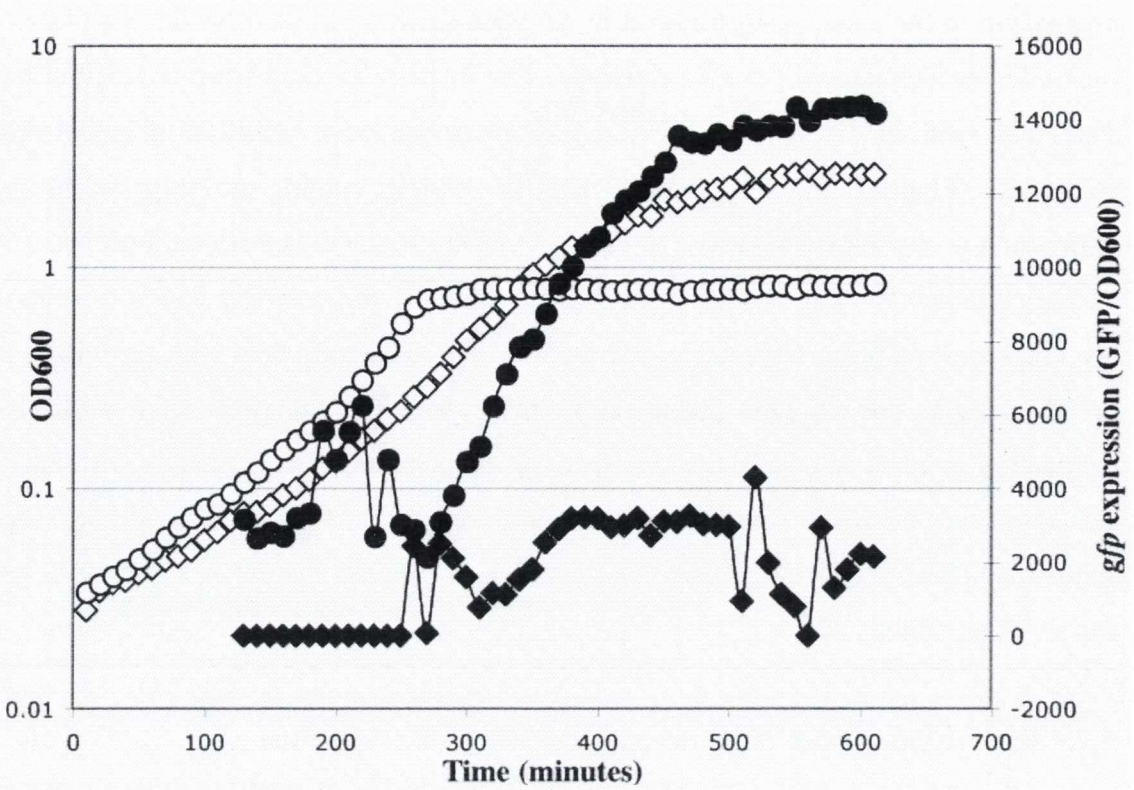


Figure 5.11 Analysis of the response of the *B. subtilis pheS* T-box regulatory element to starvation for threonine. The graph displays growth and *gfp* expression data for NF295.1 [*hisA1 trpC2 thr5*::pNF255.1 ($P_{T\text{-}box\ pheS}$-*gfp*) $Spec^R$] grown in BLM containing 100µg/ml threonine and 5µg/ml threonine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF295.1 in BLM containing 100µg/ml threonine (◇, ◆). NF295.1 in BLM containing 5µg/ml threonine (○, ●).

5.3 The *leuS* T-box regulatory element of *B. subtilis*

In order to characterize the expression of the *leuS* T-box regulatory element of *B. subtilis* we constructed a transcriptional fusion of the promoter and upstream leader region of *leuS* containing the T-box regulatory element to the GFP reporter ($P_{T\text{-box }leuS}$-*gfp*) and integrated this construct at the *leuS* locus by Campbell type mechanism into the chromosome of *B. subtilis* auxotroph strains JH642 (phenylalanine), 1A75 (leucine, isoleucine, methionine) and 1A10 (threonine). These auxotroph strains were selected to (i) test the response of the *leuS* T-box regulatory element to starvation for its cognate amino acid (ii) test the response of the *leuS* T-box regulatory element to reduced charging of phenylalanyl tRNA ($tRNA^{PHE}$ is complementary to the *leuS* T-box regulatory element's specifier codon at two out of three positions) and (iii) act as a control for induction. The $tRNA^{THR}$ anticodon is only complementary to the *leuS* T-box regulatory element's specifier codon at one out of three positions and its discriminator base differs from that of $tRNA^{LEU}$ and should inhibits its potential interaction with the T-box. Experiments to analyze the response of these strains to starvation for these amino acids were carried out as described below

5.3.1 The T-box regulatory element of the *leuS* gene of *B. subtilis* responds poorly to leucine starvation

To establish whether the *leuS* T-box regulatory element responds to leucine starvation we integrated the ($P_{T\text{-box }leuS}$-*gfp*) fusion into the chromosome of the *B. subtilis* leucine auxotroph strain 1A75. The resultant strain, NF261.1 was grown in BLM containing either 50μg/ml leucine (leucine replete) or 10μg/ml leucine (leucine limiting) and its *gfp* expression was determined. Results are shown in figure 5.12.

Growth of strain NF261.1 in 10μg/ml leucine begins to slow and plateau at approximately OD_{600} 0.8 whereas growth in the leucine replete media continues to OD_{600} 2. At the point of leucine starvation where growth plateaus, there is an increase in *gfp* expression values from ~ 0 to ~ 3000 units (Figure 5.12). However when the leucine rich culture enters stationary phase a greater response in expression is (0 to approximately 6000 units) is observed. This indicates leucine depletion and hence charged $tRNA^{LEU}$ does not result in increased transcription of the ($P_{T\text{-box }leuS}$-*gfp*) fusion. This shows that the *leuS* T-box regulatory element does not respond to starvation for leucine under the conditions tested.

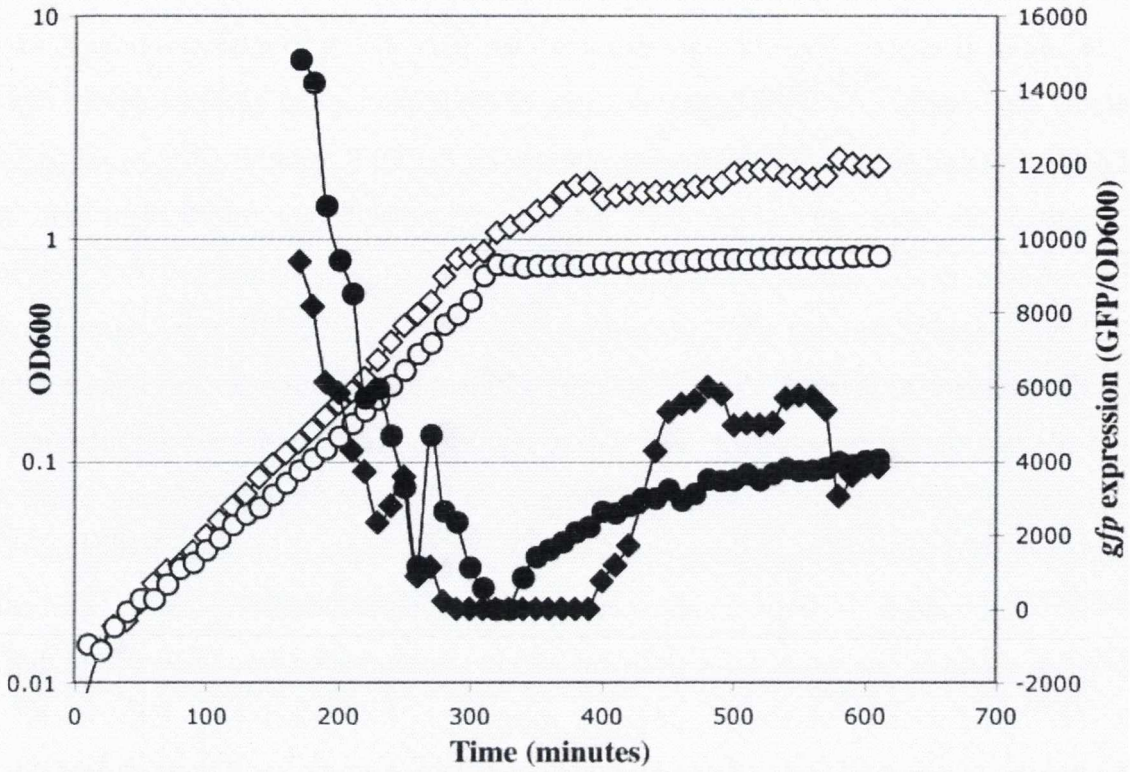


Figure 5.12 Analysis of the response of the *B. subtilis leuS* T-box regulatory element to starvation for leucine. The graph displays growth and *gfp* expression data for NF261.1 [*ilvA1 leuB8 metB5*::pNF249.4 ($P_{T\text{-}box\ leuS}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml leucine and 10µg/ml leucine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF261.1 in BLM containing 50µg/ml leucine (◇, ◆). NF261.1 in BLM containing 10µg/ml leucine (○, ●).

5.3.2 Mutation of the specifier codon of the *leuS* T-box regulatory element to a stop codon indicates that the *leuS* T-box mechanism is not responsive to the level of charged $tRNA^{LEU}$ in the cell

No increase in *gfp* expression was observed in strain NF261.1 ($P_{T\text{-}box\ leuS}$-*gfp*) in response to starvation for leucine. We wanted to test whether the ratio of charged to uncharged $tRNA^{LEU}$ played a role in the transcriptional regulation of leucyl tRNA synthetase. To investigate this we mutated the CUC specifier codon of the $P_{T\text{-}box\ leuS}$-*gfp* fusion to an UAA ochre stop codon ($P_{T\text{-}box\ leuS\ STOP}$-*gfp*) and integrated this construct into the chromosome of the *B. subtilis* leucine auxotroph strain 1A75 to create the strain NF306.1. Strain NF306.1 was grown in BLM in leucine replete and leucine limiting conditions and expression levels were established. The results are presented in figure 5.13.

Growth of strain NF306.1 in 10µg/ml leucine (leucine limiting) begins to slow and plateau at approximately OD_{600} 0.8 whereas growth in the leucine replete media continues to ~ OD_{600} 2. At the point of leucine starvation where growth plateaus, there is an increase in *gfp* expression values from ~ 0 to ~ 5000 units over a period of 100 minutes (Figure 5.13). However, when the leucine rich culture enters stationary phase a greater response in expression (0 to approximately 10000 units) is observed over the same time period. These results show a higher level of basal *gfp* expression in the $P_{T\text{-}box\ leuS\ STOP}$-*gfp* fusion than that observed in strain NF261.1 containing the wild type ($P_{T\text{-}box\ leuS}$ -*gfp* fusion (Figure 5.12). These results also show a similar profile of *gfp* expression to that seen for strain NF261.1 grown in the same conditions (Figure 5.12). This data shows that mutation of the specifier codon to a stop codon does not result in reduced expression of the $P_{T\text{-}box\ leuS}$ –*gfp* fusion as would be expected in a functional T-box regulatory element. This data combined with the data for strain NF261.1 suggests that the *leuS* T-box regulatory element does not respond to starvation for leucine and that regulation of the *leuS* gene of *B. subtilis* does not occur through a conventional T-box mechanism.

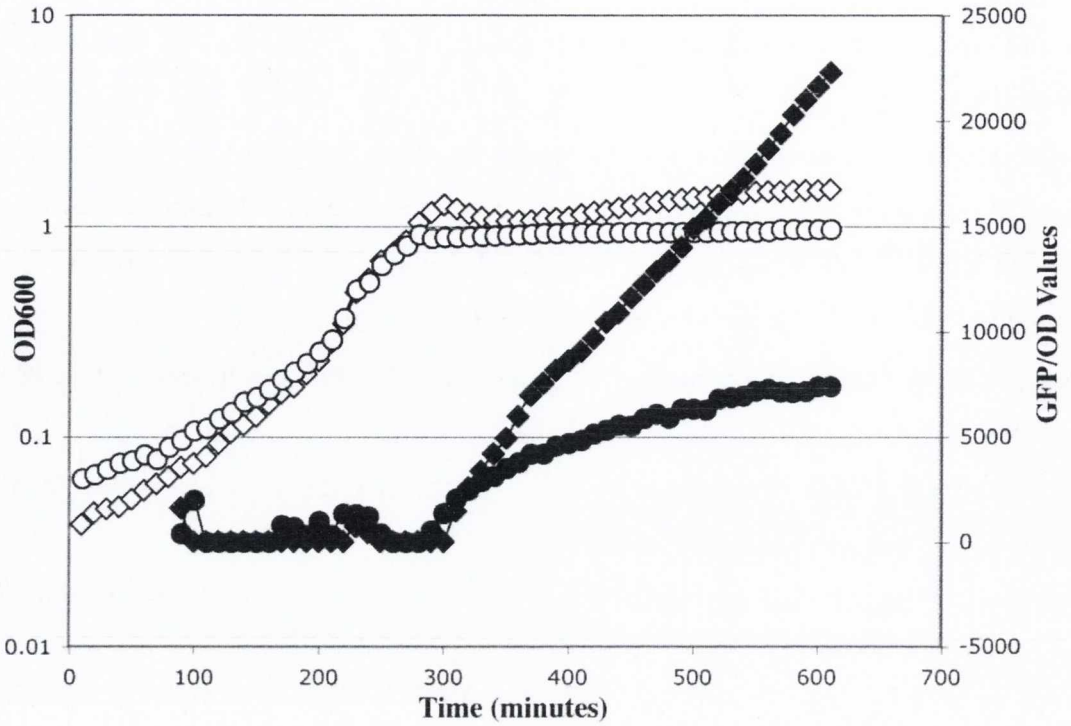


Figure 5.13 Analysis of the response of the *B. subtilis leuS* T-box regulatory element containing a UAA specifier codon to starvation for leucine. The graph displays growth and *gfp* expression data for NF306.1 [*ilvA1 leuB8 metB5*::pNF288.2 ($P_{T\text{-}box\ leuS\ STOP}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml leucine and 10µg/ml leucine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF306.1 in BLM containing 50µg/ml leucine (◇, ◆). NF306.1 in BLM containing 10µg/ml leucine (○, ●).

5.3.3 The *leuS* T-box regulatory element does not respond to starvation for phenylalanine

In order to assess the possible lack of specificity in tRNA recognition by the *leuS* T-box regulatory element the $P_{T\text{-}box\ leuS}$-*gfp* fusion was integrated into the chromosome of the *B. subtilis* strain JH642 (*pheA2*) producing the strain NF260.1. This strain would allow us to examine the effects of starvation for phenylalanine on the $P_{T\text{-}box\ leuS}$-*gfp* fusion and therefore assess whether uncharged $tRNA^{PHE}$ can interact with the *leuS* T-box regulatory element. Strain NF260.1 was grown in BLM in phenylalanine replete and phenylalanine limitation conditions as described previously and the effect on *gfp* expression was determined. The results are shown in figure 5.14.

The results show that at the point of growth cessation due to phenylalanine limitation there is a slight increase in *gfp* expression from 0 to 6000 units over the course of 100 minutes (Figure 5.14). This increase in *gfp* expression is gradual and not the rapid increase in expression characteristic of a T-box regulatory elements response to its cognate uncharged tRNA. A similar slow increase in expression is seen as the phenylalanine replete culture enters stationary phase. This indicates that this response is due to a general slow down in growth and is not a result of specific phenylalanine starvation. We conclude therefore that the *leuS* T-box regulatory element does not respond to starvation for phenylalanine.

5.3.4 Expression of the $P_{T\text{-}box\ leuS}$-*gfp* fusion is induced by starvation for threonine but not methionine or isoleucine

In testing the specificity of the response of the $P_{T\text{-}box\ leuS}$-*gfp* fusion to leucine starvation we also tested the response of the $P_{T\text{-}box\ leuS}$-*gfp* fusion to starvation for isoleucine, methionine and threonine as negative controls.

Strain NF261.1 ($P_{T\text{-}box\ leuS}$-*gfp*) is a leucine, isoleucine and methionine auxotroph. Thus it was possible to assess the effect of limitation for each of these three amino acids in this strain. In separate experiments we analyzed the effects of starvation for methionine and isoleucine on $P_{T\text{-}box\ leuS}$-*gfp* expression in strain NF261.1 (Figures 5.15 & 5.16).

Strain NF261.1 was grown in BLM containing either 50µg/ml methionine (methionine replete) or 1µg/ml methionine (methionine limiting) and its *gfp* expression was determined. The results are shown in figure 5.15.

Growth of strain NF261.1 in 1µg/ml methionine begins to slow and plateau at approximately OD_{600} 1.2 at ~ 280 minutes whereas growth in the methionine replete media

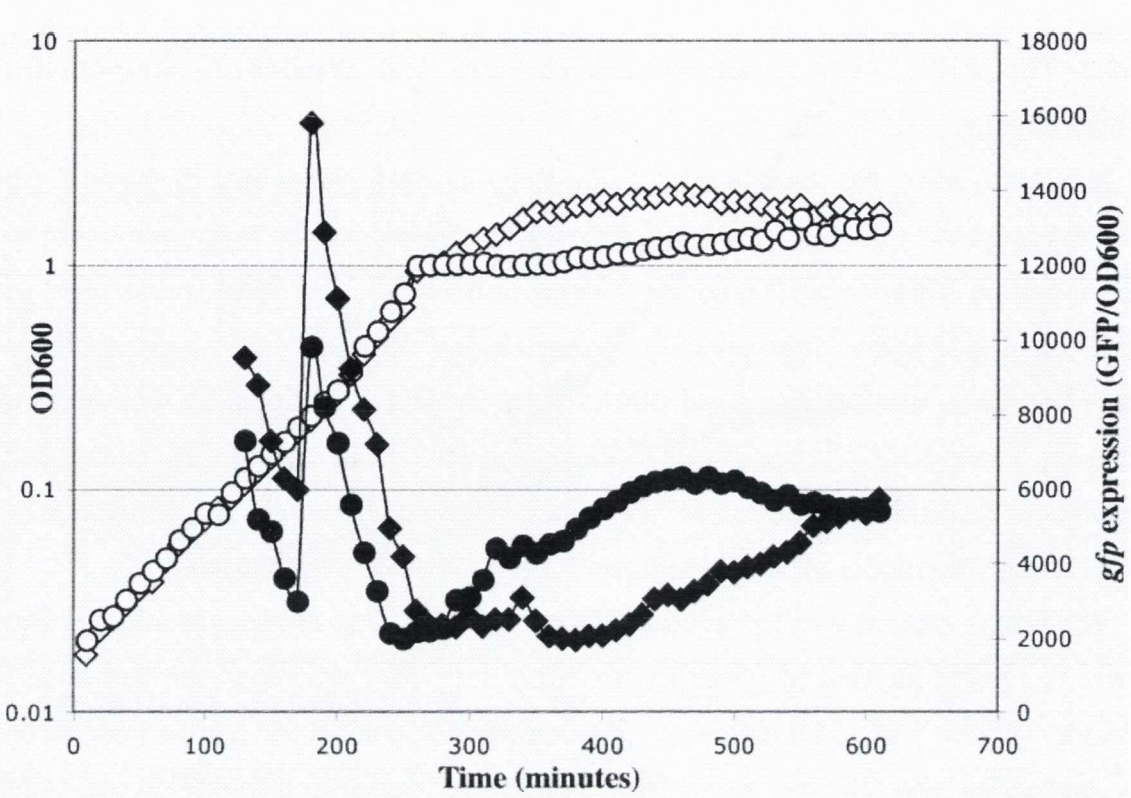


Figure 5.14 Analysis of the response of the *B. subtilis leuS* T-box regulatory element to starvation for phenylalanine. The graph displays growth and *gfp* expression data for NF260.1 [*pheA1 trpC2*::pNF249.4 ($P_{T\text{-}box\ leuS}$-*gfp*) $Spec^R$] grown in BLM containing 50μg/ml phenylalanine and 4μg/ml phenylalanine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF260.1 in BLM containing 50μg/ml phenylalanine (◇, ◆). NF260.1 in BLM containing 4μg/ml phenylalanine (○, ●).

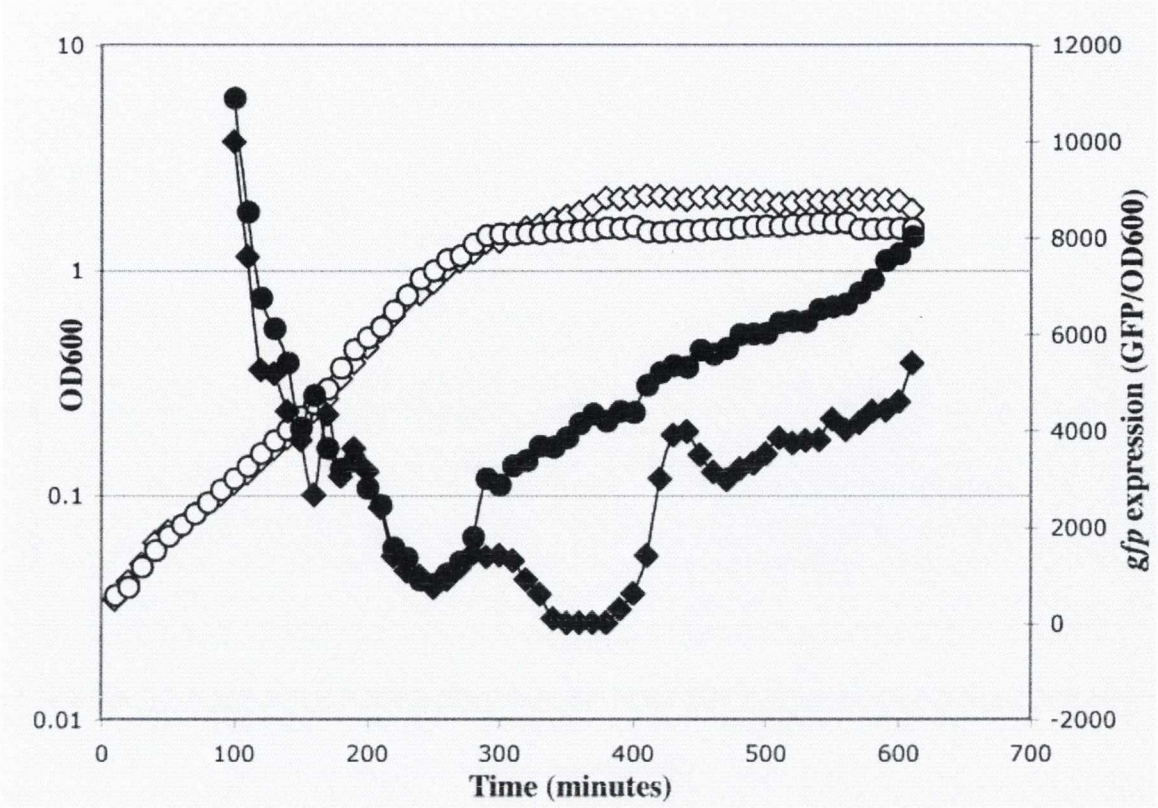


Figure 5.15 Analysis of the response of the *B. subtilis leuS* T-box regulatory element to starvation for methionine. The graph displays growth and *gfp* expression data for NF261.1 [*ilvA1 leuB8 metB5*::pNF249.4 ($P_{T\text{-}box\ leuS}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml methionine and 1µg/ml methionine . Growth is represented by open symbols. Expression values are represented by closed symbols. NF261.1 in BLM containing 50µg/ml methionine (◇, ◆). NF261.1 in BLM containing 1µg/ml methionine (○, ●).

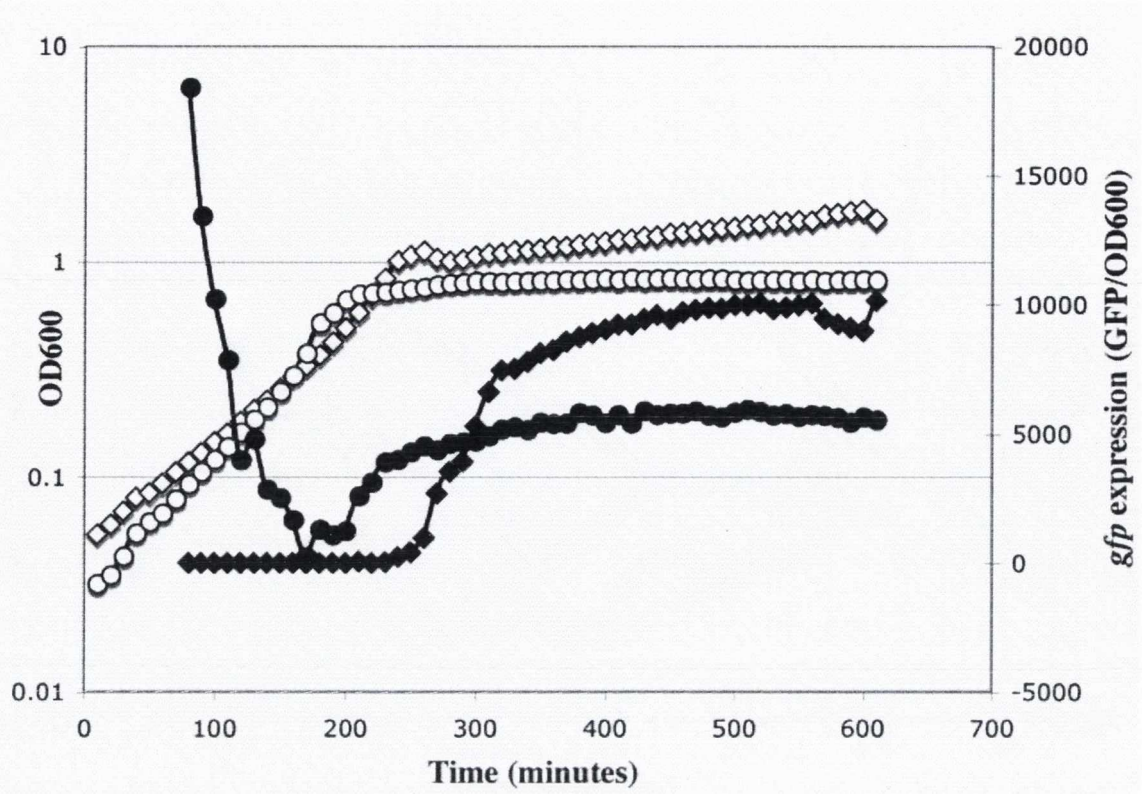


Figure 5.16 Analysis of the response of the *B. subtilis leuS* T-box regulatory element to starvation for isoleucine. The graph displays growth and *gfp* expression data for NF261.1 [*ilvA1 leuB8 metB5*::pNF249.4 ($P_{T\text{-}box\ leuS}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml isoleucine and 3µg/ml isoleucine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF261.1 in BLM containing 50µg/ml isoleucine (◇, ◆). NF261.1 in BLM containing 3µg/ml isoleucine (○, ●).

continues to OD_{600} 2. At the point of methionine starvation where growth plateaus, there is an increase in *gfp* expression from ~ 2500 to ~ 4000 units over a period of 100 minutes (Figure 5.15). This is a slow increase in expression indicating that it is not due to interaction of uncharged methionyl tRNA with the *leuS* T-box regulatory element.

To analyze the effect of starvation for isoleucine on the $P_{T\text{-}box\ leuS}$-*gfp* fusion we grew strain NF261.1 in BLM containing either 50µg/ml isoleucine (isoleucine replete) or 3µg/ml isoleucine (isoleucine limiting) and its *gfp* expression was determined. The results are shown in figure 5.16.

Growth of strain NF261.1 in 3µg/ml isoleucine begins to slow and plateau at OD_{600} 0.8 at ~200 minutes after inoculation. At this point there is an increase in *gfp* expression from ~500 units to 4000 units. This increase in *gfp* expression is also observed in the isoleucine replete culture indicating that this increase in *gfp* expression is not due to starvation for charged isoleucyl tRNA but is instead a result of a general slow down in growth.

Taken together these results show that the *leuS* T-box regulatory element of *B. subtilis* does not respond to starvation for methionine or isoleucine.

To analyze the effect of starvation for threonine on the *leuS* T-box regulatory element we integrated the $P_{T\text{-}box\ leuS}$-*gfp* fusion into the chromosome of the threonine auxotroph 1A10 to produce the strain NF292.1. We grew strain NF292.1 in BLM containing either 100µg/ml threonine (threonine replete) or 5µg/ml threonine (threonine limiting) and its *gfp* expression was determined. The results are shown in figure 5.17.

Growth of strain NF292.1 begins to slow and plateau at OD_{600} 0.9 at ~ 200 minutes after inoculation. At this point there is a rapid increase in *gfp* expression from ~ 2000 units to approximately 12000 units within 120 minutes. As the strain enters stationary phase under threonine replete conditions there is also an increase in expression. However, the increased expression is slow and gradual doubling from 4000 units to 8000 units over a period of 140 minutes. These results show a reduction in threonine availability causes an increase in expression of the $P_{T\text{-}box\ leuS}$-*gfp* fusion. This increase is unlikely to occur by interaction of uncharged $tRNA^{THR}$ with the *leuS* T-box regulatory element as we have shown that the mutation of the specifier codon has no effect on expression of the $P_{T\text{-}box\ leuS}$-*gfp* fusion indicating that uncharged tRNA does not play a role in the regulation of *leuS* expression in *B. subtilis*.

5.3.5 Conclusions about the *leuS* and *pheS* T-box regulatory elements

We assessed the expression of transcriptional fusions containing the promoter and T-box leader regions of the *pheS* and *leuS* genes of *B. subtilis* in order to assess their response to a

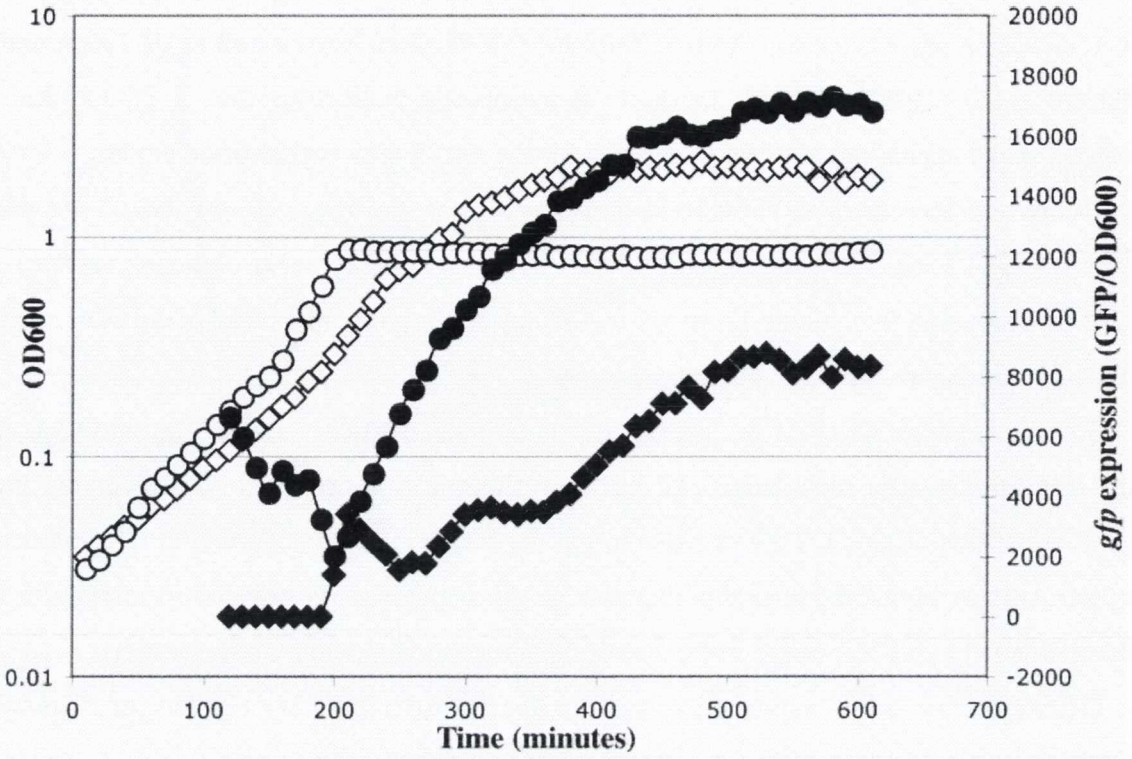


Figure 5.17 Analysis of the response of the *B. subtilis leuS* T-box regulatory element to starvation for threonine. The graph displays growth and *gfp* expression data for NF292.1 [*hisA1 thr5 trpC2*::pNF249.4 ($P_{T\text{-}box\ leuS}$-*gfp*) $Spec^R$] grown in BLM containing 100µg/ml threonine and 5µg/ml threonine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF292.1 in BLM containing 100µg/ml threonine (◇, ◆). NF292.1 in BLM containing 5µg/ml threonine (○, ●).

reduction in charging of their cognate tRNAs and also to establish if it was possible for $tRNA^{PHE}$ and $tRNA^{LEU}$ which share 2 out of 3 positions at the anticodon, to interact with the *leuS* and *pheS* T-box regulatory elements respectively.

Our results confirm that the *pheS* T-box regulatory element is functional and responds to the level of uncharged $tRNA^{PHE}$. The results also show that uncharged $tRNA^{LEU}$ cannot induce expression of a $P_{T\text{-}box\ pheS}$-*gfp* fusion. However, the data shows that expression of *pheS* is induced by starvation for threonine. The mechanism by which the response to threonine occurs is not clear.

Our analysis of the $P_{T\text{-}box\ leuS}$-*gfp* fusion shows that the putative T-box regulatory element upstream of *leuS* does not function by a canonical T-box mechanism. The data shows that the $P_{T\text{-}box\ leuS}$-*gfp* fusion does not respond to a reduction in the level of uncharged $tRNA^{LEU}$. Nor is there an induction of *gfp* expression as a result of limitation for phenylalanine, methionine or isoleucine. However, there is a response to starvation for threonine in the $P_{T\text{-}box\ leuS}$-*gfp* fusion. The observation of a response to threonine starvation in both *leuS* and *pheS* expression suggests that starvation for threonine brings about a stress response in the cell that results in increased expression of *pheS* and *leuS*.

5.4 The *hisS-aspS* T-box regulatory element of *B. subtilis*

Aminoacyl tRNA synthetase genes are generally found as the single cistron of an operon. There are exceptions, an example being the class II *lysS* of *B. subtilis*, which is the last gene of the folate operon. While it is unusual to find two synthetases co-regulated by one transcriptional system, one example is the *hisS-aspS* operon in *B. subtilis,* which is controlled by a single putative T-box mechanism located upstream of the *hisS* gene. This particular genetic arrangement is common in the *Firmicutes* where there are 58 instances of both *hisS* and *aspS* coregulated by a single putative T-box mechanism. This arrangement is also found in two *Actinobacteria, Arthrobacter* and *Arthrobacter aurescens*. We note that there are two possible structures for stem I in the *hisS-aspS* T-box regulatory element. The first contains a GAU specifier codon complementary to the anticodon of $tRNA^{ASP}$ (Figure 5.3). The second contains a CAU specifier complementary to the anticodon of $tRNA^{HIS}$ (Figure 5.4). The variable base of the T-box motif in the *hisS-aspS* T-box regulatory element is a C. Thus it may be more likely that the stem I structure containing the GAU specifier codon is correct as the discriminator base of $tRNA^{ASP}$ is a G (The discriminator base of $tRNA^{HIS}$ is a C) which would allow a canonical G-C interaction between the acceptor end of $tRNA^{ASP}$ and the T-box motif. The presence of an aspartate specifier codon implies that the histidyl tRNA synthetase responds to a depletion in the

level of charged aspartyl tRNA. Such an arrangement would be suboptimal as it would mean that the system regulating transcription of *hisS* would be unable to respond to a reduction in the level of charging of its own cognate tRNA.

In order to analyze the expression of the *hisS-aspS* T-box regulatory element of *B. subtilis* we constructed a transcriptional fusion of the promoter and upstream leader region of *hisS-aspS* containing the T-box regulatory element to the *gfp* gene ($P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp*) and integrated this construct by Campbell type mechanism into the chromosome of a number of *B. subtilis* strains as outlined below. The proposed secondary structures of the *hisS-aspS* T-box regulatory elements are shown in figures 5.3 and 5.4.

5.4.1 The *hisS-aspS* T-box regulatory element responds to limitation for histidine

To establish if the *hisS-aspS* T-box regulatory element responds to the level of charged $tRNA^{HIS}$ in the cell we integrated the $P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp* fusion onto the chromosome of the *B. subtilis* histidine auxotroph strain 1A58 to produce the strain NF282.1. Strain NF282.1 was then grown in BLM containing either 50μg/ml histidine (histidine replete) or 1μg/ml histidine (histidine limiting) and the effect on *gfp* expression was assessed. The results are presented in figure 5.18.

The data shows that as histidine becomes limiting in the histidine starved culture of strain NF282.1, there is a ten fold increase in *gfp* expression from approximately 8000 units to 80000 units (Figure 5.18). As the histidine replete (50μg/ml) culture enters stationary phase there is a slow increase in expression from approximately 8000 units to 20000 units. These results show that starvation for histidine induces increased expression of the $P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp* fusion indicating that the *hisS-aspS* T-box regulatory element responds to starvation for histidine. This result is interesting as it suggests that the alternative structure of stem I containing the CAU histidine specifier codon may be correct. This would imply a C-C interaction between the discriminator base of $tRNA^{HIS}$ and the variable base of the *hisS-aspS* T-box mechanism indicating that the interaction between the discriminator base and the variable base of the T-box sequence is not important for antitermination in the *hisS-aspS* T-box regulatory element.

5.4.2 The *hisS-aspS* T-box element does not respond to a reduction in the charging of $tRNA^{ASP}$ in a canonical T-box fashion

The *hisS-aspS* T-box regulatory element controls transcription of the *hisS* and *aspS* genes of *B. subtilis*. We have shown that this regulatory element responds to starvation for

histidine. This result suggests that our proposed secondary structure of the *hisS-aspS* T-box regulatory elemtent containing a CAU specifier codon capable of interacting with the anticodon of $tRNA^{HIS}$ may be correct. This would indicate that expression of the *hisS-aspS* genes responds to starvation for $tRNA^{HIS}$ and not $tRNA^{ASP}$. In order to investigate whether the *hisS-aspS* T-box regulatory element responds to limitation for charged $tRNA^{ASP}$ we produced a strain of *B. subtilis* with an IPTG inducible aspartyl tRNA synthetase as no aspartate auxotroph was available. We then integrated the $P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp* fusion into the chromosome of this IPTG inducible *aspS* strain to produce strain NF310.1. The level of *aspS* expression and hence of AspRS in strain NF310.1 can be controlled by altering the concentration of IPTG in the medium. This should in turn alter the level of charged $tRNA^{ASP}$ in the cell. We could then examine the effect of lowered charged $tRNA^{ASP}$ levels on the expression of the $P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp* fusion. We grew this strain (NF310.1) in LB at a range of concentrations of IPTG from 1mM to 100μM IPTG and examined the effects on expression of the $P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp* transcriptional fusion. The data is presented in figure 5.19.

The data shown represents growth curves and expression data from cultures grown in LB containing 1mM, 600μM, and 100μM IPTG. The growth profiles of cultures containing 1mM and 600μM IPTG are quite similar. Strain NF310.1 grown in LB containing 100μM IPTG has noticeably slower growth. The *gfp* expression for all cultures increases from OD_{600} 0.1 to mid-exponential growth (approximately OD_{600} 0.6). This increase in gfp expression is highest in the NF310.1 culture conraining 100μM IPTG. The *gfp* expression continues to increase after this point in the cultures containing 1mM IPTG and 600μM IPTG (Figure 5.19). However, in the culture containing 100μM IPTG there is a decrease after this point. In the culture containing 600μM IPTG *gfp* expression remains between 50000 and 60000 units throughout the growth curve while in the 1mM culture GFP activity increases further after mid-exponential growth to approximately 70000 units (Figure 5.19). The GFP activity profiles in all cultures are similar in early exponential growth. After this point GFP activity is higher in those cultures with high IPTG levels. The data obtained from this analysis is difficult to interpret. There is an increase in *gfp* expression as a result of lowered IPTG concentrations in the media between 200 and 300 minutes of the timecourse. However this peak in *gfp* expression decreases rapidly and cultures containing higher levels of IPTG are then found to higher levels of *gfp* expression. This would suggest that the *hisS-aspS* T-box mechanism is responsive to a reduction in the charging of aspartyl tRNA but this may not occur by a T-box mechanism. Our analysis may be complicated by some other regulatory feature which governs expression of *aspS*. It is also possible that the

NF310.1 culture containing 100μM IPTG acquired a suppressor mutation allowing expression of *aspS* to become IPTG independent. This would result in constitutive expression of *aspS* and may explain the decrease in GFP activity seen at approximately 300 minutes in the clture containing 100μM IPTG. If this hypothesis is correct it would indicate that expression of *hisS-aspS* is sensitive to the level of *aspS* expression and potentially the level of $tRNA^{ASP}$ charging.

5.4.3 A reduction in glutaminyl amidotransferase activity does not affect expression of the *hisS-aspS* T-box-*gfp* transcriptional fusion

As histidine occupies the same mixed codon box as glutamine (Table 3.2)) we tested if lowered levels of charged glutaminyl tRNA resulted in induction of expression of the the *hisS-aspS* T-box mechanism. As mentioned earlier, in *B. subtilis*, gln-$tRNA^{GLN}$ is formed through misacylation of $tRNA^{GLN}$ with glutamate by glutamyl tRNA synthetase. A transamidation reaction then occurs, changing glutamate to glutamine. This reaction is carried out by the GAT (glutamine amidotransferase) system encoded by the *gatCAB* genes. In order to effect a reduction in the level of $tRNA^{GLN}$ charging we placed the *gatCAB* operon under IPTG inducible control. In theory this should allow us to lower the amount of GatCAB in the cell and thus reduce the level of gln-$tRNA^{GLN}$ in the cell. We then integrated the $P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp* fusion to create the strain NF290.1. Strain NF290.1 was grown in LB containing a series of IPTG concentrations from 1mM to 100μM IPTG and the effect on *gfp* expression of the $P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp* fusion was analysed. Results are presented in figure 5.20.

The data displayed in figure 5.20 shows growth and expression data for strain NF290.1 grown with 1mM, 500μM and 100μM IPTG. Expression of the $P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp* fusion is high during early exponential growth in all three cultures. The expression decreases as the cultures approach mid-exponential growth, and then remains fairly constant throughout the rest of the growth curve. The similarity in expression levels in these three culures indicates that a reduction in the amount of gln-$tRNA^{GLN}$ in the cell does not affect the expression of the $P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp* fusion.

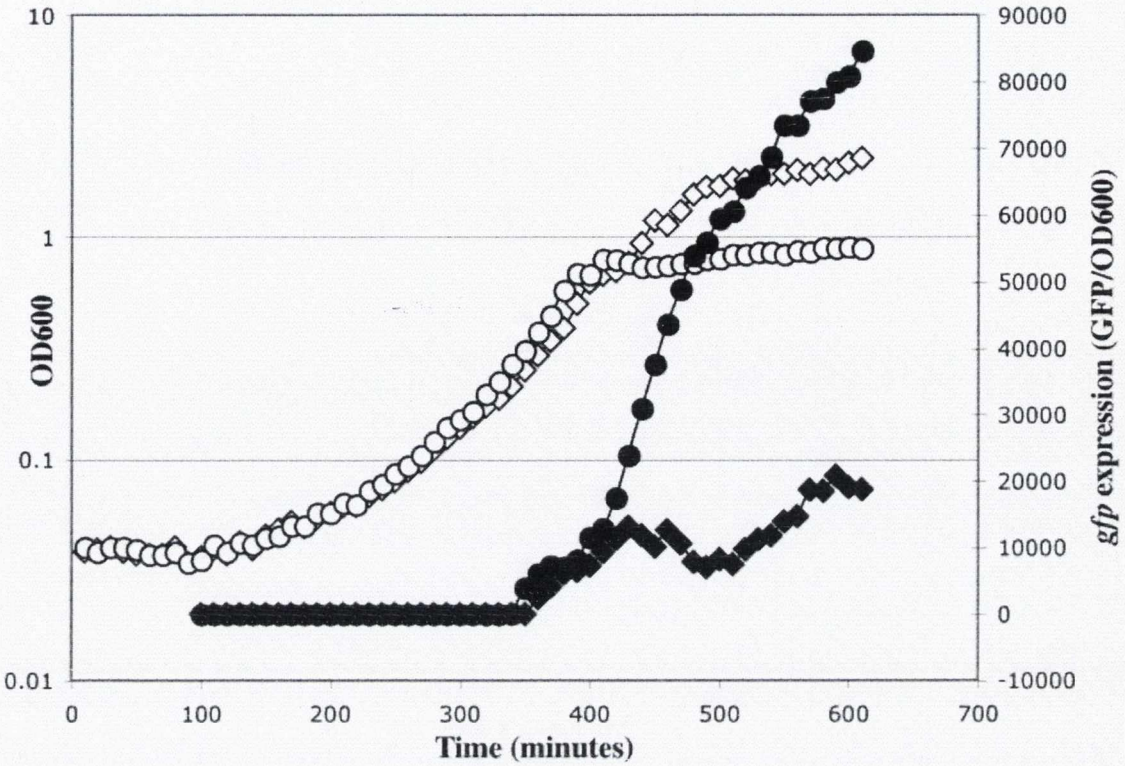


Figure 5.18 Analysis of the response of the *B. subtilis hisS-aspS* T-box regulatory element to limitation for histidine. The graph displays growth and *gfp* expression data for NF282.1 [*hisH2 trpC2*::pNF250.13 ($P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml histidine and 1µg/ml histidine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF282.1 in BLM containing 50µg/ml histidine (◇, ◆). NF282.1 in BLM containing 1µg/ml histidine (○, ●).

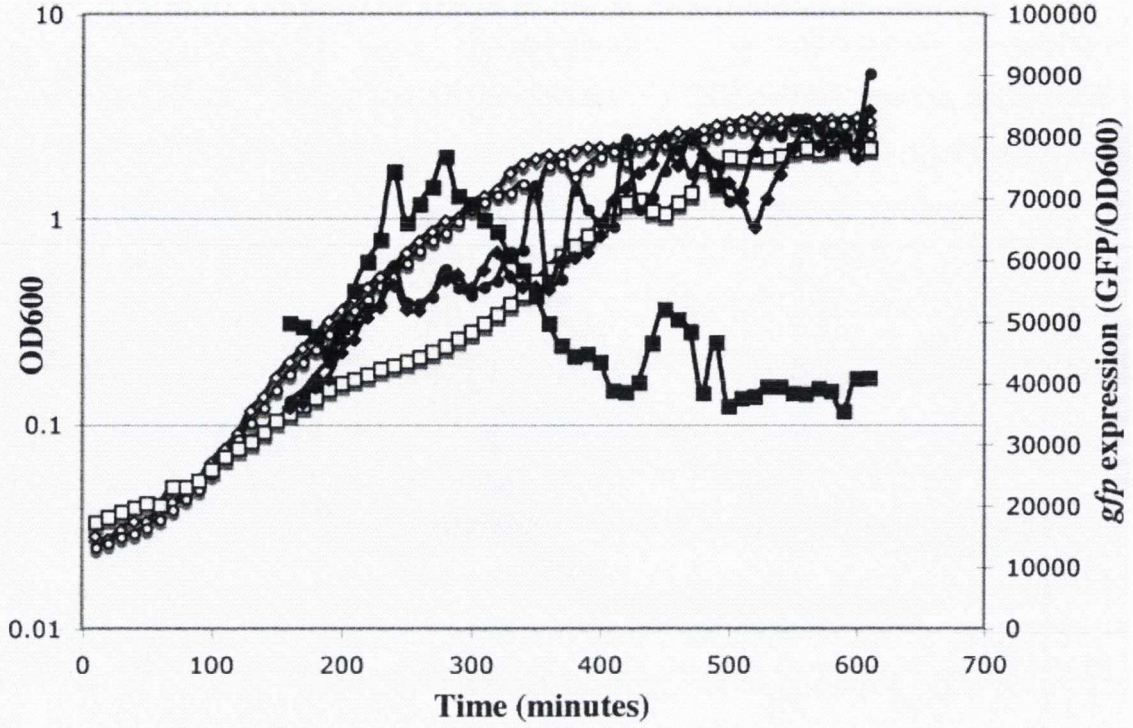


Figure 5.19 Analysis of the response of the *B. subtilis hisS-aspS* T-box regulatory element to a reduction in the level of charged $tRNA^{ASP}$. The graph displays growth and *gfp* expression data for NF310.1 [*trpC2*::pNF250.13 ($P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp*) $Spec^R$ (P_{spac}-*aspS* Em^R) pMAP65 (*penP-lacI* Phl^R)] grown in LB containing 1mM IPTG, 600μM IPTG and 100μM IPTG. Growth is represented by open symbols. Expression values are represented by closed symbols. NF310.1 in LB containing 1mM IPTG (◇, ◆). NF310.1 in LB containing 600μM IPTG (○, ●) and NF310.1 in LB containing 100μM IPTG (□, ■).

5.4.4 Conclusions about the *hisS-aspS* T-box regulatory element

We assessed the expression of a transcriptional fusion containing the promoter and T-box regulatory element located upstream of the *hisS-aspS* operon of *B. subtilis* in order to assess its response to a reduction in charging of $tRNA^{HIS}$ and $tRNA^{ASP}$. The specifier codon of the putative *hisS-aspS* T-box regulatory element may be either a CAU histidine codon or a GAU aspartate codon. Our data shows that the *hisS-aspS* T-box regulatory element responds to a reduction in $tRNA^{HIS}$ charging. This suggests that the specifier codon may be a CAU codon complementary to the anticodon of $tRNA^{HIS}$. However this would also indicate a suboptimal C-C interaction between the discriminator base of $tRNA^{HIS}$ and the variable base of the T–box motif. However, it may be the case that increased expression of the $P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp* fusion, in response to histidine limitation occurs by another mechanism and that a specifier loop containing a GAU aspartate codon is the correct structural arrangement. Further analysis of the *hisS-aspS* T-box regulatory element would be required to investigate this.

Our analysis of the response of the $P_{T\text{-}box\ hisS\text{-}aspS}$-*gfp* fusion to reduced $tRNA^{ASP}$ charging did not show a steady increase in *gfp* expression as would be expected if expression of the *hisS-aspS* genes was induced by a reduction in $tRNA^{ASP}$ charging by a T-box mechanism. However, there was a difference in the profile of *gfp* expression between strain NF310.1 grown at 1mM and 600μM IPTG and strain NF310.1 grown at 100μM IPTG. This may indicate that while the expression of the *hisS-aspS* genes is sensitive to reduced $tRNA^{ASP}$ charging, this may not occur by a T-box mechanism. This data may also show that growth of strain NF310.1 in media containing 100µM IPTG does not allow sufficient *aspS* expression and induces mutations rendering strain NF310.1 IPTG independent.

It has been reported that in *Bacillus lichenoformis* and *Clostridium thermocellum*, the specifier loop of the *hisS-aspS* T-box regulatory element contains the sequence GACAC. This sequence contains a GAC (aspartate) codon and a CAC (histidine) codon that overlap by one nucleotide. Therefore, it may be possible in these examples that the *hisS-aspS* T-box regulatory element is responsive to reduced charging of both $tRNA^{HIS}$ and $tRNA^{ASP}$ (Gutierrez-Preciado, *et al.*, 2009). This has not been shown experimentally and this possibility is not applicable to the *hisS-aspS* T-box regulatory element in B. subtilis.

In conclusion our results show that the *hisS-aspS* T-box regulatory element is responsive to limitation for histidine and that this response may be due to the interaction of uncharged $tRNA^{HIS}$ with the *hisS-aspS* T-box regulatory element by a T-box mechanism. The *hisS-aspS* T-box regulatory element also shows some response to altered levels of $tRNA^{ASP}$

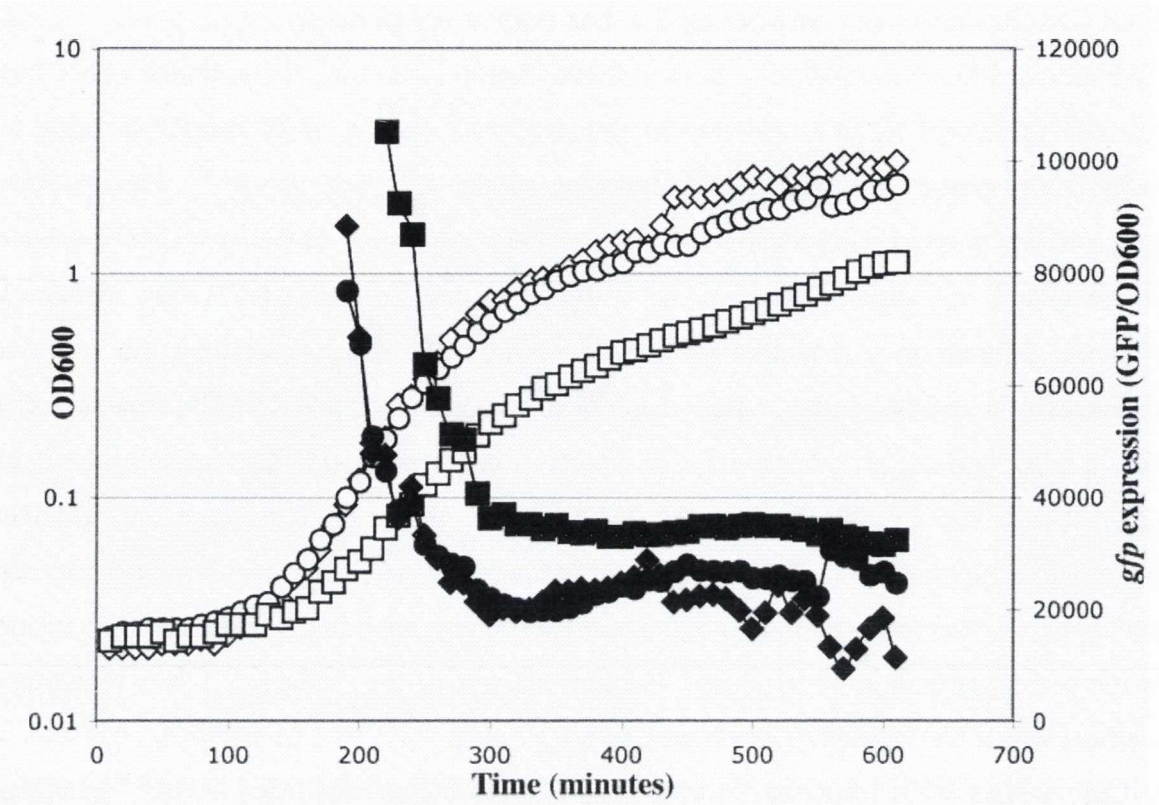


Figure 5.20 Analysis of the response of the *B. subtilis hisS-aspS* T-box regulatory element to a reduction in $tRNA^{GLN}$ charging. The graph displays growth and *gfp* expression data for NF290.1 [*trpC2*::pNF250.13 ($P_{T\text{-}box\ hisS\text{-}aspS}$-GFP) $Spec^R$ (P_{spac}-*gatCAB* Em^R) pMAP65 (*penP-lacI* Phl^R)] grown in LB containing 1mM IPTG, 500μM IPTG and 100μM IPTG. Growth is represented by open symbols. Expression values are represented by closed symbols. NF290.1 in LB containing 1mM IPTG (◇, ◆). NF290.1 in LB containing 500μM IPTG (○, ●) and NF290.1 in LB containing 100μM IPTG (□, ■).

charging. However, this response does not appear to be the result of a canonical T-box mechanism.

5.5 The *ileS* T-box regulatory element of *B. subtilis*

In order to test the expression of the *ileS* T-box regulatory element of *B. subtilis* we constructed a transcriptional fusion of the promoter and upstream leader region of *ileS* containing the T-box regulatory element to the *gfp* gene ($P_{T\text{-}box\ ileS}$-*gfp*) and integrated this construct by Campbell type mechanism into the chromosome of *B. subtilis* strains 1A75 (leucine, isoleucine and methionine auxotroph) and 1A765 (lysine auxotroph). These auxotroph strains were selected to (i) test the response of the *ileS* T-box regulatory element to starvation for its cognate amino acid (ii) test the response of the *ileS* T-box regulatory element to reduced charging of methionyl tRNA (isoleucine and methionine occupy the same mixed codon box, thus $tRNA^{MET}$ is complementary to the *ileS* T-box regulatory element's specifier codon at two out of three positions) and (iii) act as a control for induction. The $tRNA^{LYS}$ anticodon is only complementary to the *ileS* T-box regulatory element's specifier codon at one out of three positions and therefore should not be able to interact with the specifier codon to stabilize antiterminator formation. Experiments to analyze the response of these strains to starvation for these amino acids were carried out as described below.

5.5.1 The T-box regulatory element upstream of the *ileS* gene in *B. subtilis* is functional and responsive to a reduction in $tRNA^{ILE}$ charging

To assess if the *ileS* T-box regulatory element is responsive to limitation for isoleucine and thus charged $tRNA^{ILE}$ we integrated the $P_{T\text{-}box\ ileS}$-*gfp* fusion into the chromosome of the *B. subtilis* strain 1A75 (*ilvA1 leuB8 metB5*) to produce the strain NF267.1. Strain NF267.1 was then grown in BLM containing 50μg/ml isoleucine (isoleucine replete) and 3μg/ml isoleucine (isoleucine depleted) and *gfp* expression was assessed. Results are shown in figure 5.21.

Growth of strain NF267.1 in BLM containing 3μg/ml isoleucine begins to slow and plateau at approximately OD_{600} 0.8 whereas growth in the isoleucine replete media continues to OD_{600} 3. As growth begins to slow and plateau there is a rapid increase in expression of the $P_{T\text{-}box\ ileS}$-*gfp* fusion in the isoleucine limited NF267.1 culture. GFP activity increases from approximately 5000 units to ~ 23000 units over a period of 100 minutes (Figure 5.21). As the isoleucine replete NF267.1 culture enters stationary phase there is a slow increase in *gfp* expression from 5000 units to 10000 units over 100 minutes. These results show that limitation for isoleucine and thus charged $tRNA^{ILE}$ induces increased expression of a $P_{T\text{-}box\ ileS}$-*gfp* fusion. We conclude that the *ileS* T-box regulatory element is functional and responsive to reduced $tRNA^{ILE}$ charging.

5.5.2 Mutation of the specifier codon (AUC) of the isoleucyl tRNA synthetase T-box regulatory element to a UAA Stop codon inhibits induction of $P_{T\text{-}box\ ileS}$-*gfp* fusion expression by isoleucine starvation

In order to confirm that the *ileS* T-box regulatory element is responsive to a reduction in the concentration of uncharged isoleucyl tRNA, the specifier codon (AUC) of the T-box element was changed to a UAA ochre Stop codon. This mutation should prevent the uncharged isoleucyl tRNA interacting with the specifier loop. The altered transcriptional fusion ($P_{T\text{-}box\ ileS\ STOP}$-*gfp*) was transformed into a *B. subtilis* isoleucine auxotroph to produce the strain NF307.1. Strain NF307.1 was grown in BLM in isoleucine replete and isoleucine limiting conditions as previously described and the *gfp* expression was evaluated. Results are presented in figure 5.22.

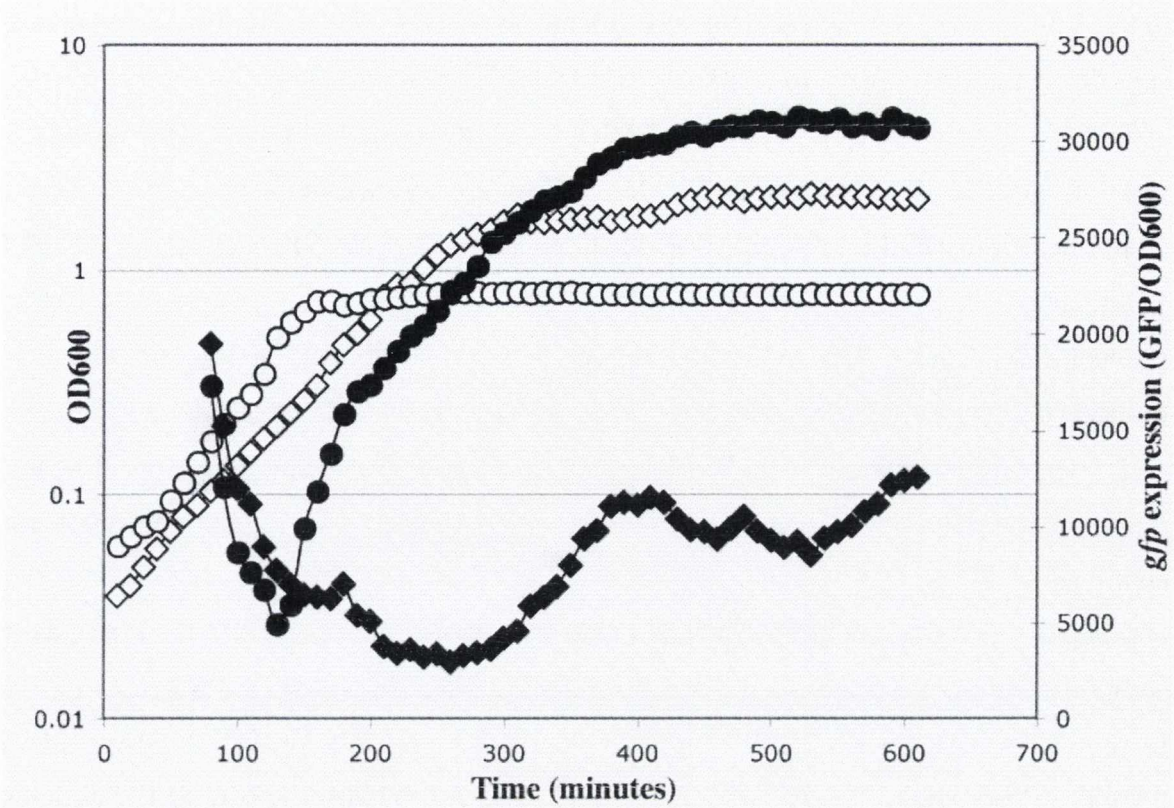


Figure 5.21 Analysis of the response of the *B. subtilis ileS* T-box regulatory element to starvation for isoleucine. The graph displays growth and *gfp* expression data for NF267.1 [*ilvA1 leuB8 metB5*::pNF253.1 ($P_{T\text{-}box\ ileS}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml isoleucine and 3µg/ml isoleucine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF267.1 in BLM containing 50µg/ml isoleucine (◇, ◆). NF267.1 in BLM containing 3µg/ml isoleucine (○, ●).

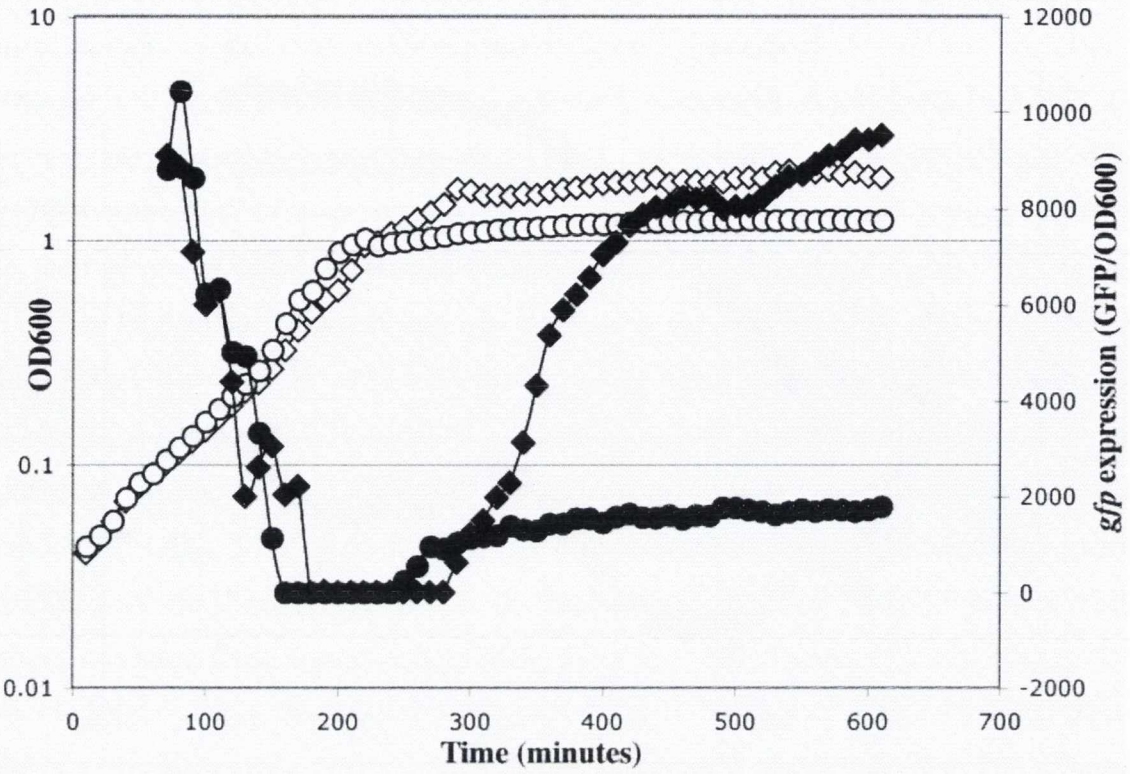


Figure 5.22 Analysis of the response of the *B. subtilis ileS* T-box regulatory element containing a UAA specifier codon to starvation for isoleucine. The graph displays growth and *gfp* expression data for NF307.1 [*ilvA1 leuB8 metB5*::pNF287.1 ($P_{T\text{-}box\ ileS\ STOP}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml isoleucine and 3µg/ml isoleucine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF307.1 in BLM containing 50µg/ml isoleucine (◇, ◆). NF307.1 in BLM containing 3µg/ml isoleucine (○, ●).

Growth of strain NF307.1 in isoleucine limiting conditions begins to slow and plateau at approximately OD_{600} 1. At this point there is no increase in expression of the $P_{T\text{-}box\ ileS\ STOP}$-*gfp* fusion. Expression then increases slightly to ~ 2000 units slightly as the isoleucine limited NF307.1 culture continues into stationary phase. As the isoleucine replete NF307.1 culture enters stationary phase there is an increase in expression from 0 to approximately 6000 units over a period of 100 minutes. These data show that there is no induction of increased *gfp* expression as a result of limitation for isoleucine in a strain containing the $P_{T\text{-}box\ ileS\ STOP}$-*gfp* fusion. Expression of this fusion in strain NF307.1 is higher when the strain is grown in isoleucine replete conditions. We conclude that the $P_{T\text{-}box\ ileS\ STOP}$-*gfp* fusion does not respond to reduced charging of $tRNA^{ILE}$ and that this is due to mutation of the specifier codon from an isoleucine codon to a stop codon that cannot interact with $tRNA^{ILE}$. This confirms that the *ileS* T-box regulatory element responds to the level of charged $tRNA^{ILE}$.

5.5.3 The *ileS* T-box regulatory element does not respond to limitation for methionine

Isoleucine and methionine are partners in the same mixed codon box. Therefore we sought to establish whether a reduction in the level of charged methionyl tRNA resulted in induction of expression of the $P_{T\text{-}box\ ileS}$-*gfp* fusion. Strain NF267.1 is a methionine auxotroph and could be used to test this.

Strain NF267.1 was grown in BLM containing 50μg/ml methionine (methionine replete) and 3μg/ml methionine (methionine limited) and its *gfp* expression was analyzed (Figure 5.23).

Growth of strain NF267.1 in methionine limiting conditions begins to plateau at ~ OD_{600} 0.9. At this point there is a slow increase in *gfp* expression from approximately 6000 units to ~ 12000 units within a period of 100 minutes. A similar increase in *gfp* expression is observed in the methionine replete NF267.1 culture at the end of exponential growth and into stationary phase. The similarities in the profiles of *gfp* expression, seen in both methionine replete and methionine limited cultures indicate that limitation for methionine has no effect on expression of the $P_{T\text{-}box\ ileS}$-*gfp* fusion. Therefore, uncharged $tRNA^{MET}$ cannot induce antitermination in the *ileS* T-box regulatory element.

5.5.4 Expression of the $P_{T\text{-}box\ ileS}$-*gfp* fusion is induced by lysine limitation

To test if induction of expression of the *ileS* T-box regulatory element is specific to isoleucine starvation and not a result of general amino acid starvation we assayed the

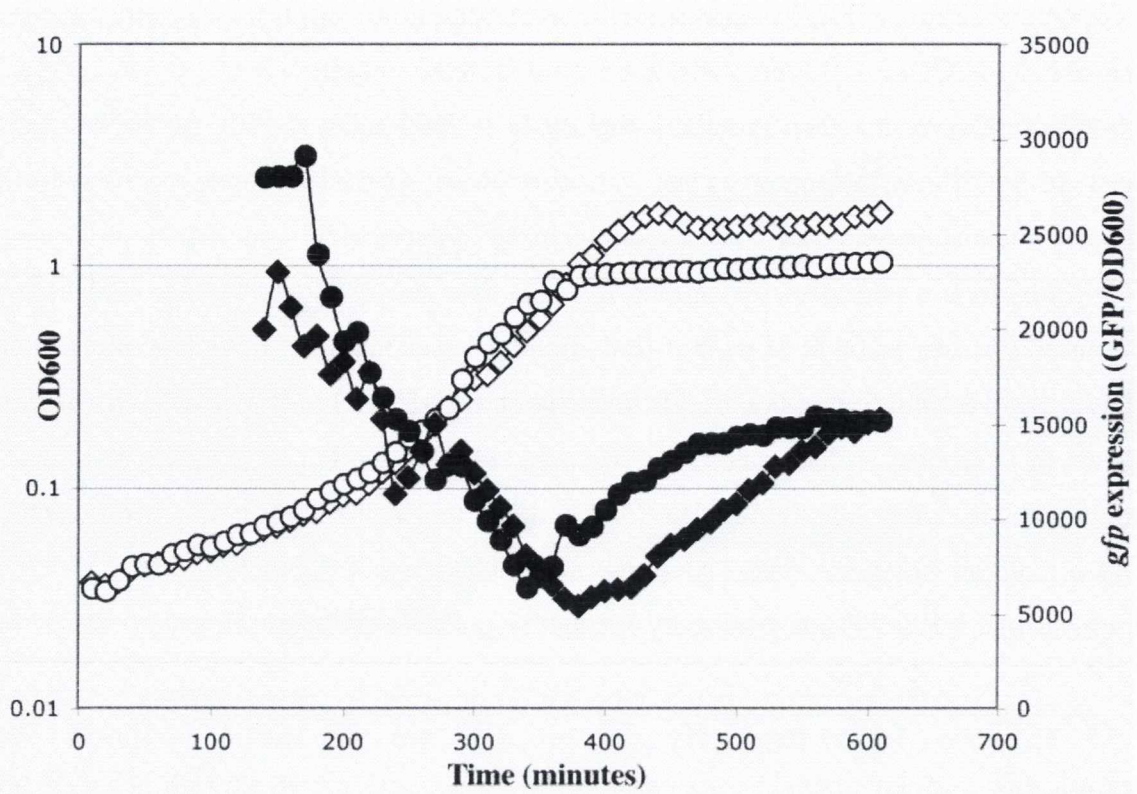


Figure 5.23 Analysis of the response of the *B. subtilis ileS* T-box regulatory element to starvation for methionine. The graph displays growth and *gfp* expression data for NF267.1 [*ilvA1 leuB8 metB5*::pNF253.1 ($P_{T\text{-}box\ ileS}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml methionine and 3µg/ml methionine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF267.1 in BLM containing 50µg/ml methionine (◇, ◆). NF267.1 in BLM containing 3µg/ml methionine (○, ●).

response of the $P_{T\text{-}box\ ileS}$-*gfp* fusion to lysine limitation. To do this we transformed the $P_{T\text{-}box\ ileS}$-*gfp* fusion into the *B. subtilis* lysine auxotroph strain 1A765 to create strain NF293.1. We then grew strain NF293.1 in BLM at 50μg/ml lysine (lysine replete) and 10μg/ml lysine (lysine limiting) and assessed the effects on *gfp* expression. The results are presented in figure 5.24.

Growth of strain NF293.1 in lysine limiting conditions begins to slow and plateau at ~ OD_{600} 0.8. At this point there is a concomitant and rapid increase in *gfp* expression from approximately 7000 units to ~ 25000 units within 100 minutes (Figure 5.24). Growth of strain NF293.1 in lysine replete conditions begins to slow and plateau and enters stationary phase at approximately 300 minutes after inoculation at OD_{600} 2. As the lysine replete culture enters stationary phase there is a slow increase in *gfp* expression from ~ 7000 units to ~ 10000 units within 100 minutes. This data shows that limitation for lysine induces increased expression of the $P_{T\text{-}box\ ileS}$-*gfp* fusion.

5.5.5 Expression of the $P_{T\text{-}box\ ileS}$-*gfp* fusion is induced by lysine limitation following mutation of the AUC (isoleucine) specifier codon to a UAA (stop) codon

We observed an increase in expression of the $P_{T\text{-}box\ ileS}$-*gfp* fusion in response to limitation for lysine. We therefore sought to test if this was due to the ability of $tRNA^{LYS}$ to interact with the *ileS* T-box mechanism. If the result in the previous experiment was due to the interaction of uncharged $tRNA^{LYS}$ with the *ileS* T-box systems specifier codon, it implied an interaction of the AUC specifier with the UUU anticodon of $tRNA^{LYS}$. As there is only one point of interaction between the specifier codon and the anticodon of lysyl tRNA this is unlikely. Mutation of the specifier codon to a UAA Stop codon would in theory increase the ability of uncharged $tRNA^{LYS}$ to interact with the specifier loop by increasing the number of nucleotides in the specifier codon that are complementary to the $tRNA^{LYS}$ anticodon. We would therefore expect an increase in the level of *gfp* expression as a result of lysine limitation using this transcriptional fusion. In order to examine this possibility we transformed the construct containing the mutated specifier codon ($P_{T\text{-}box\ ileS\ STOP}$-*gfp*) into the *B. subtilis* lysine auxotroph 1A765 and assessed the expression in lysine rich and lysine limiting conditions as described previously. The results are presented in figure 5.25.

Growth of strain NF321.1 in lysine limiting conditions begins to slow and plateau at ~ OD_{600} 0.8. At this point there is a coincident increase in *gfp* expression from approximately 200 units to ~ 15000 units over a period of 100 minutes. As growth of strain NF321.1 in lysine replete conditions begins to slow and plateau there is a small and

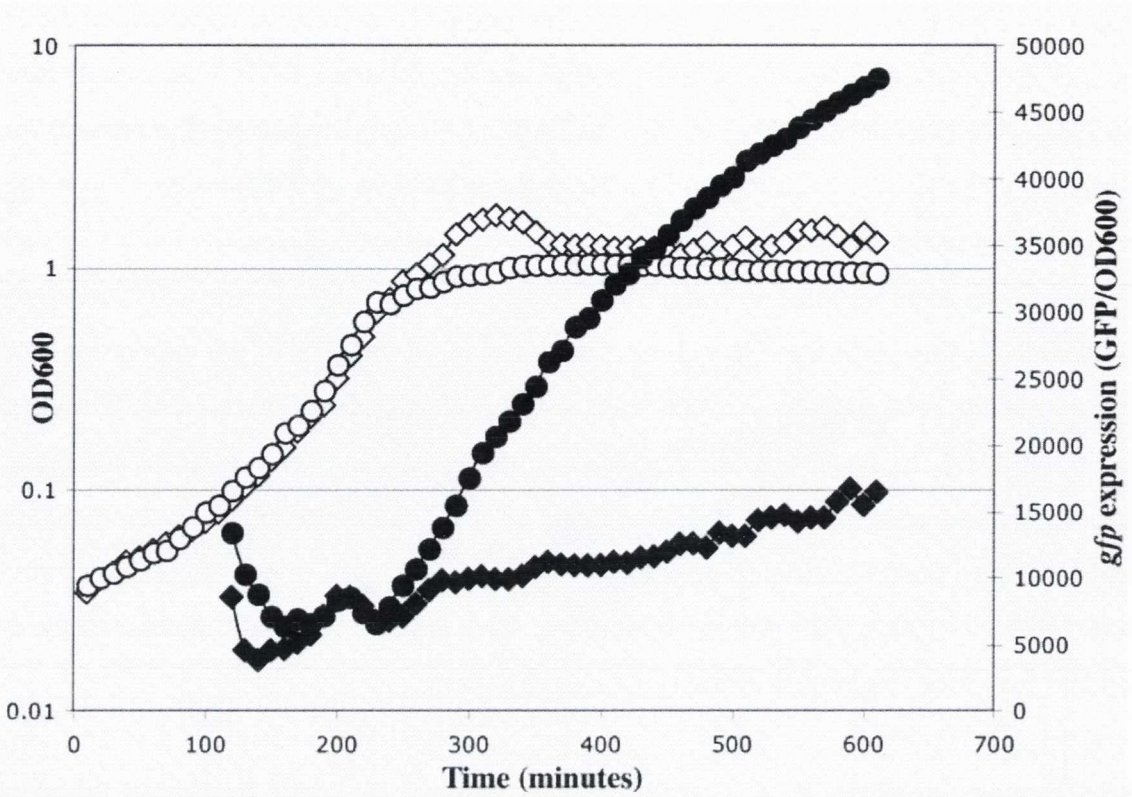


Figure 5.24 Analysis of the response of the *B. subtilis ileS* T-box regulatory element to starvation for lysine. The graph displays growth and *gfp* expression data for NF293.1 [*trpC2 lys*::pNF253.1 ($P_{T\text{-}box\ ileS}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml lysine and 10µg/ml lysine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF293.1 in BLM containing 50µg/ml lysine (◇, ◆). NF293.1 in BLM containing 10µg/ml lysine (○, ●).

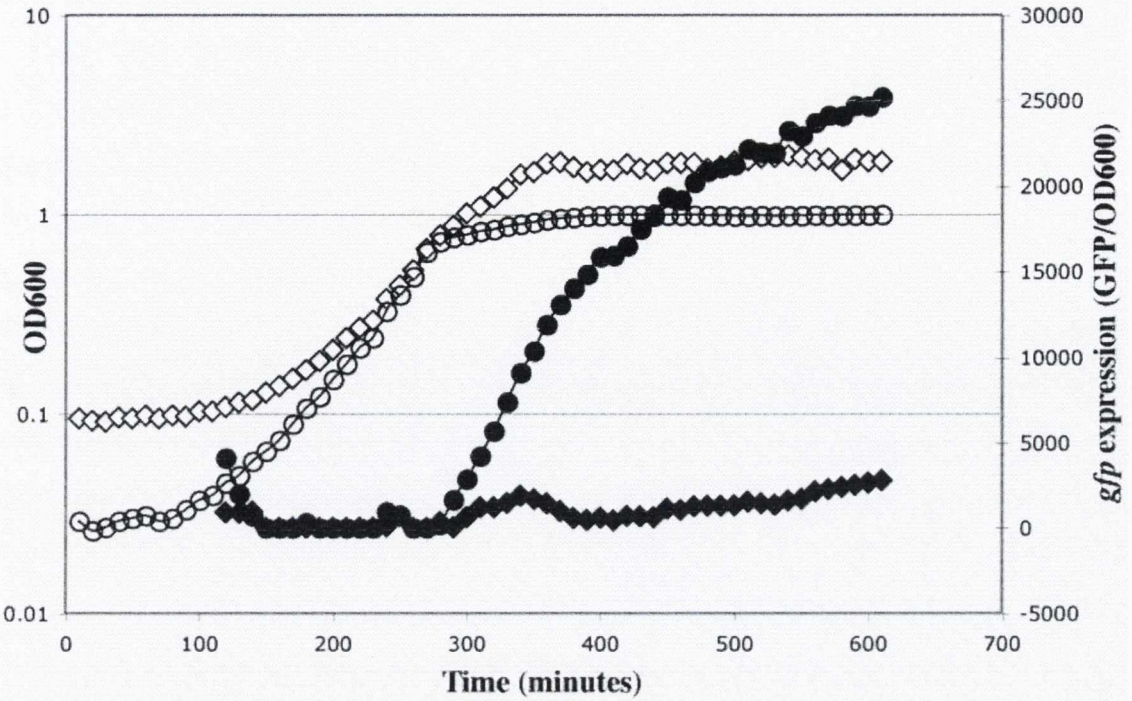


Figure 5.25 Analysis of the response of the *B. subtilis ileS* T-box regulatory element containing a UAA specifier codon to limitation for lysine. The graph displays growth and *gfp* expression data for NF321.1 [*trpC2 lys*::pNF287.1 ($P_{T\text{-}box\ ileS\ STOP}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml lysine and 10µg/ml lysine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF321.1 in BLM containing 50µg/ml lysine (◇, ◆). NF321.1 in BLM containing 10µg/ml lysine (○, ●).

gradual increase in *gfp* expression. The results of this experiment show there is an increase in *gfp* expression in the $P_{T\text{-box } ileS \text{ STOP}}$-*gfp* fusion due to limitation for lysine. A comparison of the difference in *gfp* expression between lysine replete and lysine limiting cultures of strains NF293.1 ($P_{T\text{-box } ileS}$-*gfp*) (Figure 5.24) and NF321.1 ($P_{T\text{-box } ileS \text{ STOP}}$-*gfp*) (Figure 5.25), calculated by subtraction of the *gfp* expression values of the lysine replete culture from the lysine limiting culture after the point of induction of increased *gfp* expression in the lysine limited culture, shows that the increase in expression due to lysine limitation is the same in strains NF293.1 and NF321.1 (data not shown). This shows that mutation of the specifier codon from AUC to UAA has no effect on induction of the $P_{T\text{-box } ileS}$-*gfp* fusion by limitation for lysine. This suggests that the increase in *gfp* expression seen as a result of lysine starvation is not due to interaction of $tRNA^{LYS}$ with the *ileS* T-box regulatory element and is due to some indirect effect.

5.5.6 Conclusions about the *ileS* T-box regulatory element

We investigated the expression of a transcriptional fusion containing the promoter and T-box regulatory element of the *ileS* gene of *B. subtilis* in order to assess its response to a reduction in charging of its cognate tRNA and also to test if there was a response to limitation for charged $tRNA^{MET}$ which shares 2 out of 3 nucleotides with $tRNA^{ILE}$ at the anticodon. Our results showed that the *ileS* T-box regulatory element is functional and responds to a reduction in the level of charging of its own cognate tRNA. Our analysis also showed that limitation for methionine had no effect on the $P_{T\text{-box } ileS}$-*gfp* fusion indicating that $tRNA^{MET}$ cannot induce antitermination in the *ileS* T-box regulatory element. Additionally we tested the response of the $P_{T\text{-box } ileS}$-*gfp* fusion to starvation for lysine. The results of these experiments showed that there is an increase in expression of a $P_{T\text{-box } ileS}$-*gfp* fusion in response to limitation for lysine. Following mutation of the specifier codon. a response to lysine limitation was still observed though the basal level of expression was reduced. An analysis of the increase in *gfp* expression due to lysine limitation showed no difference between strain NF293.1 and NF321.1 showing that mutation of the specifier codon has no effect on induction of expression due to lysine starvation. Thus, antitermination of the *ileS* T-box regulatory element is induced by lysine limitation but does not occur by interaction of $tRNA^{LYS}$ with the T-box regulatory element. The increase in expression of the $P_{T\text{-box } ileS}$-*gfp* fusion due to lysine limitation may occur by a number of indirect effects, for example, an indirect reduction in $tRNA^{ILE}$ charging brought about by limitation for lysine.

5.6 The *trpS* T-box regulatory element of *B. subtilis*

It has previously been shown by Steinberg that growth of a temperature sensitive *trpS* mutant at non-permissive temperatures causes increased *trpS* expression (Steinberg, 1974). This observation coupled with the discovery of a T-box antitermination mechanism upstream of the *trpS* gene lead to the hypothesis that *trpS* in *B. subtilis* is regulated by a T-box mechanism (Grundy and Henkin, 1993). However, this has not been confirmed by a direct analysis of the *trpS* T-box regulatory element.

In order to characterize the expression of the *trpS* T-box regulatory element of *B. subtilis* we constructed a transcriptional fusion of the promoter and upstream leader region of *trpS* containing the T-box regulatory element to the *gfp* gene ($P_{T\text{-}box\ trpS}$-*gfp*). This construct was integrated at the *trpS* locus by Campbell type mechanism into the chromosome of *B. subtilis* strains 168 (tryptophan auxotroph), 1A79 (tryptophan and cysteine auxotroph) and 1A75 (methionine auxotroph). These auxotroph strains were selected to (i) test the response of the *trpS* T-box regulatory element to starvation for its cognate amino acid (ii) test the response of the *trpS* T-box regulatory element to limitation for charged $tRNA^{CYS}$ ($tRNA^{CYS}$ is complementary to the specifier codon of the *trpS* T-box regulatory element at two out of three positions. Additionally there is significant sequence homology between $tRNA^{TRP}$ and $tRNA^{CYS}$ though they differ at the discriminator base) (Figure 5.26) and (iii) act as a control for induction of *gfp* expression. The $tRNA^{MET}$ anticodon is only complementary to the *trpS* T-box regulatory element's specifier codon at one out of three positions and its discriminator base is not complementary to the variable nucleotide of the *trp*S T-box sequence. Therefore uncharged $tRNA^{MET}$ should not be able to induce transcriptional readthrough of the *trpS* T-box regulatory element. Experiments to analyze the response of these strains to starvation for these amino acids were carried out as described below. The proposed secondary structure of the *trpS* T-box regulatory element is illustrated in figure 5.6.

5.6.1 The T-box regulatory element of the *trpS* gene in *B. subtilis* is functional and sensitive to tryptophan starvation

To assess if the *trpS* T-box regulatory element is functional and responds to tryptophan starvation we integrated the $P_{T\text{-}box\ trpS}$-*gfp* fusion into the chromosome of a *B. subtilis* tryptophan auxotroph to produce the strain NF277.1. We grew strain NF277.1 in BLM containing 50 µg/ml tryptophan (tryptophan replete) and 1µg/ml

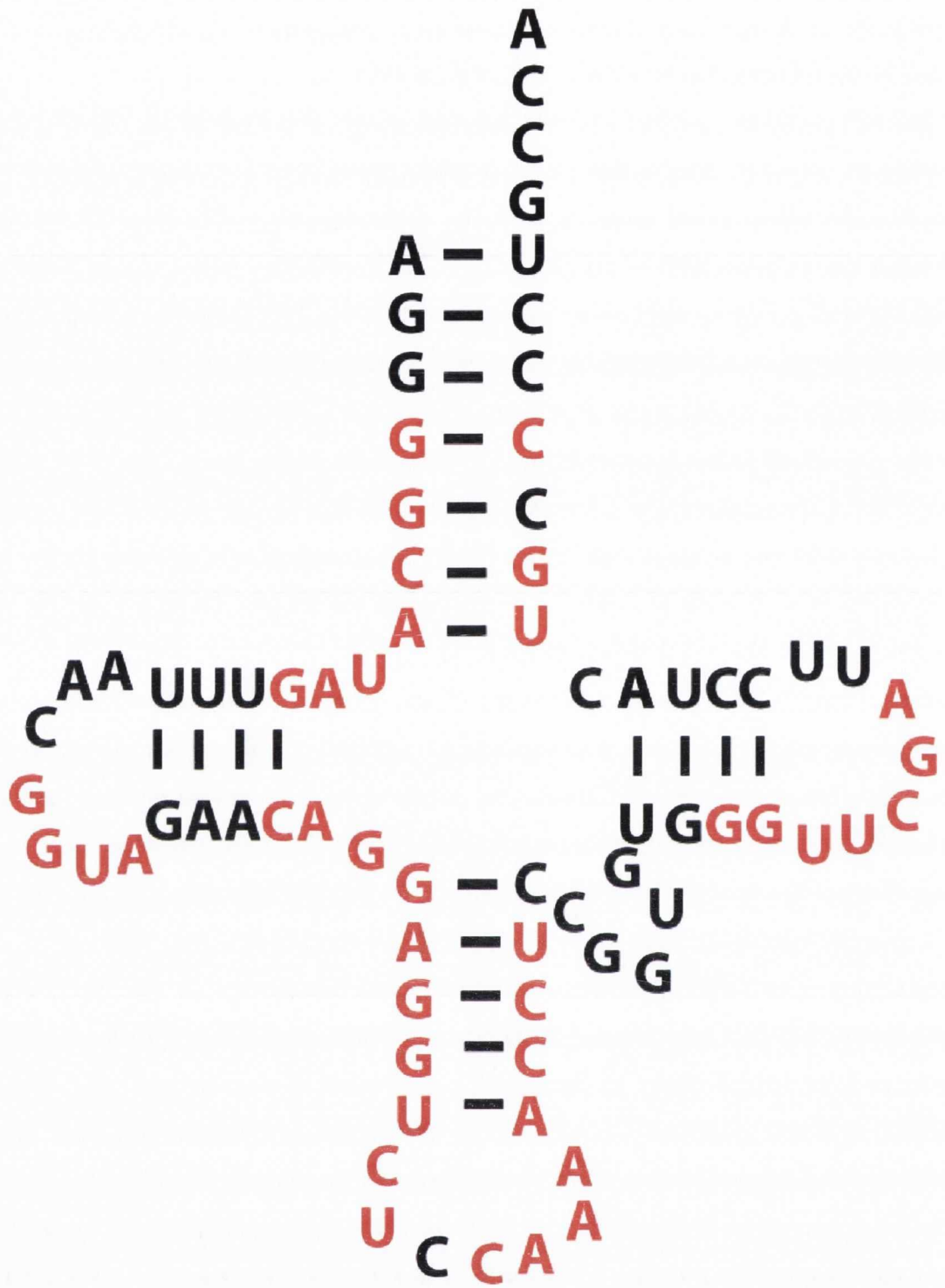


Figure 5.26 Secondary structure of *B. subtilis* $tRNA^{TRP}$. Regions of homology with $tRNA^{CYS}$ are highlighted in red.

tryptophan (tryptophan limiting) and the effect on *gfp* expression was analyzed. Results from this experiment are presented in figure 5.27.

Growth of strain NF277.1 in tryptophan limiting conditions begins to slow and plateau at ~ OD_{600} 1. At this point there is a concomitant and rapid increase in *gfp* expression with values increasing from ~ 5000 units to 50000 units within 100 minutes. The *gfp* expression profile of strain NF277.1 grown in tryptophan replete conditions does not increase in response to a slow down in growth as the culture enters stationary phase. This data shows that expression of the $P_{T\text{-}box\ trpS}$-*gfp* fusion increases in response to tryptophan limitation indicating that the *trpS* T-box regulatory element is functional and responsive to reduced $tRNA^{TRP}$ charging.

5.6.2 The T-box regulatory element upstream of the *trpS* gene in *B. subtilis* does not respond to depletion for cysteine

As cysteine and tryptophan are partners in a mixed codon box we wanted to establish whether the *trpS* T-box regulatory element could respond to starvation for cysteine. To examine this we transformed the $P_{T\text{-}box\ trpS}$-*gfp* fusion into the *B. subtilis* cysteine auxotroph 1A79 to create the strain NF257.1. Strain NF257.1 was grown in BLM containing 50μg/ml cysteine (cysteine rich) and 4μg/ml (cysteine limiting) and its *gfp* expression assessed. The data is presented in figure 5.28.

Growth of strain NF257.1 in cysteine limiting conditions begins to slow and plateau at ~ OD_{600} 0.8. At this point there is a decrease in *gfp* expression from approximately 6000 units to 3000 units. In cysteine replete conditions, the *gfp* expression profile of strain NF257.1 increases during exponential growth from 0 to ~ 10000 units before decreasing to ~ 3000 units as the culture enters stationary phase. The data shows there is no increase in gfp expression in response to cysteine limitation. We conclude therefore that the *trpS* T-box regulatory element does not respond to uncharged $tRNA^{CYS}$.

5.6.3 The *trpS* T-box regulatory element in *B. subtilis* does not respond to depletion for methionine

As a control experiment to show that the *trpS* T-box regulatory element does not respond to general amino acid starvation we integrated the $P_{T\text{-}box\ trpS}$-*gfp* fusion construct into the chromosome of the *B. subtilis* methionine auxotroph strain 1A75 to produce the strain NF302.1. Strain NF302.1 was grown in BLM containing 50μg/ml

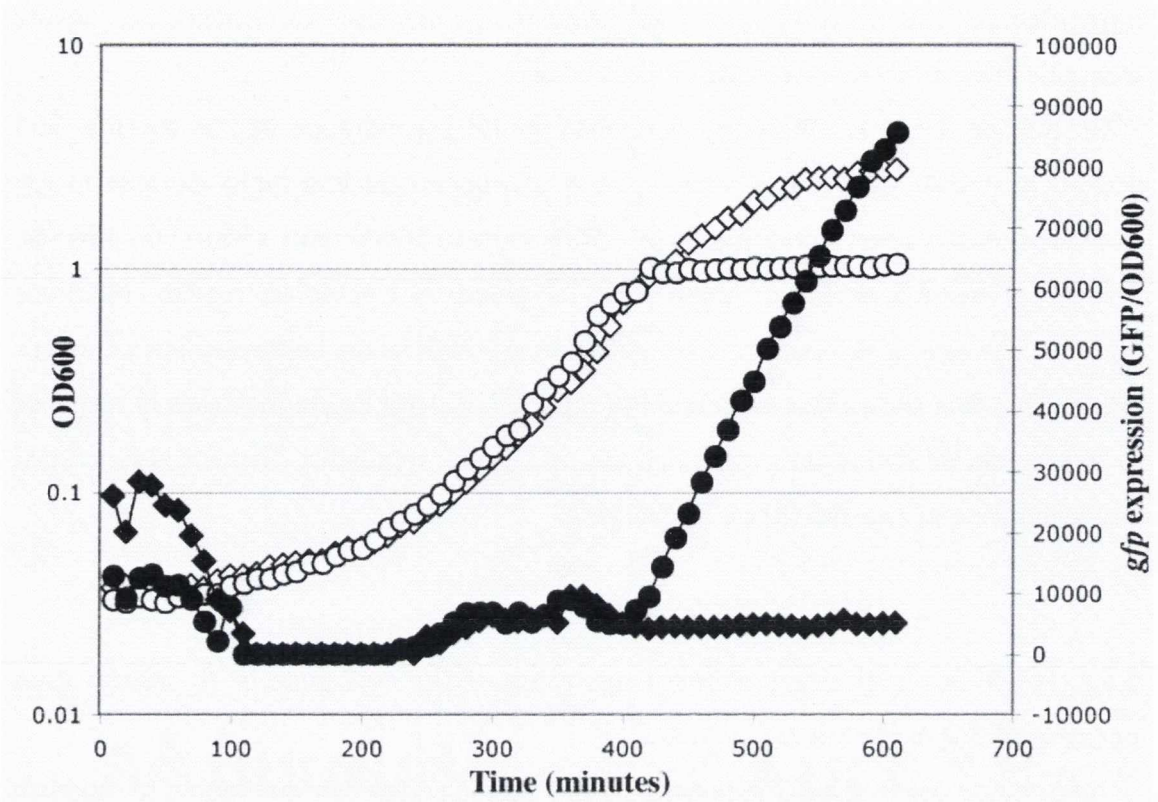


Figure 5.27 Analysis of the response of the *B. subtilis trpS* T-box regulatory element to starvation for tryptophan. The graph displays growth and *gfp* expression data for NF277.1 [*trpC2*::pNF251.1 ($P_{T\text{-}box\ trpS}$-*gfp*) $Spec^R$] grown in BLM containing 50μg/ml tryptophan and 1μg/ml tryptophan. Growth is represented by open symbols. Expression values are represented by closed symbols. NF277.1 in BLM containing 50μg/ml tryptophan (◇, ◆). NF277.1 in BLM containing 1μg/ml tryptophan (○, ●).

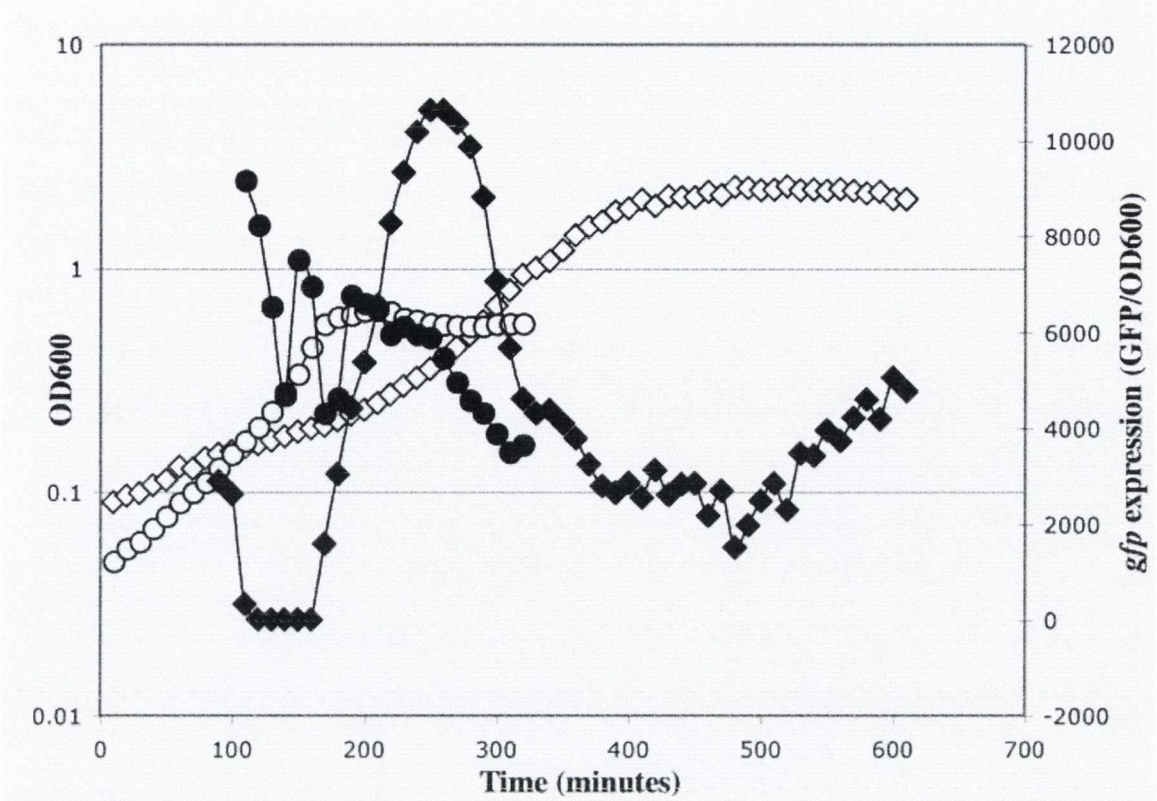


Figure 5.28 Analysis of the response of the *B. subtilis trpS* T-box regulatory element to starvation for cysteine. The graph displays growth and *gfp* expression data for NF257.1 [*cysC1 trpC2*::pNF251.1 ($P_{T\text{-}box\ trpS}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml cysteine and 4µg/ml cysteine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF257.1 in BLM containing 50µg/ml cysteine (◇, ◆). NF257.1 in BLM containing 4µg/ml cysteine (○, ●).

methionine (methionine replete) and 3µg/ml (methionine limiting) and its *gfp* expression profile was determined (Figure 5.29).

Growth of strain NF302.1 in methionine limiting conditions begins to slow and plateau at ~~ OD_{600} 0.9. At this point there is a gradual increase in *gfp* expression from ~ 200 units to 7000 units over a period of 100 minutes. A similar increase in *gfp* expression is observed in strain NF302.1 grown in methionine replete conditions as the culture enters stationary phase. This indicates that the increase in *gfp* expression is linked to a slow down in growth and is not due to interaction of uncharged $tRNA^{MET}$ with the *trpS* T-box regulatory element.

5.6.4 Conclusions about the *trpS* T-box regulatory element

Our analysis of the *trpS* T-box element of *B. subtilis* supports the prediction that the *trpS* gene is transcriptionally regulated by a T-box mechanism. Additionally we have shown that $tRNA^{CYS}$ whose anticodon is complementary to the *trpS* T-box elements specifier codon at two out of three positions and shares significant sequence homology with $tRNA^{TRP}$, cannot interact with the *trpS* T-box regulatory element to induce antitermination. Limitation for methionine also had no effect on expression of the $P_{T\text{-}box\ trpS}$-*gfp* fusion. The *trpS* T-box regulatory element appears to be specifically induced by limitation for tryptophan.

5.7 The *cysES* T-box regulatory element of *B. subtilis*

The *cysS* gene of *B. subtilis* is part of an operon containing the *gltX* gene encoding glutamyl tRNA synthetase and the *cysE* gene encoding serine O-acetyltransferase which is involved in cysteine biosynthesis (Figure 5.30). Transcription in this operon is initiated from a promoter located 43 nucleotides upstream of the *gltX* gene (Gagnon, *et al.*, 1994). There is an intergenic region between *gltX* and *cysE* that contains a T-box regulatory element that is responsive to changes in the level of charging of $tRNA^{CYS}$ (Gagnon, *et al.*, 1994). We constructed a transcriptional fusion of the *gltX* promoter and *cysES* T-box regulatory element to the *gfp* gene ($P_{gltX\text{-}T\text{-}box\text{-}cysES}$-*gfp*) and integrated this construct into the amylase locus of the *B. subtilis* auxotroph strains 168 (*trpC2*) and 1A79 (*trpC2 cysC1*). These auxotroph strains were selected to (i) confirm that the *cysES* T-box responds to a reduction in $tRNA^{CYS}$ charging (ii) test the response of the *cysES* T-box regulatory element to a reduction in the level of charged $tRNA^{TRP}$ (tryptophan and cysteine share the same mixed codon box and therefore $tRNA^{TRP}$ is complementary to the specifier codon of the *cysES* T-box regulatory element at two out

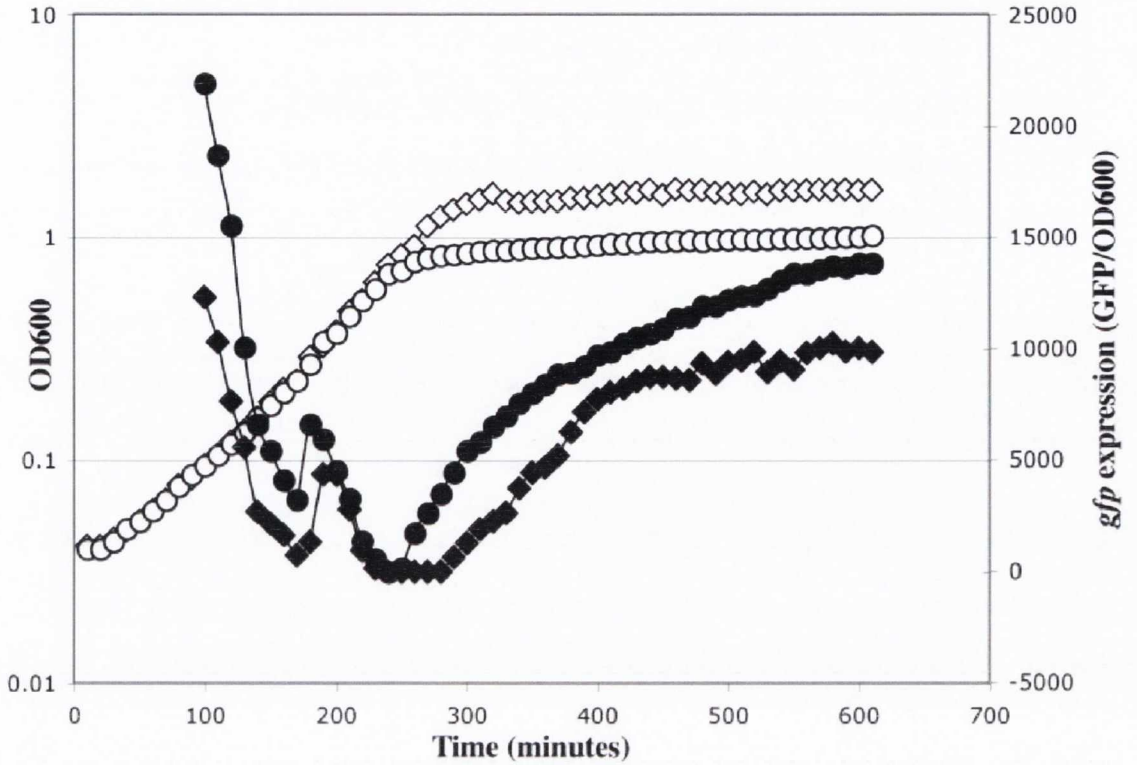


Figure 5.29 Analysis of the response of the *B. subtilis trpS* T-box regulatory element to starvation for methionine. The graph displays growth and gfp expression data for NF302.1 [*ilvA1 leuB8 metB5*::pNF251.1 ($P_{T\text{-}box\ trpS}$-*gfp*) $Spec^R$] grown in BLM containing 50μg/ml methionine and 3μg/ml methionine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF302.1 in BLM containing 50μg/ml methionine (◇, ◆). NF302.1 in BLM containing 3μg/ml methionine (○, ●).

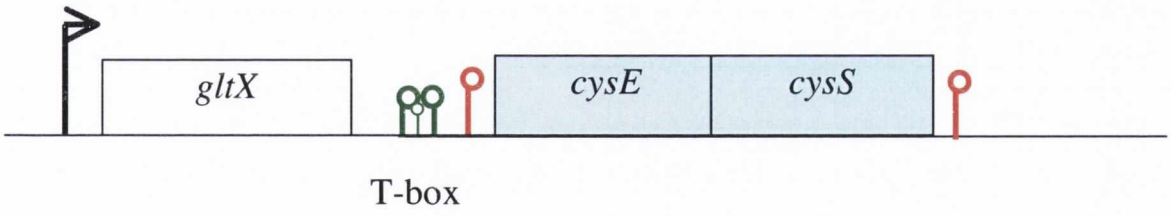


Figure 5.30 The *gltX-cysES* operon of *B. subtilis*. The promoter is indicated by the black arrow. Terminator structures are indicated by red lollipops. Stem loop structures that are part of the *cysES* T-box regulatory element are indicated in green.

of three positions). Experiments to analyze the response of these strains to cysteine and tryprophan limitation were carried out as described below. The putative secondary structure of the *cysES* T-box regulatory element is shown in figure 5.7.

5.7.1 The *cysES* T-box regulatory element responds to depletion for cysteine

To assess whether the *cysES* T-box regulatory element responds to starvation for cysteine the plasmid pNF329.1 ($P_{gltX\text{-T-box-}cysES}$-*gfp*) was integrated onto the chromosome of the *B. subtilis* cysteine auxotroph strain 1A79 (*cysC1 trpC2*) to create the strain NF337.1. This strain was grown in BLM containing a range of cysteine concentrations and the *gfp* expression was determined.

Our attempts to work with this strain met with limited success. It was difficult to grow overnight cultures to a high enough optical density to allow inoculation for the experiment. In addition, when we were able to inoculate cultures we found that those cultures containing higher concentrations of cysteine experienced a much greater lag phase than those cultures with lower cysteine levels, often lasting upto 5 hours. In experiments where we successfully achieved cysteine limitation and cessation of growth we found a very slight induction reponse of the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion (Figure 5.31). In this experiment we saw an induction of the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion from zero to 4500 units of activity at the point of growth cessation in response to cysteine starvation before rapidly falling back to zero. The cysteine replete culture grows more slowly initially but reaches a higher overall OD_{600}. The expression activity in this strain stays at zero throughout the growth curve and shows a slight increase in stationary phase. These results indicate that the *cysES* T-box regulatory element shows some response to cysteine starvation.

5.7.2 Supplementation of the media with methionine alters the response of the *cysES* T-box regulatory element to cysteine starvation

Following a study of the literature relating to cysteine metabolism we postulated that addition of methionine may help the cysteine auxotroph cultures to grow, as methionine can be converted back into cysteine. We found that addition of 5µg/ml methionine allowed us to consistently grow the cysteine auxotroph strains while also allowing us to deplete for cysteine. With this approach we carried out experiments on NF337.1 [*cysC1 trpC2 amyE*::pNF329.1 ($P_{gltX\text{-T-box }cysES}$ -*gfp*) $Spec^R$] to assess the effect of cysteine starvation on the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion with a constant level of methionine of

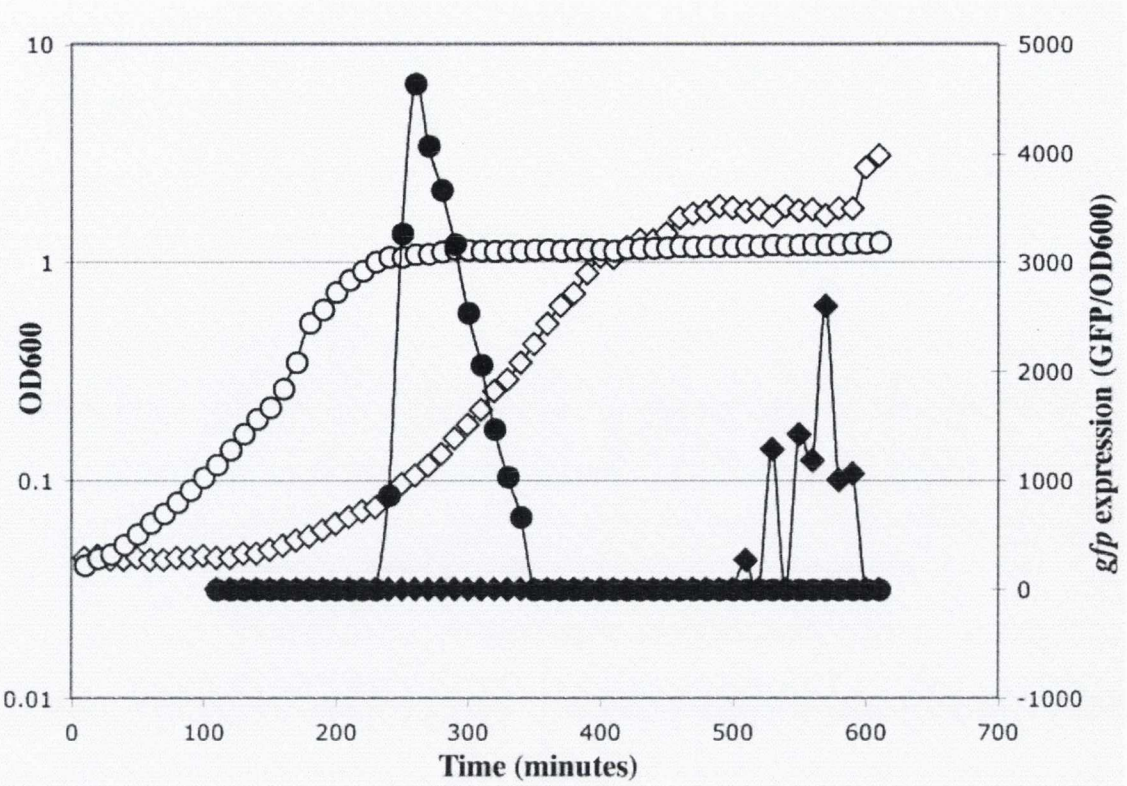


Figure 5.31 Analysis of the response of the *B. subtilis gltX* promoter and *cysES* T-box regulatory element to starvation for cysteine in BLM. The graph displays growth and *gfp* expression data for NF337.1 [*cysC1 trpC2*::pNF329.1 ($P_{gltX\text{-}T\text{-}box\text{-}cysES}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml cysteine and 1µg/ml cysteine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF337.1 in BLM containing 20µg/ml cysteine (◇, ◆). NF337.1 in BLM containing 1µg/ml cysteine (○, ●).

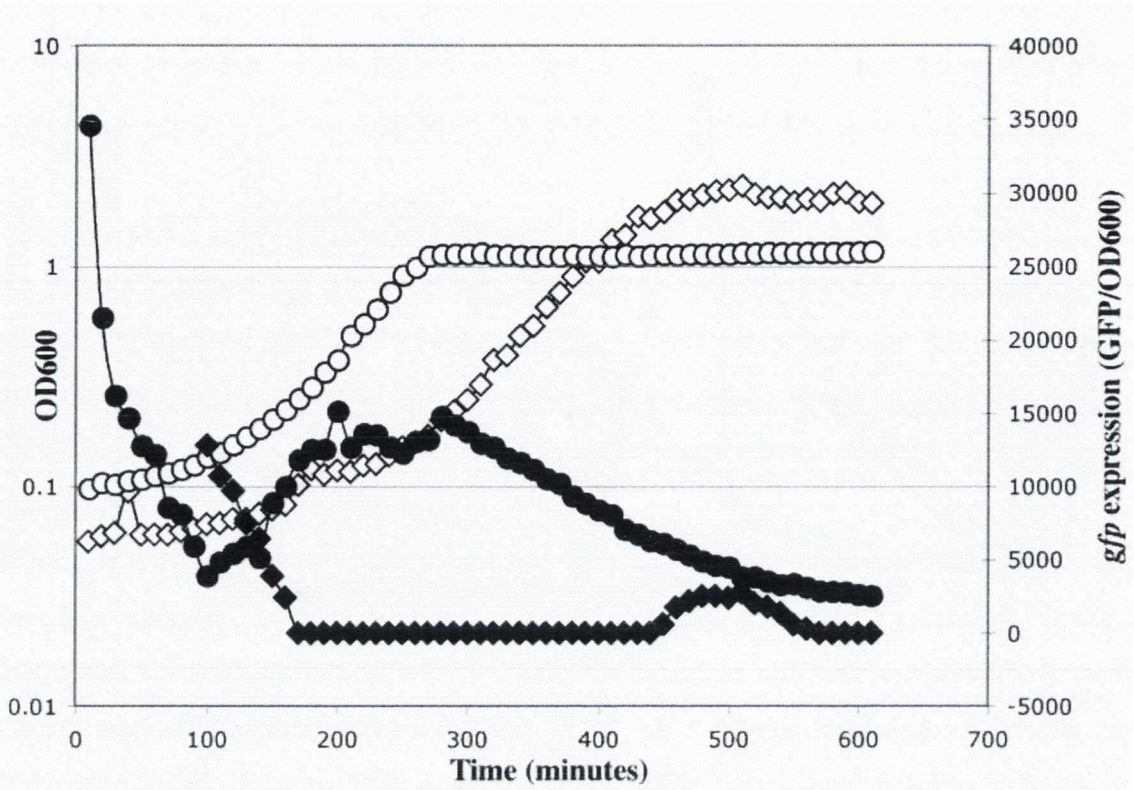


Figure 5.32 Analysis of the response of the *B. subtilis gltX* promoter and *cysES* T-box regulatory element to starvation for cysteine in BLM containing 5µg/ml methionine. The graph displays growth and *gfp* expression data for NF337.1 [*cysC1 trpC2*::pNF329.1 ($P_{gltX\text{-}T\text{-}box\text{-}cysES}$-*gfp*) $Spec^R$] grown in BLM containing 20µg/ml cysteine and 1µg/ml cysteine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF337.1 in BLM containing 20µg/ml cysteine (◇, ◆). NF337.1 in BLM containing 1µg/ml cysteine (○, ●).

5μg/ml. The results are illustrated in figure 5.32 and show data from cultures containing 20μg/ml and 1μg/ml cysteine.

Here we see again that cultures containing higher concentrations of cysteine take longer to achieve exponential growth. The data shows that in the cysteine limited culture there is a steady increase in *gfp* expression during early exponential growth to approximately 15000 units. This level of activity remains at approximately the same value until the culture becomes limited for cysteine at OD_{600} 1.1. At this point there is a steady and concomitant decrease in *gfp* expression that continues into cysteine limitation induced stationary phase. In the cysteine replete culture there is no *gfp* expression during exponential growth. There is an increase in expression from zero to approximately 3000 units as the culture enters stationary phase which then decreases to zero once more as the culture continues into stationary phase. These results present a profile of *gfp* expression in the cysteine limited culture that is contrary to what is normally observed in T-box regulatory elements that are limited for their charged cognate tRNA. Limitation for cysteine should induce increased expression of the $P_{gltX\text{-}T\text{-}box\text{-}cysES}$-*gfp* fusion in a functional *cysES* T-box regulatory element at the point of cysteine limitation. However, in this case limitation for cysteine in the presence of methionine causes a reduction in expression of the $P_{gltX\text{-}T\text{-}box\text{-}cysES}$-*gfp* fusion at the point of cysteine limitation. However, there is increased expression of the $P_{gltX\text{-}T\text{-}box\text{-}cysES}$-*gfp* fusion in the NF337.1 culture limited for cysteine during exponential growth.

Methionine can be converted to cysteine via the AdoMet recycling pathway. The *cysE* gene product, O acetyl serine is involved in this process. Thus, conversion of methionine to cysteine would require an upregulation of *cysE* expression. However, the cessation of growth in the NF337.1 cysteine limited culture at OD_{600} 1.1 indicates that the culture is starved for cysteine. Thus, an increase in expression of the $P_{gltX\text{-}T\text{-}box\text{-}cysES}$-*gfp* fusion would be expected.

Further experimental analysis is required to work out the effect that the presence of methionine has on the expression of the *cysES* genes in B. subtilis.

5.7.3 The $P_{gltX\text{-}T\text{-}box\text{-}cysES}$-*gfp* fusion is induced by starvation for tryptophan

Tryptophan and cysteine occupy the same mixed codon box. Therefore, their codons and the tRNA anticodons that they interact with are identical at the first two positions. We sought to test whether $tRNA^{TRP}$ could interact with the *cysES* T-box regulatory element to induce antitermination in response to limitation for tryptophan. To analyze the effect of tryptophan starvation on the *cysES* T-box regulatory element we integrated

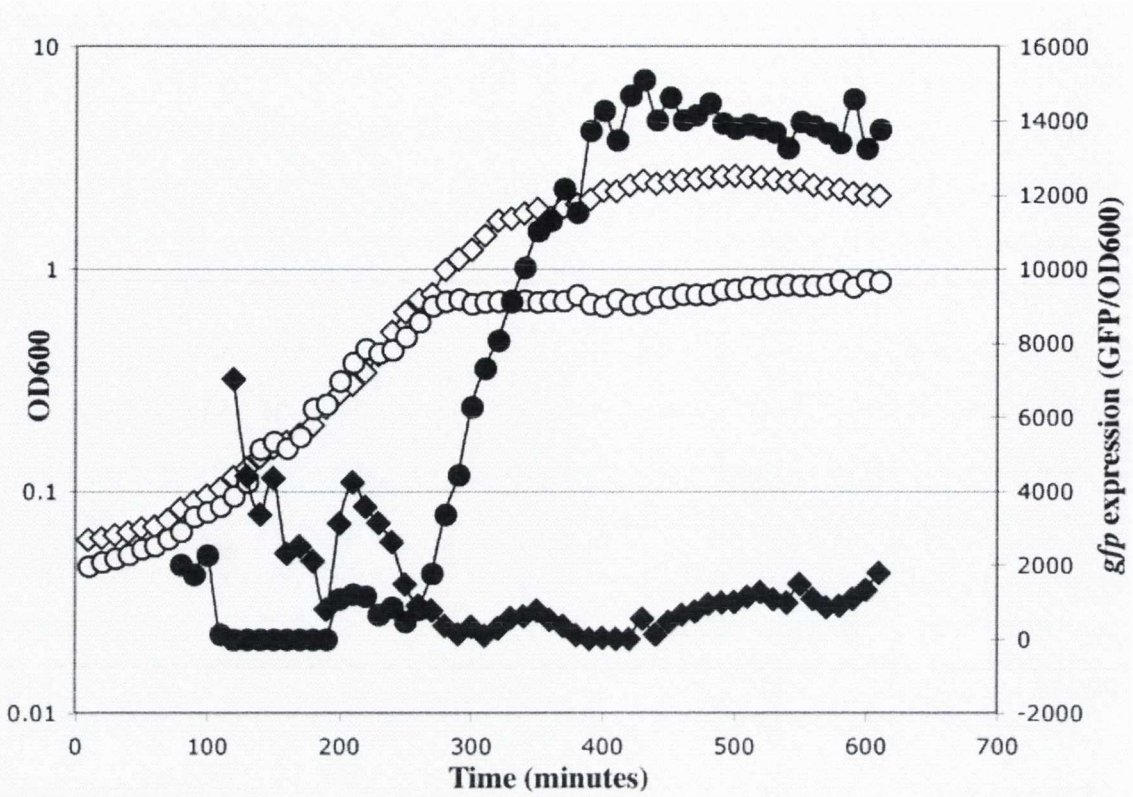


Figure 5.33 Analysis of the response of the *B. subtilis gltX* promoter and *cysES* T-box regulatory element to starvation for tryptophan. The graph displays growth and *gfp* expression data for NF333.1 [*trpC2*::pNF329.1 ($P_{gltX\text{-}T\text{-}box\text{-}cysES}$-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml tryptophan and 1µg/ml tryptophan. Growth is represented by open symbols. Expression values are represented by closed symbols. NF333.1 in BLM containing 50µg/ml tryptophan (◇, ◆). NF333.1 in BLM containing 1µg/ml tryptophan (○, ●).

the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion onto the chromosome of *B. subtilis* 168 to produce strain NF333.1. Strain NF333.1 was grown in BLM containing 50µg/ml tryptophan (tryptophan replete) and 1µg/ml tryptophan (tryptophan limiting) and the effect of starvation for tryptophan on *gfp* expression was examined. The results are presented in figure 5.33.

Growth of strain NF333.1 in tryptophan limiting conditions begins to slow and plateau at ~ OD_{600} 0.8. At this point there is a concomitant and rapid increase in *gfp* expression from approximately 1000 units to ~ 12000 units over the course of 100 minutes. The *gfp* expression profile of strain NF333.1 grown in tryptophan replete conditions is unaltered by the slow down in growth of the culture as it enters stationary phase and remains at approximately 1000 units throughout the growth curve. The results from this experiment show that expression of the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion is induced by limitation for tryptophan.

5.7.4 Mutation of the specifier codon of the *cysES* T-box regulatory element does not affect expression of the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion in response to cysteine or tryptophan limitation

We hypothesized that it may be possible for uncharged $tRNA^{TRP}$ to induce antitermination of the *cysES* T-box regulatory element as the anticodons of $tRNA^{TRP}$ and $tRNA^{CYS}$ are identical at two of the three bases of the anticodon and only differ at the wobble base. In addition, $tRNA^{CYS}$ and $tRNA^{TRP}$ share a high degree of sequence homology in their single stranded stem loop regions, which have been shown to be important in tRNA-leader RNA interactions (van de Guchte, *et al.*, 2001). However they differ at the discriminator base. $tRNA^{TRP}$ has a G at this position whereas $tRNA^{CYS}$ has a U. This would mean a sub-optimal G-A interaction for $tRNA^{TRP}$ at the *cysES* T-box sequence. Nonetheless we observed an increase in *gfp* expression in the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion as a result of tryptophan starvation. We sought to establish if the increase in expression of the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion as a result of tryptophan starvation was due to interaction of uncharged $tRNA^{TRP}$ with the *cysES* T-box regulatory element. To test this we constructed $P_{gltX\text{-T-box-}cysES}$-*gfp* fusions with mutations at the specifier codon. If uncharged $tRNA^{TRP}$ can interact with the *cysES* T-box regulatory element to induce antitermination, then mutation of the specifier codon should alter the *gfp* expression in response to tryptophan starvation. The $P_{gltX\text{-T-box-}cysES}$-*gfp* fusions containing altered specifier codons would also allow us to assess the effect of mutations of the specifier codon on the ability of $tRNA^{CYS}$ to interact with the *cysES*

T-box regulatory element. For example mutation of the UGC specifier codon for cysteine to a UGG codon for tryptophan should improve the ability of $tRNA^{TRP}$ to interact with the specifier loop whereas mutation of the specifier codon to a UAA stop codon should have a negative effect on the ability of either $tRNA^{TRP}$ or $tRNA^{CYS}$ to interact with the specifier loop. Also, a change to the specifier codon might result in significantly lower gfp expression in the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion.

To assess the effect of specifier codon mutations on the induction of expression of the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion we generated a number of new $P_{gltX\text{-T-box-}cysES}$-*gfp* fusions containing mutated specifier codons. The mutated specifier codons were UGU (cysteine), UGG (tryptophan) and the UGA and UAA stop codons. These altered $P_{gltX\text{-T-box-}cysES}$-*gfp* fusions were integrated in single copy onto the chromosome of the *B. subtilis* strains 168 (*trpC2*) and 1A79 (*cysC1 trpC2*) and their *gfp* expression profiles in response to tryptophan and cysteine starvation were assessed. For a description of these strains see materials and methods.

Strains containing the altered $P_{gltX\text{-T-box-}cysES}$-*gfp* fusions integrated into the *B. subtilis* strain 1A79 (*cysC1 trpC2*) were NF335.2 ($P_{gltX\text{-T-box-}cysES\ UGG}$-*gfp*), NF336.1 ($P_{gltX\text{-T-box-}cysES\ UAA}$-*gfp*), NF337.1 ($P_{gltX\text{-T-box-}cysES\ UGC}$-*gfp*) and NF338.1 ($P_{gltX\text{-T-box-}cysES\ UGU}$-*gfp*). Data from the experiments carried out with these strains is presented in figure 5.34.

Growth curves were carried out in BLM as previously described with 5µg/ml methionine added to the medium. The data shows that mutation of the specifier codon of the *cysES* T-box regulatory element has no effect on the induction of *gfp* expression as a result of limitation for cysteine. In all the strains tested (NF335.2, NF336.1, NF337,1 and NF338.1) there is an increase in expression during exponential growth from approximately 200 units to ~20000 units when grown in BLM with a limiting amount of cysteine (Figure 5.34A-D). This activity then falls as the cultures reach late exponential growth and decreases steadily as the cultures become limited for cysteine and growth is halted. In the cysteine replete cultures there is little or no *gfp* expression at any point in the growth curve.

This result indicates that while the *cysES* T-box regulatory element responds to limitation for cysteine, it is unlikely that this response occurs via interaction of uncharged $tRNA^{CYS}$ with the *cysES* T-box regulatory element. Mutation of the specifier codon to the codons UGU and UGG resulted in a virtually identical profile of *gfp* expression. The possibility that this is because the wobble base of the specifier codon is non-discriminatory in terms of tRNA anticodon recognition, and can therefore read UGC, UGU and UGG equally well is ruled out, as the level of *gfp*

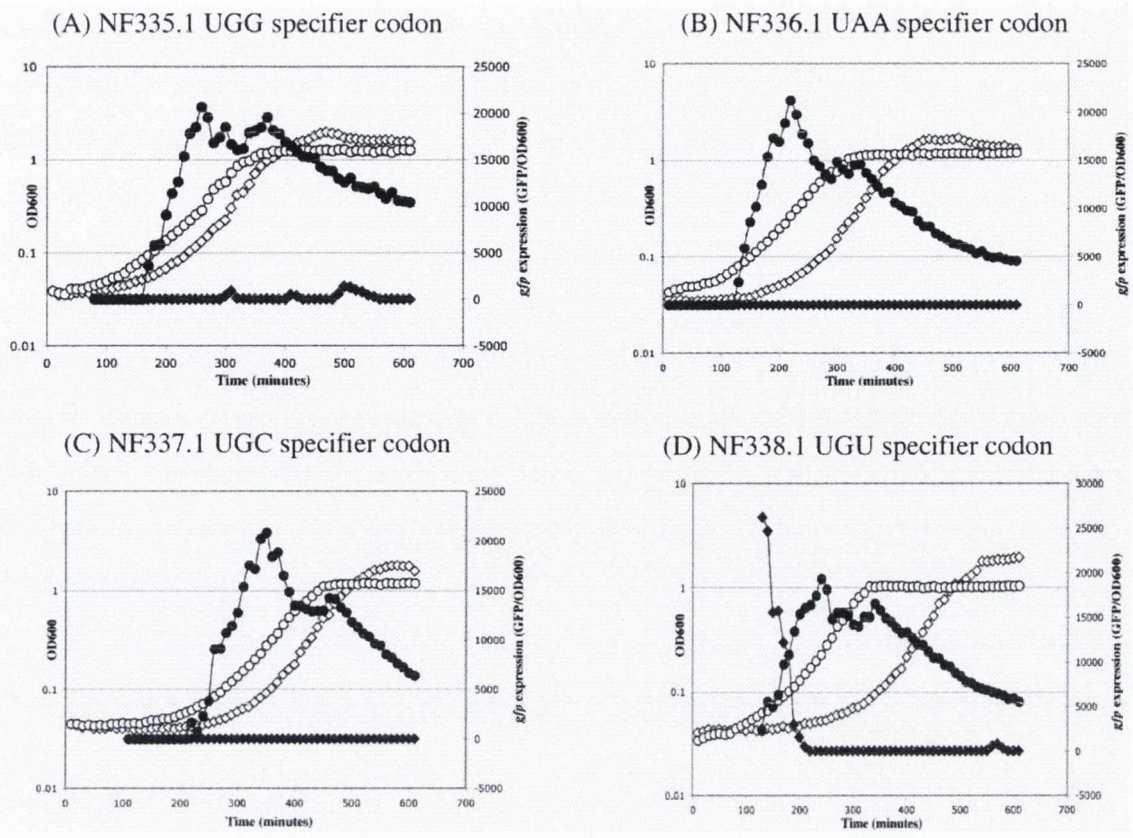


Figure 5.34 Analysis of the response of the *B. subtilis gltX* promoter and *cysES* T-box regulatory elements containing the specifier codons UGG, UAA, UGC, and UGU to starvation for cysteine during growth in BLM containing 5µg/ml methionine. The graphs (A-D) display growth and *gfp* expression data for NF335.1 NF336.1, NF337.1 and NF338.1 containing the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion with the UGG, UAA, UGC and UGU specifier codons respectively, grown in BLM containing 20µg/ml cysteine and 1µg/ml cysteine. Growth is represented by open symbols. Expression values are represented by closed symbols. Strains grown in BLM containing 20µg/ml cysteine are represented by (◇, ◆). Strains grown in BLM containing 1µg/ml cysteine are represented by (○, ●).

expression for the $P_{gltX\text{-T-box-}cysES\ UAA}$-*gfp* fusion containing the UAA specifier codon which contains only one complementary base to the $tRNA^{CYS}$ anticodon resulted in an identical expression profile. Regulation of *cysES* transcription does not appear to be by a T-box mechanism.

To test the response of the altered $P_{gltX\text{-T-box-}cysES}$-*gfp* fusions to limitation for tryptophan, strains containing the altered $P_{gltX\text{-T-box-}cysES}$-*gfp* fusions were integrated into the *B. subtilis* strain 168 (*trpC2*). The resultant strains were NF331.1 ($P_{gltX\text{-T-box-}cysES\ UGG}$-*gfp*), NF332.1 ($P_{gltX\text{-T-box-}cysES\ UAA}$-*gfp*), NF333.1 ($P_{gltX\text{-T-box-}cysES\ UGC}$-*gfp*), NF334.1 ($P_{gltX\text{-T-box-}cysES\ UGU}$-*gfp*) and NF347.1 ($P_{gltX\text{-T-box-}cysES\ UGA}$-*gfp*). These strains were grown in BLM containing either 50μg/ml tryptophan or 1μg/ml tryptophan as described previously. Data from the experiments carried out with these strains is presented in figure 5.35A-E.

Growth of strains NF331.1, NF332.1, NF333.1, NF334.1 and NF347.1 in tryptophan limiting conditions begins to slow and plateau at ~ OD_{600} 0.8-1.0. As growth in these cultures plateaus there is a coincident increase in *gfp* expression from around zero to approximately 10000 units over the course of 100 minutes. This increase in *gfp* expression in response to tryptophan limitation is similar for strains NF331.1, NF332.1, NF333.1, NF334.1 and NF347.1 and is therefore not affected by mutations of the specifier codon. This suggests that $tRNA^{TRP}$ does not interact with the *cysES* T-box regulator element and that the increase in expression of the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion in response to tryptophan limitation does not occur by a T-box mechanism.

5.7.5 Regulation of expression of the *cysES* genes occurs at the initiation of transcription at the *gltX* promoter

Data from our analysis of the *cysES* T-box regulatory element had shown that mutation of the specifier codon had no effect on induction of expression of a $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion as a result of cysteine limitation. We concluded that regulation of expression of the *cysES* genes does not occur by a T-box mechanism. This led us to question whether or not the *gltX* promoter plays a role in the regulation of expression of the *cysES* genes.

In order to examine this we made a transcriptional fusion of the *gltX* promoter region to the GFP reporter. The resultant plasmid was pNF339 (P_{gltX}-*gfp*). This plasmid was then integrated in single copy into the chromosome of the *B. subtilis* strain 1A79 (*cysC1 trpC2*) in order to examine the effects of cysteine starvation on the P_{gltX}-*gfp* fusion. The resultant strain was NF345.1 (*cysC1 trpC2* P_{gltX}-*gfp*).

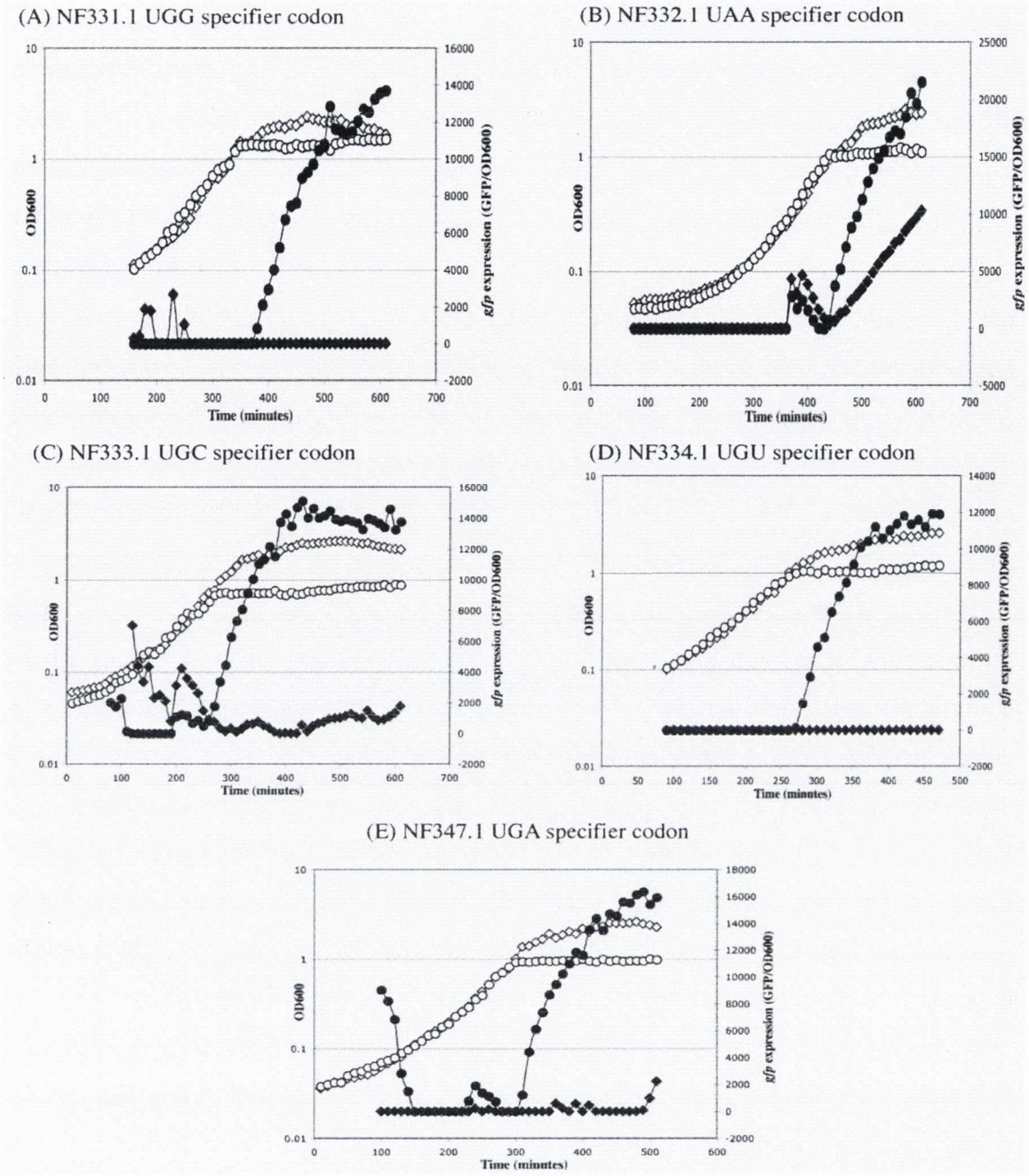


Figure 5.35 Analysis of the response of the *B. subtilis gltX* promoter and *cysES* T-box regulatory elements containing the specifier codons UGG, UAA, UGC, UGU and UGA to starvation for tryptophan . The graphs (A-E) display growth and *gfp* expression data for strains NF331.1 NF332.1, NF333.1, NF334.1 and NF347.1 containing the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion with the UGG, UAA, UGC, UGU and UGA specifier codons respectively. Cultures were grown in BLM containing 50µg/ml tryptophan and 1µg/ml tryptophan. Growth is represented by open symbols. Expression values are represented by closed symbols. Strains grown in BLM containing 50µg/ml tryptophan are represented by (◇, ◆). Strains grown in BLM containing 1µg/ml tryptophan are represented by (○, ●).

Strain NF345.1 was grown in BLM containing a range of cysteine concentrations and the *gfp* expression was determined. The data is presented in figure 5.36A.

Growth of strain NF345.1 in cysteine limiting conditions begins to slow and plateau at ~ OD_{600} 0.9. At this point there is a coincident and rapid increase in *gfp* expression from approximately 2000 units to ~ 6000 units. Expression of the P_{gltX}-*gfp* fusion then decreases as the culture continues into stationary phase. This is a very similar pattern of *gfp* expression to that seen in strain NF337.1 containing the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion grown in the same conditions (Figure 5.31). In the cysteine replete NF345.1 culture, *gfp* expression reaches ~ 2000 units during exponential growth. This level then falls to ~ 500 units before increasing to 2000 units again as the culture enters stationary phase (Figure 5.36A). Strain NF337.1 grown in similar conditions shows no *gfp* expression until late in stationary phase.

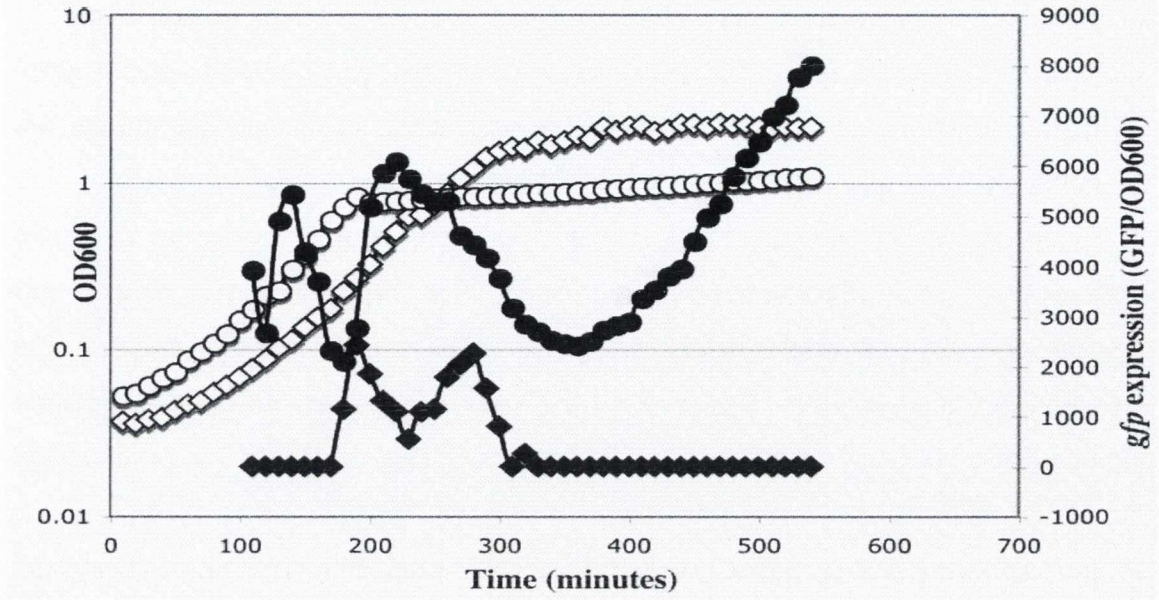
Strain NF345.1 was also grown in BLM containing a range of cysteine concentrations and a constant level of methionine of 5µg/ml and the gfp expression was determined. The data is presented in figure 5.36B.

Growth of strain NF345.1 in cysteine limiting conditions begins to plateau at ~ OD_{600} 1. Expression of the P_{gltX}-*gfp* fusion in this culture increases to approximately 25000 units during exponential growth and then begins to decrease shortly after growth is halted due to cysteine limitation. In cysteine replete conditions there is little or no *gfp* expression during exponential growth. As the culture enters stationary phase there is a brief increase in *gfp* expression to approximately 3000 units before activity decreases again. The profile of *gfp* expression in strain NF345.1 in cysteine limiting and replete conditions with constant 5µg/ml methionine is similar to that seen in NF337.1 grown under the same conditions (Figure 5.32). These results show that the profile of *gfp* expression in strains containing the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion and the P_{gltX}-*gfp* fusion are very similar. This indicates that the *cysES* T-box regulatory element has little effect on the regulation of transcription initiated from the *gltX* promoter and that the *gltX* promoter may be the primary regulatory element governing *cysES* expression.

5.7.6 Expression of the P_{gltX}-*gfp* fusion is induced by limitation for tryptophan

Limitation for tryptophan induces increased expression of a $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion. We sought to establish if the response to tryptophan limitation also occurs at the *gltX* promoter rather than at the *cysES* T-box regulatory element. To investigate this we

(A)



(B)

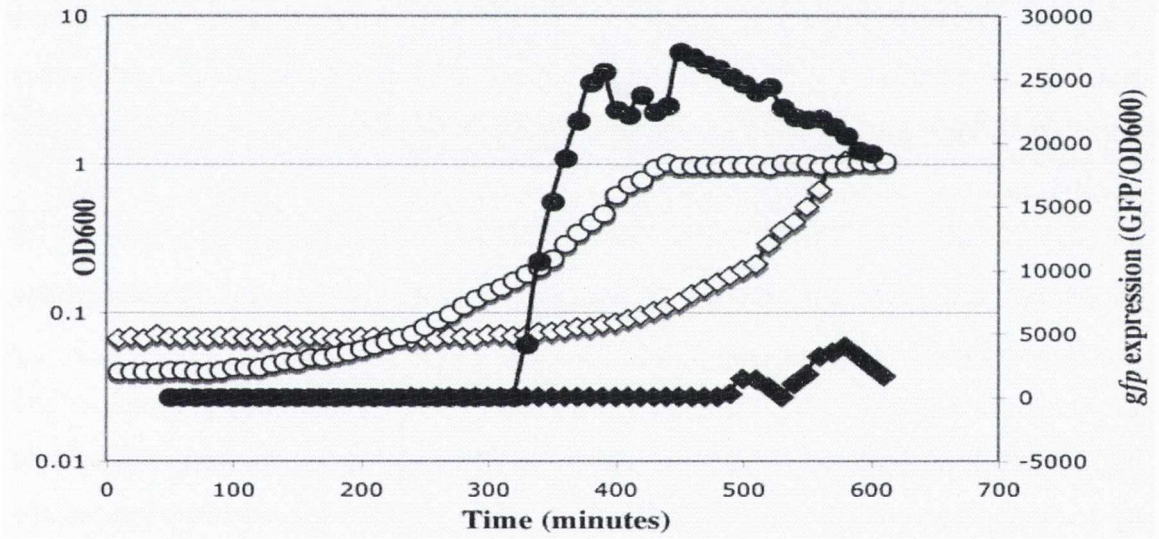


Figure 5.36 Analysis of the response of the *B. subtilis gltX* promoter to limitation for cysteine both with and without addition of 5µg/ml methionine. (A) Growth and *gfp* expression data for NF345.1 [*cysC1 trpC2*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$] grown in BLM containing 20µg/ml cysteine and 1µg/ml cysteine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF345.1 in BLM containing 20µg/ml cysteine (◇, ◆). NF345.1 in BLM containing 1µg/ml cysteine (○, ●). (B) Growth and *gfp* expression data for NF345.1 grown in BLM containing 5µg/ml methionine with either 20µg/ml cysteine or 1 µg/ml cysteine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF345.1 in BLM containing 20µg/ml cysteine (◇, ◆). NF345.1 in BLM containing 1µg/ml cysteine

transformed the plasmid pNF339 into *B. subtilis* strain 168 in single copy at the amylase locus to produce the strain NF343.1 (*trpC2* P_{gltX}-*gfp*). Strain NF343.1 was grown in BLM containing 50μg/ml tryptophan (tryptophan replete) and 1μg/ml tryptophan (tryptophan limiting) and *gfp* expression was assessed. The results are presented in figure 5.37.

Growth of strain NF343.1 in tryptophan limiting conditions begins to slow and plateau at ~ OD_{600} 0.9. As growth is halted there is a concomitant increase in *gfp* expression from ~ 6500 units to approximately 18000 units within 100 minutes. The P_{gltX}-*gfp* fusion is expressed at a level of approximately 5000 units throughout exponential growth in the tryptophan replete culture (Figure 5.37). The level of *gfp* expression then slowly rises in stationary phase.

The profiles of *gfp* expression of strain NF343.1 (P_{gltX}-*gfp*) in tryptophan limiting and replete conditions are similar to those observed in strain NF333.1 ($P_{gltX\text{-T-box-}cysES}$-*gfp*) under the same growth conditions (Figure 5.33). This shows that the *cysES* T-box regulatory element plays little or no role in regulation of expression of the *cysES* genes under conditions of tryptophan limitation and that control of expression is exerted at the *gltX* promoter.

5.7.7 Expression of the the *gltX-cysES* operon is induced by limitation for phenylalanine, methionine and lysine

Expression of the *gltX-cysES* genes is induced by limitation for cysteine and tryptophan. We sought to establish whether depletion for any other amino acids could induce expression of the P_{gltX}-*gfp* fusion. In order to test this we transformed the plasmid pNF339 (P_{gltX}-*gfp*) into the *B. subtilis* strains JH642 (*pheA1*), 1A75 (*ilvA1 leuB8 metB5*) and 1A765 (*lys trpC2*) to produce the strains NF344.1 (*pheA1* P_{gltX}-*gfp*), NF351.1 (*ilvA1 leuB8 metB5* P_{gltX}-*gfp*) and NF352.1 (*lys trpC2* P_{gltX}-*gfp*) and analyzed the effect of starvation for phenlalanine, methionine and lysine on *gfp* expression in the respective strains. The results are presented in figure 5.38.

Growth of strains NF344.1, NF351.1 and NF352.1 begins to slow and plateau at ~ OD_{600} 0.9-1.0 as they become limited for their respective essential amino acid. In each case, as growth plateaus, there is a coincident increase in *gfp* expression. The level of induction is greatest in strain NF352.1 starved for lysine. Expression of the P_{gltX}-*gfp* fusion increases from 1000 units to ~ 19000 units within 100 minutes in this strain. In strains NF344.1 and NF351.1 *gfp* expression reaches 9000 units and 12000 units respectively 100 minutes after induction due to amino acid limitation (Figure 5.38).

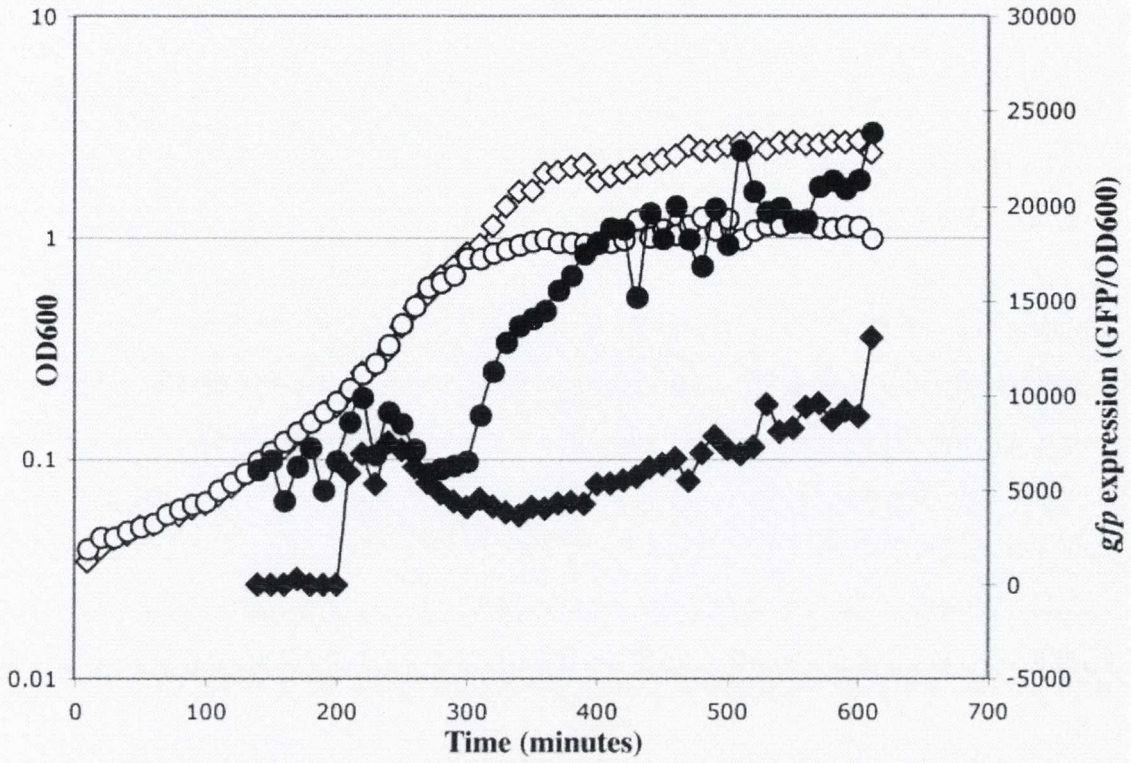


Figure 5.37 Analysis of the response of the *B. subtilis gltX* promoter to limitation for tryptophan. The graph displays growth and *gfp* expression data for strain NF343.1 [*trpC2*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml tryptophan and 1µg/ml tryptophan. Growth is represented by open symbols. Expression values are represented by closed symbols. NF343.1 in BLM containing 50µg/ml tryptophan (◇, ◆). NF343.1 in BLM containing 1µg/ml tryptophan (○, ●).

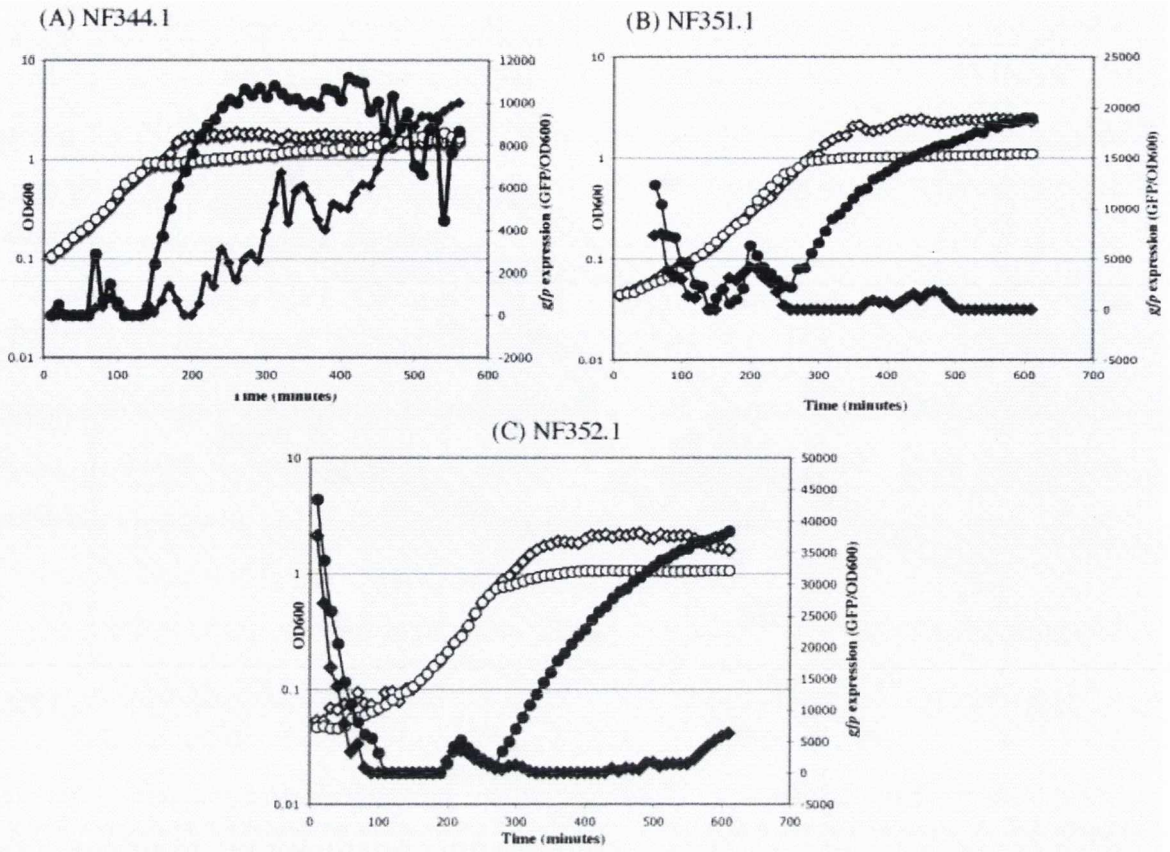


Figure 5.38 Analysis of the response of the *B. subtilis gltX* promoter to starvation for phenylalanine, methionine and lysine. (A) The graph displays growth and *gfp* expression data for NF344.1 [*pheA1 trpC2*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml phenylalanine and 4µg/ml phenylalanine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF344.1 in BLM containing 50µg/ml phenylalanine (◇, ◆). NF344.1 in BLM containing 4µg/ml phenylalanine (○, ●). (B) The graph displays growth and *gfp* expression data for NF351.1 [*ilvA1 leuB8 metB5*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml methionine and 3µg/ml methionine. Growth is represented by open symbols. Expression values are represented by closed symbols. NF351.1 in BLM containing 50µg/ml methionine (◇, ◆). NF351.1 in BLM containing 3µg/ml methionine. (C) The graph displays growth and *gfp* expression data for NF352.1 [*lys trpC2*::pNF339.1 (P_{gltX}-*gfp*) $Spec^R$] grown in BLM containing 50µg/ml lysine and 10µg/ml lysine. Growth is represented by open symbols. Exression values are represented by closed symbols. NF352.1 in BLM containing 50µg/ml lysine (◇, ◆). NF352.1 in BLM containing 10µg/ml lysine.

The level of *gfp* expression in strains NF351.1 and NF352.1 grown in BLM containing replete levels of their essential amino acids peaks during mid-exponential growth at a level of ~ 5000 units. This level of activity then decreases and remains low during stationary phase. The level of expression in strain NF344.1 grown in phenylalanine replete conditions is low during exponential growth but then increases as growth slows and the culture enters stationary phase. This increase in expression occurs at a slower rate than that seen in strain NF344.1 grown in phenylalanine limiting conditions.

These results show that expression of the P_{gltX}-*gfp* fusion is greater in amino acid limiting conditions relative to amino acid replete conditions. This shows that expression of the P_{gltX}-*gfp* fusion is induced by general amino acid starvation.

5.7.8 An inverted repeat sequence present in the *gltX* promoter region may be involved in the regulation of expression of the *gltX* and *cysES* genes

Following the observation that a P_{gltX}-*gfp* fusion responds to general amino acid starvation we sought to establish whether any DNA/RNA secondary structures exist at the *gltX* promoter locus that could possibly be involved in the regulation of transcription. In order to investigate this we examined the *gltX* promoter region nucleotide sequence for any potential regulatory structures or domains.

Our analysis uncovered an inverted repeat seqeunce (Figure 5.39) that overlaps with the –10 region of the *gltX* promoter. Such a sequence could potentially bind a protein which could compete with RNA polymerase for binding to the *gltX* promoter and thus have a role in the regulation of transcription initiation. Further analysis is required to validate this hypothesis.

```
ACTTGTTTTGATGCCGGGTCTATTTGGTGGTAGAATAGATTCATACATTTT
      -35                          -10
```

Figure 5.39 Sequence of the *gltX* promoter region. The -35 and -10 regions are highlighted in red. The inverted repeat sequence is indicated in green.

5.7.9 Conclusions about the *cysES* T-box regulatory element

It has previously been reported that the expression of a $P_{thrS\text{-}T\text{-}box\text{-}cysES}$-*lacZ* fusion is repressed 3-fold upon overexpression of the *cysS* gene *in trans* (Gagnon, *et al.*, 1994). This suggests that antitermination in the *cysES* T-box regulatory element is responsive to the level of charged $tRNA^{CYS}$ in the cell. Results from our analysis of the *cysES* T-box regulatory element suggest that this is not the case. Mutation of the specifier codon should have an effect on expression of the $P_{gltX\text{-}T\text{-}box\text{-}cysES}$-*gfp* fusion. However, our analysis showed that these mutations had no effect on induction of expression as a result of cysteine limitation. This suggests that $tRNA^{CYS}$ does not interact with the specifier codon and that regulation of *cysES* expression does not occur by a T-box mechanism. Additionally we showed that expression of the $P_{gltX\text{-}T\text{-}box\text{-}cysES}$-*gfp* fusion is induced by limitation for tryptophan. This led us to investigate if the regulation of expression of the *cysES* genes occurs at the *gltX* promoter. Experiments to test the response of the P_{gltX}-*gfp* fusion to limitation for the amino acids cysteine, tryptophan, phenylalanine, methionine and lysine all showed that there is an induction of *gfp* expression in response to amino acid limitation. This shows that initation of transcription from the *gltX* promoter is responsive to general amino acid limitation.

Chapter 6

Discussion

6.1 Introduction

The T-box antitermination mechanism is a method of transcriptional control which alters the level of expression of genes involved in amino acid biosynthesis, amino acid transport and tRNA charging in response to the ratio of charged to uncharged tRNA in the cell. This method of regulation is most frequently found to control aminoacyl tRNA synthetase genes and the T-box regulatory elements of both *tyrS* and *thrS* in *B. subtilis* are well characterized. Our work began by determining the frequency at which expression of the each AARS gene is controlled by a T-box mechanism. The results showed that expression of lysyl tRNA synthetases by a T-box mechanism occurs along with *gltX* at the lowest frequency. It was interesting and unusual therefore, to identify a T-box regulatory element upstream of the *lysK* gene of *B. cereus* (Ataide, *et al.*, 2005). This led us to investigate the expression of the *lysK* gene of *B. cereus* in order to (i) establish if the *lysK* T-box regulatory element is functional and responds to a reduction in $tRNA^{LYS}$ charging (ii) establish if a T-box regulated *lysK/S* gene can support growth of *B. subtilis* and (iii) assess what features of the *lysK* T-box regulatory element may make it suboptimal for regulation of *lysK/S* expression

6.2 Distribution of T-box regulation amongst the AARS genes

The discovery of a putative T-box regulatory element upstream of the *B. cereus lysK* gene was the first instance of a T-box regulated lysyl tRNA synthetase gene to be identified (Ataide, et al., 2005). This discovery led us to carry out a search among 891 completely sequenced bacterial genomes in order to determine the relative distribution of T-box regulation for the various AARS genes. The results showed that *lysS/K* genes are rarely regulated by T-box mechanisms when compared with other AARS genes. For example, while there are 146 and 103 instances of T-box regulation for *ileS* and *pheS* respectively, putative T-box regulatory elements were identified upstream of genes coding for lysyl tRNA synthetase in only six instances. Four of theses T-box regulatory elements were found upstream of *lysK* genes in *B. cereus* and *B. thuringiensis* species whereas the other two were found upstream of *lysS* genes in *S. thermophilum* and *C. beijerinckii*. An interesting feature of this is that in all bacteria that contain a T-box regulated *lysS/K* gene, a second copy of *lysS/K* is encoded in the genome. This suggests perhaps that one copy is responsible for the housekeeping function of $tRNA^{LYS}$ charging while the other may have some other role in the cell. In the case of *B. cereus* it is known that the class II *lysS* encodes the gene responsible for $tRNA^{LYS}$ charging during exponential growth while the T-box regulated class I *lysK* is expressed predominantly in

stationary phase (Ataide, *et al.*, 2005). A similar arrangement may be found in *S. thermophilum* and *C. beijerinckii*. We noted that in the case of *C. beijerinckii* which encodes two class II *lysS* genes, the two copies have only 45% sequence similarity. *C. acetobutylicum* which is *C. beijerinckii's* closest relative (Collins, *et al.*, 1994) only encodes one *lysS* gene which is more closely related to the non T-box regulated *lysS* gene of *C. beijerinckii*. This may be an indication that, as is the case for *B. cereus*, the T-box regulated *lysS* gene does not encode the housekeeping LysRS. These observations suggest the possibility that a single *lysS* gene, regulated by a T-box mechanism, is not optimal for growth.

Work carried out by Shaul, *et al*, (2006) indicated that the class I *lysK* found in *B. cereus* and *B. thuringiensis* clusters with the class I LysK from the archaebacterium *Pyrococcus*. Therefore, LysK is likely to have been acquired by *B. cereus* and *B. thuringiensis* by lateral gene transfer from *Pyrococcus*. As there are no T-box regulatory elements in archaebacteria, this raises the question as to how the *lysK* in *B. cereus* and *B. thuringiensis* came to be T-box regulated. It is possible that the T-box regulatory element found upstream of the *B. cereus lysK* gene might be derived from an asparagine T-box regulatory element. In theory, this conversion of specificity from $tRNA^{ASN}$ to $tRNA^{LYS}$ would only require a single point mutation at the specifier codon to convert it from an AAC asparagine codon to an AAA codon specific for $tRNA^{LYS}$. With these considerations in mind, we chose to establish (i) if the *lysK* T-box regulatory element is functional in a canonical way; (ii) whether a bacterium in which the expression of the class II *lysS* was controlled by a T-box mechanism was viable and (iii) whether the *lysK* T-box regulatory element is responsive to the level of charged asparaginyl tRNA.

6.3 The *lysK* T-box regulatory element is functional

Analysis of the sequence of the T-box regulatory element upstream of the *lysK* gene of *B. cereus* showed that it contains the essential features of a functional T-box regulatory element, namely a specifier codon complementary to the anticodon of $tRNA^{LYS}$ and a canonical T-box sequence. In view of our analysis showing that regulation of *lysS/K* genes by a T-box regulatory element is unusual, we investigated whether the *B. cereus lysK* T-box regulatory element is functional.

To investigate this we examined expression of $P_{T\text{-}box\ lysK}$-*lacZ* reporter fusions and tested (i) is the $P_{T\text{-}box\ lysK}$-*lacZ* fusion inducible by starvation for lysine in a lysine auxotroph strain (ii) is the $P_{T\text{-}box\ lysK}$-*lacZ* fusion inducible by reduced charging of $tRNA^{LYS}$ in a strain containing an IPTG inducible *lysS* gene and (iii) is expression of the is the $P_{T\text{-}box\ lysK}$-*lacZ*

fusion in response to reduced $tRNA^{LYS}$ charging affected by mutation of the specifier codon. The results showed that the $P_{T\text{-}box\ lysK}$-*lacZ* fusion was induced by reduced $tRNA^{LYS}$ charging in both a lysine auxotroph background and in a strain containing an IPTG inducible *lysS* gene. This showed that the *lysK* T-box regulatory element is functional and responds to reduced charging of $tRNA^{LYS}$. Mutations of the specifier codon from an AAA lysine codon to an AAG lysine codons and the AAU and AAC codons for asparagine all resulted in lower expression of the $P_{T\text{-}box\ lysK}$-*lacZ* fusion is response to lysine starvation further demonstrating that the lysK T-box regulatory element is functional and sensitive to the level of $tRNA^{LYS}$ charging. We verified specificity by showing that reduced growth rate was not responsible for the increase in expression seen in the $P_{T\text{-}box\ lysK}$-*lacZ* fusion by growing a strain containing the $P_{T\text{-}box\ lysK}$-*lacZ* fusion is phenylalanine limiting conditions. Analysis of the expression of the $P_{T\text{-}box\ lysK}$-*lacZ* fusion in phenylalanine starvation conditions showed no increase in β-galactosidase activity upon cessation of growth indicating that the induction of expression was specific to limitation for charged $tRNA^{LYS}$, and not due to reduced growth rate. We conclude from this analysis that the *lysK* T-box regulatory element from *B. cereus* is functional and responsive to reduced charging of $tRNA^{LYS}$.

6.4 Control of *lysS* expression by the *lysK* T-box regulatory element can support growth of *B. subtilis*

In *B. subtilis*, the *lysS* gene is the last gene in an operon containing genes for folate biosynthesis and trytophan metabolism. This genetic arrangement is widely conserved amongst the *Bacilli* including *B. cereus*. As there are no identified cases of a bacterium containing a single *lysS/K* gene that is regulated by a T-box mechanism we were curious as to whether such a genetic arrangement could support growth of *B. subtilis*. To test this we asked (i) can the *lysK* gene from *B. cereus* under the control of its T-box regulatory element support growth of *B. subtilis* and (ii) can LysRS support growth of *B. subtilis* when expression is regulated by a T-box mechanism. The results showed that strains containing either a T-box regulated *lysK* or a T-box regulated *lysS* were viable. The $P_{T\text{-}box\ lysK}$-*lysK* containing strain was shown to have a slow growth phenotype in rich medium and could not be grown in minimal media. This phenotype is likely to be caused by the reduced catalytic efficiency of the class I *lysK* enzyme relative to the class II enzyme and is not due to any feature of the T-box regulatory element (Wang, *et al.*, 2006). Our analysis of the strain containing the $P_{T\text{-}box\ lysK}$-*lysS* arrangement showed that this strain grew similarly to wild type in both nutrient rich and nutrient limiting conditions. Therefore, while it is rare to

find a T-box regulated *lysS/K* gene, a strain containing a single *lysS/K* gene whose expression is regulated by a T-box mechanism is viable.

6.5 Why is the expression of *lysS/K* not commonly regulated by a T-box mechanism?

Our analysis of a *B. subtilis* strain where expression of *lysS* is controlled by a T-box mechanism showed that this regulatory arrangement is viable and supports normal growth. Given that this method of regulation is so commonly utilized to control expression of AARS genes in *B. subtilis*, we wanted to investigate why this method of regulation occurs so infrequently with *lysS* genes. We noted that those AARS genes not regulated by a T-box mechanism in *B. subtilis* have cognate amino acids that are generally found in mixed codon boxes. The AARS genes not regulated by a T-box mechanism in *B. subtilis* are *argS, asps, asnS, metS, lysS*, proS and *gltX*. All these genes code for synthetases that charge tRNAs with amino acids from mixed codon boxes with the exception of *proS*. We therefore proposed that *lysK* may not be regulated by a T-box mechanism due to a lack of specificity, i.e. the *lysK* T-box regulatory element may not be able to distinguish between its cognate tRNA and the tRNA for the amino acid in the same mixed codon box ($tRNA^{ASN}$) as they share two out of three nucleotides at the anticodon. A prediction of this "mixed codon box hypothesis" is that $tRNA^{ASN}$ can interact with the *lysK* T-box regulatory element and promote anti-termination. We had also noted the possibility that the *lysK* T-box regulatory element may have evolved from a T-box regulatory element responsive to $tRNA^{ASN}$. Analysis of the nucleotide sequences of the asparagine synthase gene *asnA* and *lysK* T–box regulatory elements in *B. cereus* show some sequence homology but not enough to say with certainty that the *lysK* T-box regulatory element evolved from a T-box regulatory element specidic for asparagine. Additionally, analysis of the sequence and secondary structure of both $tRNA^{ASN}$ and $tRNA^{LYS}$ from *B. subtilis* show extensive sequence similarity between the two in the single stranded stem loop regions of the D loop and TφC loop regions.

AAT	**Asn**
AAC	**Asn**
AAA	**Lys**
AAG	**Lys**

Table 6.1 The mixed codon box for lysine and asparagine. Nucleotides highlighted in red indicate the similarity between the codons and therefore the anticodons of lysine and asparagines cognate tRNAs.

As these regions have been shown to be important in the T-box leader-tRNA interaction in some systems (Van de Guchte, *et al.*, 2001; Grundy, *et al.*, 2000), this is an additional indication that it may be possible that the *lysK* T-box leader region can recognize and interact with $tRNA^{ASN}$ as well as $tRNA^{LYS}$. Therefore, we asked whether the *lysK* T-box regulatory element is responsive to $tRNA^{ASN}$.

As no asparagine auxotroph was available we constructed a *B. subtilis* strain containing an IPTG inducible *asnS* gene which allowed us to alter the level of *asnS* expression and thus AsnRS levels in the cell and analyzed the effect of reduced $tRNA^{ASN}$ charging on the $P_{T\text{-}box\ lysK}$-*lacZ* fusion. Our results showed that a reduction in *asnS* expression results in increased expression of a $P_{T\text{-}box\ lysK}$-*lacZ* fusion suggesting that $tRNA^{ASN}$ is capable of interacting with the *lysK* T–box regulatory element to trigger antitermination. As a control experiment to show that a reduction in AsnRS levels in the cell did not cause a reduction in $tRNA^{LYS}$ charging and in so doing cause the increase in expression of the $P_{T\text{-}box\ lysK}$-*lacZ* fusion at low IPTG concentrations, we performed tRNA charging northern experiments. The data obtained from these analyses failed to detect small changes in the level of $tRNA^{LYS}$ charging and thus could not rule out definitively the possibility that a reduction in *asnS* expression also has an effect on the levels of charged $tRNA^{LYS}$. However, the *lacZ* data indicates that the *lysK* T-box regulatory element from *B. cereus* may be sensitive to the leve of charged $tRNA^{ASN}$.

We then carried out experiments to assess if the third position of the specifier codon of the *lysK* T-box regulatory element was non-discriminating in terms of its interaction with tRNA, i.e. would mutations at this position have any effect on antitermination by either $tRNA^{LYS}$ or $tRNA^{ASN}$. The results of this analysis show that mutation of the nucleotide at the third position of the specifier codon to a pyrimidine result in significantly reduced expression of the $P_{T\text{-}box\ lysK}$-*lacZ* fusion. This result was unexpected as the presence of a C or a U at the third position of the specifier codon produces asparagine codons. This result may indicate a requirement for a purine at the third position in the specifier loop to allow correct presentation of the specifier codon for tRNA binding. Analysis of the T-box loop on the 5' side of the anti-terminator structure has shown that the conserved ACC nucleotides of the loop are important to stabilize the conformation of the UGGN nucleotides for interaction with the acceptor end of tRNA (Gerdeman, *et al*, 2002). It is possible that replacement of a purine at the third position of the specifier codon with a pyrimidine also affects the conformation, thus preventing correct presentation of the specifier codon. Experiments to assess the effect of a purine-pyrimidine (G-C) mutation at the second position of the specifier loop showed that this resulted in reduced expression of

the $P_{T\text{-}box\ lysK}$-*lacZ* fusion in conditions of both reduced $tRNA^{LYS}$ and $tRNA^{ASN}$ charging, indicating that mutations of the specifier loop can alter the ability of tRNA to interact with the specifier codon.

Our analysis of the expression of the $P_{T\text{-}box\ lysK}$-*lacZ* fusions in an IPTG inducible *asnS* background suggest that the reason for the relative lack of T-box regulation for *lysS/K* genes may be due to a lack of specificity of the *lysS/K* T-box regulatory elements.

6.6 Investigation of the specificity of induction of T-box regulated AARS genes for amino acids in mixed codon boxes in *B. subtilis*

The observation that $tRNA^{ASN}$ may interact with the *lysK* T-box regulatory element of *B. cereus* to trigger antitermination prompted us to test whether the expression of other T-box regulated AARS encoding genes in *B. subtilis*, whose cognate amino acids are in mixed codon boxes responded to limitation for the other amino acid in that mixed codon box. We therefore constructed transcriptional fusions of the T-box regulatory elements located upstream of *leuS, pheS, hisS-aspS, ileS, trpS* and *cysES* and analyzed the effect of limitation of cognate amino acids and the other amino acid in each mixed codon box.

TTT	Phe	ATT	Ile	CAT	His	TGT	Cys
TTC	Phe	ATC	Ile	CAC	His	TGC	Cys
TTA	Leu	ATA	Ile	CAA	Gln	TGA	Stop
TTG	Leu	ATG	Met	CAG	Gln	TGG	Trp

Table 6.2 The mixed codon boxes for phenylalanine and leucine, isoleucine and methionine, histidine and glutamine and cysteine and tryptophan. Nucleotides in red indicate the similarities between the codons. The codons in each box differ only in the third "wobble" position.

The results of these analyses showed that none of the fusions tested were induced by limitation for the non-cognate amino acid that occupies the same mixed codon box as their cognate amino acid. For example, the $P_{T\text{-}box\ pheS}$-*gfp* fusion was induced by limitation for phenyalanine but not leucine. Analysis of the expression of the $P_{T\text{-box}\ pheS\ STOP}$-*gfp* fusion containing a UAA stop codon at the specifier codon position, showed an increase in *gfp* activity in response to phenylalanine limitation. The increase in *gfp* expression observed was significantly lower than that of the wild-type $P_{T\text{-box}\ pheS}$ -*gfp* fusion, however the increase in expression does indicate that $tRNA^{PHE}$ can still interact with the *pheS* T-box regulatory element when the specifier codon is a stop codon to induce antitermination. While it is possible that other features of $tRNA^{PHE}$ may stabilize its interaction with the

pheS T-box regulator element, this result suggests that there may be some flexibility in terms of tRNA anticodon-specifier codon interactions in some T-box regulatory elements.

6.7 Conclusions about the "mixed codon box hypothesis"

The data obtained from our analysis of the *B. cereus lysK* T-box regulatory element suggests that it is induced by depletion for charged $tRNA^{ASN}$ in addition to the cognate $tRNA^{LYS}$. However, when we tested to see if this is a feature of other T-box elements we found that it does not appear to be a general phenomenon. Why then does this situation exist for the *lysK* of *B. cereus*?

One possible explanation is that LysKRS is produced predominantly in stationary phase and therefore is not responsible for charging $tRNA^{LYS}$ during exponential growth. Efforts to knockout the *lysK* gene were unsuccessful indicating that it is required for some other stationary phase function. It has been shown to function in combination with LysRS to aminoacylate a novel tRNA known as $tRNA^{OTHER}$ (Ataide, *et al.*, 2005). The function of this tRNA has not yet been elucidated but work in our lab has shown that it is not essential (Foy and Devine, unpublished results). It has been suggested that the class I LysKRS may be involved in cell wall modification during stationary phase. Therefore LysKRS may play a role in the survival of the cell in challenging conditions and its being sensitive to starvation for two amino acids may be an advantage when cells are stressed and have entered stationary phase.

6.8 A number of T-box regulatory elements in *B. subtilis* appear to be non-responsive to reduced charging of their cognate tRNA

A characteristic feature of genes regulated by T-box regulatory elements is that their expression is induced upon starvation for their cognate amino acid. Our analysis of T-box regulatory elements in *B. subtilis* revealed that in some cases, they did not respond to limitation for their cognate amino acid as expected.

An analysis of the *leu*S coding region carried out by Vander Horn and Zahler suggested that the *leuS* gene may be regulated by a transcriptional attenuation mechanism possibly involving a T-box mechanism (Vander Horn and Zahler, 1992). However this was not demonstrated experimentally. Our investigation of the *leuS* T-box regulatory element showed little or no increase in *gfp* expression of the $P_{T\text{-}box\ leuS}$-*gfp* fusion in response to starvation for leucine. Furthermore, mutation of the specifier codon to a stop codon did not produce a change in *gfp* expression as a result of leucine limitation, as would be expected if *leuS* is regulated by a T-box mechanism. These results indicate that tRNA induced anti-

termination does not play a major role in the regulation of LeuRS synthesis under the conditions of our experiment. However, in some cases, T-box regulation is just one aspect of a complex system of regulation. For example, investigation of the *ilv-leu* T-box (possesses the same specifier codon as the *leuS* T-box regulatory element) regulatory element of *B. subtilis* showed it to be functional and responsive to decreases in the concentration of leucine in the cell by a T-box mechanism (Marta, *et al.*, 1996). Further work on *ilv-leu* transcriptional regulation has shown that transcriptional control of the *ilv-leu* operon is complex and is subject to regulation by CodY, CcpA, TnrA and the stringent response in addition to T-box regulation (Tojo, *et al.*, 2005). Therefore, it is possible that there are many levels of regulation involved in *leuS* expression. We can rule out a role for the stringent response in this as it has been shown that the stringent response does not affect production of AARS genes in *B. subtilis* (Eymann, *et al.*, 2002).

Our analysis of the putative *cysES* T-box regulatory element also failed to produce a canonical T-box response to limitation for the cognate amino acid. Our analysis was complicated by a difficulty in achieving consistent growth of the cysteine auxotroph strain 1A79. Limitation for cysteine in the strain NF337.1 containing the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion resulted in a brief induction of *gfp* expression (Figure 5.30). Supplementation of NF337.1 cultures with methionine improved the growth of these cultures while still allowing us to starve the cells for cysteine. The *gfp* expression profile in cysteine limited cultures of strain NF337.1 containing methionine was significantly different to that observed in cultures without methionine (Figure 5.31). Cysteine limitation experiments performed with methionine in the media resulted in expression of the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion during exponential growth which then decreased in stationary phase as the cells became limited for cysteine. Neither of the *gfp* induction profiles of strain NF337.1 limited for cysteine in the presence or absence of methionine is a canonical T-box response to amino acid starvation. It has been shown that overproduction of CysRS represses *cysES* transcription (Gagnon, *et al.*, 1994). This result suggests that the *cysES* T-box regulatory element might be responsive to the level of charged $tRNA^{CYS}$. However, while our results show that expression of the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion is responsive to the level of cysteine in the cell, this response does not appear to occur by a T-box mechanism. A further indication of this was provided by analysis of the effects of mutations at the specifier codon. Mutation of the *cysES* T-box regulatory elements specifier codon to any of the four codons UGG, UGU, UGA and UAA, had no effect on induction of *gfp* expression as a result of cysteine starvation. Interestingly, starvation for tryptophan produced an increase in *gfp* expression of the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion that was also independent of the sequence of the specifier

codon. This prompted us to investigate if transcription initiation was affected by general amino acid starvation. An analysis of the effect of cysteine, tryptophan, lysine, methionine and phenylalanine starvation on the *gltX* promoter showed that promoter activity was increased in response to amino acid limitation in all cases. The profiles and degree of *gfp* activity were similar under the same growth conditions for strains containing either the the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion or the P_{gltX}-*gfp* fusion, indicating that the *cysES* T-box regulatory element has little effect on modulating the level of transcription from the *gltX* promoter into the *cysES* genes.

There has been extensive research on sulfur metabolism in *B. subtilis*. This research has revealed that expression of genes involved in cysteine metabolism are controlled primarily by two proteins, CysK and CymR (Tanous, *et al.*, 2008). These two proteins form a complex and bind a loosely conserved consensus sequence found in the promoter region of a number of genes involved in cysteine metabolism to repress expression. The *cysE* gene product, OAS, plays an important role in this regulation. OAS inhibits the formation of the CysK-CymR complex and as a result stimulates expression of genes involved in cysteine biosynthesis. OAS also plays a direct role in cysteine formation as it is reduced with sulfide by the O acetyl-serine thiol lyase activity of CysK to produce cysteine (Even, *et al.*, 2006; Tanous, *et al.*, 2008). All proposed models to explain the cysteine metabolism regulation assume that expression of *cysE* is induced in response to cysteine limitation by a T-box mechanism (Albanesi, *et al.*, 2005; Even, *et al.*, 2006; Tanous, *et al.*, 2008). Our data shows that induction of *cysE* (and *cysS*) is indeed responsive to cysteine limitation. However, this response does not appear to occur by a T-box mechanism. The change in expression of the *cysES* genes as a result of the presence of methionine in the medium is also interesting. The presence of methionine resulted in increased expression of the *cysES* genes, but only in limiting cysteine conditions. This observation is logical as OAS is required for the conversion of methionine to cysteine by the AdoMet pathway (Hullo, *et al.*, 2008). However, under our experimental conditions, at the point of cysteine limitation in cultures containing methionine, expression of the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion is reduced. The reasons for this are unclear. A reduction in expression of *cysES* would indicate that the cell has sufficient cysteine. However, this reduction in expression is coincident with cessation of growth brought about by experimental conditions designed to produce cysteine limitation. It is possible that the experimental conditions used produced some other response in the cell that caused the cessation of growth and that the conversion of methionine to cysteine in the cell resulted in a reduction of expression of the $P_{gltX\text{-T-box-}cysES}$-*gfp* fusion. The exact nature of the regulation of this process remains to be elucidated.

Further analysis of this pathway may help to explain the effects of methionine on *cysES* expression.

The putative T-box regulatory elements of *leuS* and *cysES* in *B. subtilis* are not the first predicted transcription antitermination elements found to be non-functional. Mansilla *et al* (2000) have shown that the S-box regulatory element upstream of the *cysH* gene that is involved in cysteine biosynthesis in *B. subtilis* is non-functional and that regulation in this system occurs at the level of transcription initiation. Transcription of *cysH* is stimulated by OAS and repressed by cysteine, though this repression can only occur in the absence of OAS (Mansilla, *et al.*, 2000). A non-functional T-box regulatory element has also been described in *Clostridium acetobutylicum*. The T-box regulatory element upstream of the *ubiG* operon involved in methionine to cysteine conversion has been shown not to be important for the regulation of expression of genes in this operon. Analysis of the sequence of this regulatory element predicted that reduced $tRNA^{CYS}$ charging should result in increased levels of expression. However, experimental analysis of this system has shown that the T-box regulatory element does not play a major role in the regulation of transcription and that antisense transcription is the primary determinant of regulation of *ubiG* gene expression (Andre, *et al.*, 2008).

Thus it has been shown that in the case of genes related to sulfur metabolism there are multiple regulatory layers on control of expression and that the identification of a recognized regulatory element does not necessarily mean downstream genes will be subject to its control.

6.9 Expression of a number of AARS encoding genes in *B. subtilis* responds to limitation for non-cognate amino acids

It has been shown previously in work carried out in *E. coli* that expression of AARS genes is induced upon starvation for their cognate amino acids. This induction of expression of a given synthetase gene is specific to limitation for its cognate amino acid and not to general amino acid starvation (Nass & Neidhardt, 1967). Results from our analyses of T-box elements in *B. subtilis* confirms the specificity of the T-box regulatory element. Each T-box element responds to their cognate tRNA only, despite high degree of homology between tRNAs that are charged with amino acids from the same mixed codon box, at the anticodon and in some cases other regions of the tRNA. However, results from our analyses of T-box regulated AARS genes in *B. subtilis* shows that in the case of the *leuS, pheS, ileS* and *cysES* genes, induction of expression occurs in response to limitation for non-cognate amino acids. This induction of expression is not due to general amino acid

starvation but appears to respond to limitation for particular amino acids (Table 6.3). For example, in the case of *leuS*, expression of the $P_{T\text{-}box\ leuS}$-*gfp* fusion was induced by limitation for threonine, but not limitation for methionine, isoleucine, phenylalanine or leucine. The mechanism by which this occurs is unclear. As mentioned earlier, expression of AARS genes is unaffected by the stringent response so a *relA* dependant mechanism can be ruled out (Eymann, *et al.*, 2002). An analysis of the biosynthetic pathways of leucine, phenylalanine and threonine shows that they have no intermediates in common, thus an indirect limitation for the cognate amino acid can also be excluded as a possibility.

Induction of expression of *ileS* and *cysES* was also observed to occur in response to limitation for a non-cognate amino acid. In the case of *ileS*, limitation for lysine induces increased expression. For *cysES* expression was induced by phenylalanine, tryptophan, methionine and lysine starvation. Similarly to the examples of *leuS* and *pheS*, the non-cognate amino acids involved in expression of *ileS* and *cysES* have no role in the biosynthesis of isoleucine or cysteine.

The induction of expression observed in the above examples was shown to occur independently of a T-box mechanism in the case of the *leuS, ileS* and *cysES* genes. This indicates that the increase in expression may be due to increased initiation of transcription at the promoter. This was shown to be the case for the *cysES* genes whose promoter activity (P_{glxX}) was increased by limitation for all amino acids tested. Why an increase in expression of a given AARS gene in response to limitation for a non-cognate amino acid might occur, is an interesting question. An increase in expression of an AARS gene by limitation for a non-cognate amino acid would appear to be a sub-optimal response in the sense that metabolic energy is devoted to the synthesis of a protein which may not alleviate the stress which induces its production. As amino acid limitation causes significant stress to the cell, it is likely that the increase in expression of the AARS genes mentioned induces a number of indirect pleiotropic effects which result in upregulation of certain AARSs. It is also possible that efficient regulation of the expression of certain AARS genes does not require a response to limitation for their cognate amino acid. An example of this would be the regulation of *cysS* in a number of *Clostridium* species. In *C. acetobutylicum, C. perfringens, C. difficile* and *C. tetani, cysS* and *proS* are coregulated by a T-box mechanism that appears to be responsive only reduced charging of $tRNA^{Pro}$ (Gutierrez-Preciado, *et al.*, 2009). This has not been shown experimentally but may be an indication that in some cases, regulation of expression of an AARS gene need not be sensitive to limitation for is cognate amino acid.

Table 6.3 Table showing the T-box regulatory elements analyzed in this study and their response to limitation by different amino acids. Lowercase y indicates an induction less than that observed with the cognate amino acid

Species	**T-box transcriptional fusion**	**Specifier codon**	**Induction by amino acid limitation**				
			Lysine	**Asparagine[1]**		**Phenylalanine**	
B. cereus	*lysK-lacZ*	wt	Y	Y		N	
		AAG	y	y			
		AAC	N	N			
		AAU	N	N			
			Lysine	**Asparagine[1]**		**Tryptophan**	
B. cereus	*asnS-lacZ*	wt	N	N		N	
			Histidine	**Glutamine[2]**		**Aspartate[3]**	
B. subtilis	*hisS-aspS-gfp*	wt	Y	N		N	
			Phenylalanine	**Leucine**	**Threonine**	**Methionine**	**Isoleucine**
B. subtilis	*pheS-gfp*	wt	Y	N	y		
		UAA	Y	N			
B. subtilis	*leuS-gfp*	wt	N	N	Y	N	N
		UAA		N			
			Isoleucine	**Methionine**	**Lysine**		
B. subtilis	*ileS-gfp*	wt	Y	N	Y		
		UAA	N	N	Y		
			Tryptophan	**Cysteine**	**Methionine**	**Phenylalanine**	**Lysine**
B. subtilis	*trpS-gfp*	wt	Y	N	N		
B. subtilis	*gltX-cysES-gfp*	wt	Y	Y			
		UGG	Y	Y			
		UAA	Y	Y			
		UGA	Y	Y			
		UGU	Y	Y			
	gltX-gfp		Y	Y	Y	Y	Y

1 IPTG inducible *asnS* strain
2 IPTG inducible *gatCAB* strain
3 IPTG inducible *aspS* strain

A clear understanding of the reasons behind the increase in expression of some AARS genes in response to certain non-cognate amino acids requires further experimental analysis.

6.10 Summary

We set out to investigate the function of the *B. cereus lysK* T-box regulatory element with a view to shedding light on the reason for the relative paucity of T–box regulation for *lysK/S* genes. Our results indicate that the *lysK* T–box regulatory element is functional and responsive to the level of charging of $tRNA^{LYS}$. The data also shows that a *B. subtilis* strain containing either a T-box regulated *lysK* or T-box regulated *lysS* gene is viable. The observation that $tRNA^{ASN}$ can induce antitermination in the *lysK* T-box regulatory element of *B. cereus* may indicate that a lack of specificity is a factor in the relative lack of T-box regulation for *lysS/K* genes. This led us to investigate a number of T-box regulatory elements in *B. subtilis* in order to assess whether they could efficiently distinguish between tRNAs that are identical at two out of three positions at the anticodon. These analyses confirmed the specificity of the T-box mechanism and indicates that the lack of specificity in tRNA recognition seen with the *B. cereus lysK* T-box regulatory element is a special case. Our analyses of T-box regulatory elements in *B. subtilis* revealed that the *leuS* and *cysES* T-box regulatory elements do not function by a canonical T-box mechanism under our experimantal conditions. Furthermore these analyses show that in some cases transcription of tRNA synthetases in *B. subtilis* can respond to starvation for more than one amino acid. Additionally there is a suggestion that while some of these AARS genes may have been T-box regulated in the past (*leuS* and *cysES*), the role of the T-box mechanism in their transcriptional control may have been superceded by other regulation pathways, for example those involved in global regulation of sulfur metabolism in the case of the *cysES* genes.

6.11 Future prospets

The interaction of the anticodon and specifier codon in T-box regulatory elements is critical for specificity of regulation of genes controlled by T-box mechanisms, with an exception in *lysK* of *B. cereus* which appears to be a special case. Codon-anticodon interactions have been studied extensively, particularly their role in translation within the ribosome. An important feature emerging from these analyses is the role of tRNA

modification at the anticodon in codon recognition and discrimination. Most amino acids are coded for by multiple codons, as many as six in the case of leucine and serine. This degeneracy of the code requires that the third "wobble" base of the tRNA anticodon can form non Watson-Crick base-pairing interactions with a number of different nucleotides so tRNA can recognize multiple codons and translation can occur efficiently. Modification of bases in tRNA's and especially of bases within the anticodon loop play a crucial role in the specificity of codon-anticodon interactions. For example the xmo^{5}U modification of uridine allows it to bind A and wobble to G and U. In some cases these modifications restrict wobble as in the case of the lysidine modification of cytosine which restricts the basepairing of the C to A instead of G (Agris, 2004; Agris, *et al.*, 2007). These tRNA modifications function in many cases to alter the structure of the anticodon to allow it to interact with both mRNA codons and ribosomal RNA and proteins (Agris, 2004). These interactions within the ribosome are essential to restrict entry of tRNAs to those tRNAs that have both the correct anticodon and the correct charged amino acid into the acceptor site. Codon-anticodon interactions outside the ribosome are not subject to these restrictions, potentially allowing greater wobble of nucleotides at the anticodon.

We would like to analyze the contribution of these tRNA modifications at the anticodon to the specificity of anticodon-specifier codon interactions within T-box regulatory elements. An example of such an analysis would be the effect of removal of the queuosine modification of the G nucleotide at the tRNA^{ASN} anticodon on the ability of tRNA^{ASN} to interact with the *lysK* T-box regulatory element. An examination of the effect of removal of tRNA modifications of the anticodon of tRNA^{LYS} on its ability to induce antitermination of the *lysK* T-box regulatory element would also be interesting. Analyses of this sort may indicate a compromise between regulation of gene expression and translational fidelity.

In conclusion, the T-box regulatory element is highly specific to reduced levels of charging of its cognate tRNA. The T-box regulatory element of *lysK* in *B. cereus* is an exception to this rule as it displays a lack of specificity in tRNA recognition and is induced by reduced charging of tRNA^{ASN} and tRNA^{LYS}. Our work also demonstrates the importance of experimental verification of putative T-box regulatory elements prior to definitively ascribing a role for these in regulation.

References

Agris, P. F. 2004. Decoding the genome: a modified view. Nucleic Acids Res **32:**223-38.

Agris, P. F., F. A. Vendeix, and W. D. Graham. 2007. tRNA's wobble decoding of the genome: 40 years of modification. J Mol Biol **366:**1-13.

Albanesi, D., M. C. Mansilla, G. E. Schujman, and D. de Mendoza. 2005. Bacillus subtilis cysteine synthetase is a global regulator of the expression of genes involved in sulfur assimilation. J Bacteriol 187:7631-8.

Alexandrov, A., I. Chernyakov, W. Gu, S. L. Hiley, T. R. Hughes, E. J. Grayhack, and E. M. Phizicky. 2006. Rapid tRNA decay can result from lack of nonessential modifications. Mol Cell **21:**87-96.

Anagnostopoulos, C., and J. Spizizen. 1961. Requirements for Transformation in Bacillus Subtilis. J Bacteriol **81:**741-6.

Andre, G., S. Even, H. Putzer, P. Burguiere, C. Croux, A. Danchin, I. Martin-Verstraete, and O. Soutourina. 2008. S-box and T-box riboswitches and antisense RNA control a sulfur metabolic operon of Clostridium acetobutylicum. Nucleic Acids Res **36:**5955-69.

Andrulis, I. L., S. Evans-Blackler, and L. Siminovitch. 1985. Characterization of single step albizziin-resistant Chinese hamster ovary cell lines with elevated levels of asparagine synthetase activity. J Biol Chem **260:**7523-7.

Antoniewski, C., B. Savelli, and P. Stragier. 1990. The spoIIJ gene, which regulates early developmental steps in Bacillus subtilis, belongs to a class of environmentally responsive genes. J Bacteriol **172:**86-93.

Aslanidis, C., and P. J. d. Jong. 1990. Ligation-independant cloning of PCR products (LIC-PCR). Nucleic Acids Research **18:**6069-6075.

Ataide, S. F., B. C. Jester, K. M. Devine, and M. Ibba. 2005. Stationary-phase expression and aminoacylation of a transfer-RNA-like small RNA. EMBO Rep **6:**742-7.

Baldwin, A., and P. Berg. 1966. Transfer Ribonucleic Acid inducedHydrolysis of Valyladenylate bound to Isoleucyl Ribonucleic Acid Synthetase. The Journal of Biological Chemistry **241:**839-845.

Blouin, S., J. Mulhbacher, J. C. Penedo, and D. A. Lafontaine. 2009. Riboswitches: ancient and promising genetic regulators. Chembiochem **10:**400-16.

Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem **72:**248-54.

Collins, M. D., P. A. Lawson, A. Willems, J. J. Cordoba, J. Fernandez-Garayzabal, P. Garcia, J. Cai, H. Hippe, and J. A. Farrow. 1994. The phylogeny of the genus Clostridium: proposal of five new genera and eleven new species combinations. Int J Syst Bacteriol **44:**812-26.

Condon, C., M. Grunberg-Manago, and H. Putzer. 1996a. Aminoacyl-tRNA synthetase gene regulation in Bacillus subtilis. Biochimie **78:**381-9.

Condon, C., H. Putzer, and M. Grunberg-Manago. 1996b. Processing of the leader mRNA plays a major role in the induction of thrS expression following threonine starvation in Bacillus subtilis. Proc Natl Acad Sci U S A **93:**6992-7.

Condon, C., H. Putzer, D. Luo, and M. Grunberg-Manago. 1997. Processing of the Bacillus subtilis thrS leader mRNA is RNase E-dependent in Escherichia coli. J Mol Biol **268:**235-42.

Crick, F. H. 1958. On protein synthesis. Symp Soc Exp Biol **12:**138-63.

Elf, J., D. Nilsson, T. Tenson, and M. Ehrenberg. 2003. Selective charging of tRNA isoacceptors explains patterns of codon usage. Science **300:**1718-22.

Eriani, G., M. Delarue, O. Poch, J. Gangloff, and D. Moras. 1990. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. Nature **347:**203-6.

Even, S., P. Burguiere, S. Auger, O. Soutourina, A. Danchin, and I. Martin-Verstraete. 2006. Global control of cysteine metabolism by CymR in Bacillus subtilis. J Bacteriol **188:**2184-97.

Even, S., O. Pellegrini, L. Zig, V. Labas, J. Vinh, D. Brechemmier-Baey, and H. Putzer. 2005. Ribonucleases J1 and J2: two novel endoribonucleases in B.subtilis with functional homology to E.coli RNase E. Nucleic Acids Res **33:**2141-52.

Eymann, C., G. Homuth, C. Scharf, and M. Hecker. 2002. Bacillus subtilis functional genomics: global characterization of the stringent response by proteome and transcriptome analysis. J Bacteriol **184:**2500-20.

Fauzi, H., K. D. Jack, and J. V. Hines. 2005. In vitro selection to identify determinants in tRNA for Bacillus subtilis tyrS T box antiterminator mRNA binding. Nucleic Acids Res **33:**2595-602.

Frugier, M., M. Ryckelynck, and R. Giege. 2005. tRNA-balanced expression of a eukaryal aminoacyl-tRNA synthetase by an mRNA-mediated pathway. EMBO Rep **6:**860-5.

Gagnon, Y., R. Breton, H. Putzer, M. Pelchat, M. Grunberg-Manago, and J. Lapointe. 1994. Clustering and co-transcription of the Bacillus subtilis genes encoding the aminoacyl-tRNA synthetases specific for glutamate and for cysteine and the first enzyme for cysteine biosynthesis. J Biol Chem **269:**7473-82.

Garst, A. D., A. Heroux, R. P. Rambo, and R. T. Batey. 2008. Crystal structure of the lysine riboswitch regulatory mRNA element. J Biol Chem **283:**22347-51.

Gerdeman, M. S., T. M. Henkin, and J. V. Hines. 2002. In vitro structure-function studies of the Bacillus subtilis tyrS mRNA antiterminator: evidence for factor-

independent tRNA acceptor stem binding specificity. Nucleic Acids Res **30:**1065-72.

Gerdeman, M. S., T. M. Henkin, and J. V. Hines. 2003. Solution structure of the Bacillus subtilis T-box antiterminator RNA: seven nucleotide bulge characterized by stacking and flexibility. J Mol Biol **326:**189-201.

Gibson, J., R. Poole, M. Hughes, and J. Rees. 1984. Filamentous growth of Escherichia coli K12 elicited by dimeric, mixed-valence complexes of ruthenium. Arch Microbiol **139:**265-271.

Glaser, P., F. Kunst, M. Arnaud, M. P. Coudart, W. Gonzales, M. F. Hullo, M. Ionescu, B. Lubochinsky, L. Marcelino, I. Moszer, and et al. 1993. Bacillus subtilis genome project: cloning and sequencing of the 97 kb region from 325 degrees to 333 degrees. Mol Microbiol **10:**371-84.

Graffe, M., J. Dondon, J. Caillet, P. Romby, C. Ehresmann, B. Ehresmann, and M. Springer. 1992. The specificity of translational control switched with transfer RNA identity rules. Science **255:**994-6.

Grandoni, J. A., S. A. Zahler, and J. M. Calvo. 1992. Transcriptional regulation of the ilv-leu operon of Bacillus subtilis. J Bacteriol **174:**3212-9.

Grundy, F. J., S. C. Lehman, and T. M. Henkin. 2003. The L box regulon: lysine sensing by leader RNAs of bacterial lysine biosynthesis genes. Proc Natl Acad Sci U S A **100:**12057-62.

Grundy, F. J., J. A. Collins, S. M. Rollins, and T. M. Henkin. 2000. tRNA determinants for transcription antitermination of the Bacillus subtilis tyrS gene. Rna **6:**1131-41.

Grundy, F. J., and T. M. Henkin. 1998. The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in gram-positive bacteria. Mol Microbiol **30:**737-49.

Grundy, F. J., M. T. Haldeman, G. M. Hornblow, J. M. Ward, A. F. Chalker, and T. M. Henkin. 1997. The Staphylococcus aureus ileS gene, encoding isoleucyl-tRNA synthetase, is a member of the T-box family. J Bacteriol **179:**3767-72.

Grundy, F. J., and T. M. Henkin. 1992. Characterization of the Bacillus subtilis rpsD regulatory target site. J Bacteriol **174:**6763-70.

Grundy, F. J., and T. M. Henkin. 1994. Conservation of a transcription antitermination mechanism in aminoacyl-tRNA synthetase and amino acid biosynthesis genes in gram-positive bacteria. J Mol Biol **235:**798-804.

Grundy, F. J., and T. M. Henkin. 2004. Kinetic analysis of tRNA-directed transcription antitermination of the Bacillus subtilis glyQS gene in vitro. J Bacteriol **186:**5392-9.

Grundy, F. J., and T. M. Henkin. 2004. Regulation of gene expression by effectors that bind to RNA. Curr Opin Microbiol **7:**126-31.

Grundy, F. J., and T. M. Henkin. 1993. tRNA as a positive regulator of transcription antitermination in B. subtilis. Cell **74:**475-82.

Grundy, F. J., S. E. Hodil, S. M. Rollins, and T. M. Henkin. 1997. Specificity of tRNA-mRNA interactions in Bacillus subtilis tyrS antitermination. J Bacteriol **179:**2587-94.

Grundy, F. J., T. R. Moir, M. T. Haldeman, and T. M. Henkin. 2002. Sequence requirements for terminators and antiterminators in the T box transcription antitermination system: disparity between conservation and functional requirements. Nucleic Acids Res **30:**1646-55.

Grundy, F. J., S. M. Rollins, and T. M. Henkin. 1994. Interaction between the acceptor end of tRNA and the T box stimulates antitermination in the Bacillus subtilis tyrS gene: a new role for the discriminator base. J Bacteriol **176:**4518-26.

Grundy, F. J., W. C. Winkler, and T. M. Henkin. 2002. tRNA-mediated transcription antitermination in vitro: codon-anticodon pairing independent of the ribosome. Proc Natl Acad Sci U S A **99:**11121-6.

Grundy, F. J., M. R. Yousef, and T. M. Henkin. 2005. Monitoring uncharged tRNA during transcription of the Bacillus subtilis glyQS gene. J Mol Biol **346:**73-81.

Guerout-Fleury, A., K. Shazand, N. Frandsen, and P. Stragier. 1995. Antibiotic-resistance cassettes for Bacillus subtilis. Gene **180:**335-336.

Gutierrez-Preciado, A., T. M. Henkin, F. J. Grundy, C. Yanofsky, and E. Merino. 2009. Biochemical features and functional implications of the RNA-based T-box regulatory mechanism. Microbiol Mol Biol Rev **73:**36-61.

Harwood, C., and S. Cutting. 1990. Molecular Biological Methods for Bacillus. Wiley, NY.

Henkin, T. M. 1994. tRNA-directed transcription antitermination. Mol Microbiol **13:**381-7.

Henkin, T. M., B. L. Glass, and F. J. Grundy. 1992. Analysis of the Bacillus subtilis tyrS gene: conservation of a regulatory sequence in multiple tRNA synthetase genes. J Bacteriol **174:**1299-306.

Hoagland, M. 1955. An enzymic mechanism for amino acid activation in animal tissues. Biochim Biophys Acta **16:**288-9.

Hoagland, M. B., M. L. Stephenson, J. F. Scott, L. I. Hecht, and P. C. Zamecnik. 1958. A soluble ribonucleic acid intermediate in protein synthesis. J Biol Chem **231:**241-57.

Hornstra, L. M., Y. P. de Vries, W. M. de Vos, T. Abee, and M. H. Wells-Bennik. 2005. gerR, a novel ger operon involved in L-alanine- and inosine-initiated germination of Bacillus cereus ATCC 14579. Appl Environ Microbiol **71:**774-81.

Hullo, M., S. Auger, O. Soutourina, O. Barzu, M. Yvon, A. Danchin, and I. Martin-Verstraete. 2007. Conversion of methionine to cysteine in Bacillus subtilis and its regulation. J Bacteriol **189:**187-197.

Ibba, M., H. D. Becker, C. Stathopoulos, D. L. Tumbula, and D. Soll. 2000. The adaptor hypothesis revisited. Trends Biochem Sci **25:**311-6.

Ibba, M., and D. Soll. 2000. Aminoacyl-tRNA synthesis. Annu Rev Biochem **69:**617-50.

Ibba, M., and D. Soll. 2001. The renaissance of aminoacyl-tRNA synthesis. EMBO Rep **2:**382-7.

Jack, K., M. JA, and J. Hines. 2008. Characterizing riboswitch function: identification of Mg2+ binding site in T box antiterminator RNA. Biochem Biophys Res Commun **370:**306-310.

Jakubowski, H., and A. R. Fersht. 1981. Alternative pathways for editing non-cognate amino acids by aminoacyl-tRNA synthetases. Nucleic Acids Res **9:**3105-17.

Jakubowski, H., and E. Goldman. 1992. Editing of errors in selection of amino acids for protein synthesis. Microbiol Rev **56:**412-29.

Jester, B. C., J. D. Levengood, H. Roy, M. Ibba, and K. M. Devine. 2003. Nonorthologous replacement of lysyl-tRNA synthetase prevents addition of lysine analogues to the genetic code. Proc Natl Acad Sci U S A **100:**14351-6.

Lawrence, J. S., and W. W. Ford. 1916. Spore-bearing Bacteria in Milk. J Bacteriol **1:**277-320 51.

Lestienne, P., J. A. Plumbridge, M. Grunberg-Manago, and S. Blanquet. 1984. Autogenous repression of Escherichia coli threonyl-tRNA synthetase expression in vitro. J Biol Chem **259:**5232-7.

Levengood, J., S. F. Ataide, H. Roy, and M. Ibba. 2004. Divergence in noncognate amino acid recognition between class I and class II lysyl-tRNA synthetases. J Biol Chem **279:**17707-14.

Loftfield, R. B. 1963. The Frequency of Errors in Protein Biosynthesis. Biochem J **89:**82-92.

Luo, D., C. Condon, M. Grunberg-Manago, and H. Putzer. 1998. In vitro and in vivo secondary structure probing of the thrS leader in Bacillus subtilis. Nucleic Acids Research **26:**5379-5387.

Mansilla, M. C., D. Albanesi, and D. de Mendoza. 2000. Transcriptional control of the sulfur-regulated cysH operon, containing genes involved in L-cysteine biosynthesis in Bacillus subtilis. J Bacteriol **182:**5885-92.

Marta, P. T., R. D. Ladner, and J. A. Grandoni. 1996. A CUC triplet confers leucine-dependent regulation of the Bacillus subtilis ilv-leu operon. J Bacteriol **178:**2150-3.

Mayaux, J. F., G. Fayat, M. Panvert, M. Springer, M. Grunberg-Manago, and S. Blanquet. 1985. Control of phenylalanyl-tRNA synthetase genetic expression. Site-directed mutagenesis of the pheS, T operon regulatory region in vitro. J Mol Biol **184:**31-44.

Miller, J. 1972. Experiments in molecular genetics. Cold Spring Harbor Press, Cold Spring Harbor, NY.

Moine, H., P. Romby, M. Springer, M. Grunberg-Manago, J. P. Ebel, B. Ehresmann, and C. Ehresmann. 1990. Escherichia coli threonyl-tRNA synthetase and tRNA(Thr) modulate the binding of the ribosome to the translational initiation site of the thrS mRNA. J Mol Biol **216:**299-310.

Moine, H., P. Romby, M. Springer, M. Grunberg-Manago, J. P. Ebel, C. Ehresmann, and B. Ehresmann. 1988. Messenger RNA structure and gene regulation at the translational level in Escherichia coli: the case of threonine:tRNAThr ligase. Proc Natl Acad Sci U S A **85:**7892-6.

Mulhbacher, J., and D. A. Lafontaine. 2007. Ligand recognition determinants of guanine riboswitches. Nucleic Acids Res **35:**5568-80.

Nass, G., and F. Neidhardt. 1967. Regulation of formation of aminoacyl ribonucleic acid synthetases in Escherichia coli. Biochim Biophys Acta **134:**347-359.

Nelson, A. R., T. M. Henkin, and P. F. Agris. 2006. tRNA regulation of gene expression: interactions of an mRNA 5'-UTR with a regulatory tRNA. Rna **12:**1254-61.

O'Donoghue, P., and Z. Luthey-Schulten. 2003. On the evolution of structure in aminoacyl-tRNA synthetases. Microbiol Mol Biol Rev **67:**550-73.

O'Donoghue, P., A. Sethi, C. R. Woese, and Z. A. Luthey-Schulten. 2005. The evolutionary history of Cys-tRNACys formation. Proc Natl Acad Sci U S A **102:**19003-8.

Pelchat, M., and J. Lapointe. 1999a. Aminoacyl-tRNA synthetase genes of Bacillus subtilis: organization and regulation. Biochem Cell Biol **77:**343-7.

Pelchat, M., and J. Lapointe. 1999b. In vivo and in vitro processing of the Bacillus subtilis transcript coding for glutamyl-tRNA synthetase, serine acetyltransferase, and cysteinyl-tRNA synthetase. Rna **5:**281-9.

Petit, M., E. Dervyn, M. Rose, K. Entian, S. McGovern, S. Ehrlich, and C. Bruand. 1998. PcrA is an essential DNA helicase of Bacillus subtilis fulfilling functions both in repair and rolling-circle replication. Mol Microbiol **29:**261-273.

Putney, S., and P. Schimmel. 1981. An aminoacyl tRNA synthetase binds to a specific DNA sequence and regulates its gene transcription. Nature **291:**1497-1501.

Putzer, H., C. Condon, D. Brechemier-Baey, R. Brito, and M. Grunberg-Manago. 2002. Transfer RNA-mediated antitermination in vitro. Nucleic Acids Res **30:**3026-33.

Putzer, H., N. Gendron, and M. Grunberg-Manago. 1992. Co-ordinate expression of the two threonyl-tRNA synthetase genes in Bacillus subtilis: control by transcriptional antitermination involving a conserved regulatory sequence. Embo J **11:**3117-27.

Putzer, H., S. Laalami, A. A. Brakhage, C. Condon, and M. Grunberg-Manago. 1995. Aminoacyl-tRNA synthetase gene regulation in Bacillus subtilis: induction, repression and growth-rate regulation. Mol Microbiol **16:**709-18.

Reader, J. S., D. Metzgar, P. Schimmel, and V. de Crecy-Lagard. 2004. Identification of four genes necessary for biosynthesis of the modified nucleoside queuosine. J Biol Chem **279:**6280-5.

Rollins, S. M., F. J. Grundy, and T. M. Henkin. 1997. Analysis of cis-acting sequence and structural elements required for antitermination of the Bacillus subtilis tyrS gene. Mol Microbiol **25:**411-21.

Romby, P., C. Brunel, J. Caillet, M. Springer, M. Grunberg-Manago, E. Westhof, C. Ehresmann, and B. Ehresmann. 1992. Molecular mimicry in translational control of E. coli threonyl-tRNA synthetase gene. Competitive inhibition in tRNA aminoacylation and operator-repressor recognition switch using tRNA identity rules. Nucleic Acids Res **20:**5633-40.

Romby, P., J. Caillet, C. Ebel, C. Sacerdot, M. Graffe, F. Eyermann, C. Brunel, H. Moine, C. Ehresmann, B. Ehresmann, and M. Springer. 1996. The expression of E.coli threonyl-tRNA synthetase is regulated at the translational level by symmetrical operator-repressor interactions. Embo J **15:**5976-87.

Ryckelynck, M., R. Giege, and M. Frugier. 2005. tRNAs and tRNA mimics as cornerstones of aminoacyl-tRNA synthetase regulations. Biochimie **87:**835-45.

Salazar, J., I. Ahel, O. Orellana, D. Tumbula-Hansen, R. Krieger, L. Daniels, and D. Soll. 2003. Coevolution of an aminoacyl-tRNA synthetase with its tRNA substrates. Proc Natl Acad Sci USA **100:**13863-8.

Sambrook, J., E. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Lab. Press, Plainview, NY.

Schimmel, P. 2008. Development of tRNA synthetases and connection to genetic code and disease. Protein Sci **17:**1643-52.

Shaul, S., R. Nussinov, and T. Pupko. 2006. Paths of lateral gene transfer of lysyl-aminoacyl-tRNA synthetases with a unique evolutionary transition stage of prokaryotes coding for class I and II varieties by the same organisms. BMC Evol Biol **6:**22.

Springer, M., J. F. Mayaux, G. Fayat, J. A. Plumbridge, M. Graffe, S. Blanquet, and M. Grunberg-Manago. 1985. Attenuation control of the Escherichia coli phenylalanyl-tRNA synthetase operon. J Mol Biol **181:**467-78.

Springer, M., J. A. Plumbridge, J. S. Butler, M. Graffe, J. Dondon, J. F. Mayaux, G. Fayat, P. Lestienne, S. Blanquet, and M. Grunberg-Manago. 1985. Autogenous control of Escherichia coli threonyl-tRNA synthetase expression in vivo. J Mol Biol **185:**93-104.

Steinberg, W. 1974. Temperature-induced derepression of tryptophan biosynthesis in a tryptophanyl-transfer ribonucleic acid synthetase mutant of Bacillus subtilis. J Bacteriol **117:**1023-34.

Stulke, J., R. Hanschke, and M. Hecker. 1993. Temporal activation of beta-glucanase synthesis in Bacillus subtilis is mediated by the GTP pool. J Gen Microbiol **139:**2041-5.

Tanous, C., O. Soutourina, B. Raynal, M. F. Hullo, P. Mervelet, A. M. Gilles, P. Noirot, A. Danchin, P. England, and I. Martin-Verstraete. 2008. The CymR regulator in complex with the enzyme CysK controls cysteine metabolism in Bacillus subtilis. J Biol Chem **283:**35551-60.

Tojo, S., T. Satomura, K. Morisaki, J. Deutscher, K. Hirooka, and Y. Fujita. 2005. Elaborate transcription regulation of the Bacillus subtilis ilv-leu operon involved in the biosynthesis of branched-chain amino acids through global regulators of CcpA, CodY and TnrA. Mol Microbiol **56:**1560-73.

Vagner, V., E. Dervyn, and S. Ehrlich. 1998. A vector for systematic gene inactivation in Bacillus subtilis. Microbiology **144:**3097-3104.

van de Guchte, M., S. Ehrlich, and A. Chopin. 2001. Identity elements in tRNA-mediated transcription antitermination: implication of tRNA D- and T-arms in mRNA recognition. Microbiology **147:**1223-1233.

Vander Horn, P. B., and S. A. Zahler. 1992. Cloning and nucleotide sequence of the leucyl-tRNA synthetase gene of Bacillus subtilis. J Bacteriol **174:**3928-35.

Wang, S., M. Praetorius-Ibba, S. Ataide, H. Roy, and M. Ibba. 2006. Discrimination of cognate and non-cognate substrates at the active site of class I lysyl-tRNA synthetase. Biochemistry **45:**3646-3652.

Wels, M., T. Groot Kormelink, M. Kleerebezem, R. J. Siezen, and C. Francke. 2008. An in silico analysis of T-box regulated genes and T-box evolution in prokaryotes, with emphasis on prediction of substrate specificity of transporters. BMC Genomics **9:**330.

Winkler, W., A. Nahvi, and R. R. Breaker. 2002. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. Nature **419:**952-6.

Winkler, W. C., F. J. Grundy, B. A. Murphy, and T. M. Henkin. 2001. The GA motif: an RNA element common to bacterial antitermination systems, rRNA, and eukaryotic RNAs. Rna **7:**1165-72.

Woese, C. R., G. J. Olsen, M. Ibba, and D. Soll. 2000. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. Microbiol Mol Biol Rev **64:**202-36.

Yousef, M. R., F. J. Grundy, and T. M. Henkin. 2005. Structural transitions induced by the interaction between tRNA(Gly) and the Bacillus subtilis glyQS T box leader RNA. J Mol Biol **349:**273-87.

Yousef, M. R., F. J. Grundy, and T. M. Henkin. 2003. tRNA requirements for glyQS antitermination: a new twist on tRNA. Rna **9:**1148-56.