MOLECULAR ANALYSIS AND REGULATION OF THE *CLOSTRIDIUM ACETOBUTYLICUM* GLUTAMINE SYNTHETASE GENE *GLNA*, CLONED IN *ESCHERICHIA COLI*

Paul J. D. Janssen

B.Sc.(Eng.)

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Faculty of Science, University of Cape Town,

CAPE TOWN

JANUARY, 1990

CERTIFICATION OF SUPERVISOR

In terms of paragraph 9 of "General regulations for the degree of Ph. D." I, as supervisor of the candidate, P. J. D. Janssen, certify that I approve of the incorporation in this thesis of material that has already been published or submitted for publication.

Signed

Professor. D. R. Woods
Deputy Vice-Chancellor (Research)
Director Microbial Research Units
University of Cape Town

"*In nature's infinite book of secrecy A little I can read*"

William Shakespeare
(*In* Antony and Cleopatra)

"*Chance favours the trained mind*"

Louis Pasteur

CONTENTS

ACKNOWLEDGEMENTS

I am indebted to my supervisor, Prof. David Woods, for his encouragement, helpful criticism, and patience. His contagious enthusiasm added considerable impetus to this study.

My gratitude also extends to Prof. David Jones, who was co-supervisor in the first years of this study, not only for his guidance and constant interest, but also for the lengthy discussions we had. Thanks Dave.

Thanks to Dr. Edward Rybicki of the Virology Section, who was very helpful with the phylogenetic studies.

I would also like to thank the people in the Anaerobe Unit who have been extremely kind to me. A special thanks to Joe Santangelo, who has shown me the way in computer software, to Dr. Harold Zappe for the exchange of ideas, to Dr. Heide Goodman, Dr. Vallerie Abratt, Eldie Strydom and Mary-Lou Henderson for providing moral support and 'tolerating' me in the lab, in spite of me being a bit grumpy and tense the last few months.

Many thanks to William Bourn of the Streptomycetes group for meticulous proofreading.

I am indebted to Drs. Johann Botterman and Patrick Stanssens (Plant Genetic Systems, Belgium) for sending me strains, plasmids, and phasmids.

Last but not least, I would like to express my greatest gratitude towards my parents, for their unfailing spiritual support and financial assistance throughout the years of Learning. I also would like to thank them, and my sisters, Marijke and Tinneke, for their love and understanding, and the long letters from home.

I acknowledge the financial assistance which I received from the C.S.I.R., South Africa.

ABBREVATIONS

A	adenine
aa(s)	amino acid(s)
ABE	acetone-butanol-ethanol
Ac	acetate
Ap	ampicillin
ATP	adenosine 5'-triphosphate
bp(s)	base pair(s)
BSA	bovine serum albumine
C	cytosine
C°	Celsius grades
CBM	clostridial basal medium
cDNA	complementair DNA
cfu	colony-forming units
Cm	chloramphenicol
CsCl	cesium chloride
C-terminal	carboxy terminal end of a protein
DEPC	diethylpyrocarbonate
DNA	deoxyribonucleic acis
DNAse	deoxyribonuclease
DTT	1,4-dithio-L-threitol
ds	double stranded
EDTA	ethylenediaminetetra-acetic acid
EtBr	etidium bromide
EtOH	ethanol
g	gram
G	guanine
GDH	glutamate dehydrogenase
GS	glutamine synthetase
GOGAT	glutamine-oxoglutarate amido transferase
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase pairs
Km	kanamycin
Km	Michaelis constant
L	liter
LB	Luria-Bertani (medium)
min	minute
MOPS	3-(N-morpholini) propanesulfonic (acid)
M_r	apparent relative molecular mass
N-terminal	amino terminal end of a protein
nt(s)	nucleotide(s)
OD	optical density
ORF	open reading frame
ori	origin of replication

p	plasmid/ phasmid
PAGE	polyacrylamide gel electrophoresis
PEG	polyethylene glycol
pfu	plaque-forming units
PolIk	Klenow fragment of E. coli DNA polymerase I
R	resistance
RBS	ribosome binding site
rDNA	recombinant DNA
RF	replicative form DNA (of M13 bacteriophage)
RNAse	ribonuclease
rpm	revolution per minute
s	second
S	sensitive
SD	Shine-Dalgarno (sequence)
SDS	sodium dodecyl sulphate
ss	single stranded
SSPE	sodiumchloride-sodiumphosphate-EDTA solution
T	thymine
Tc	tetracycline
TCA	tricarboxylic acid
Tris	Tris(hydroxymethyl)aminomethane
U	units of enzyme activity
U	uracil
UV	ultraviolet (light)
v/v	volume to volume
w/v	weight to volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
::	novel joint
( )	plasmid carrier state
Δ	deletion

ABSTRACT

Clostridium acetobutylicum P262 is an endospore-forming Gram-positive anaerobic bacterium, which has been used in industrial fermentation for the production of acetone and butanol from carbohydrate substrates for many years. This study forms part of a research program towards a better understanding of the genetics and molecular biology of *C. acetobutylicum*, which in the long term will lead to the commercial improvement of the ABE fermentation. Although research over the past few years has been focused on the genes and enzymes involved in the various pathways of solvent production and substrate utilization, it is also of interest to study genes involved in nitrogen metabolism, as it was earlier established that nitrogen levels in *C. acetobutylicum* are important for solventogenesis and sporulation. The main objective of this study was the characterization and molecular analysis of the C. acetobutylicum *glnA* gene.

The nucleotide sequence of the *glnA* coding region, as well as up- and downstream regulatory regions, was determined. The *glnA* gene consisted of an ORF of 1332-bp and the deduced primary aa sequence of 444 residues was compared with the aa sequence of ten other bacterial GS enzymes. The highest sequence similarity was observed with the *Bacillus subtilis* and *Bacillus cereus* GS aa sequences, and a phylogenetic study placed the *C. acetobutylicum* GS in the same group as

the GS of the two *Bacilli* spp., indicating a close evolutionary relationship between the three enzymes.

In *E. coli* and related Gram-negative organisms, regulation of *glnA* gene expression is accomplished by the action of NtrB and NtrC and activation of *glnA* transcription requires a specific sigma factor, σ^{54}. In addition, GS activity is modulated by a complex process involving the adenylylation/ deadenylylation of GS subunits. In *B. subtilis*, no such an elaborate control of GS activity is present and the regulation of *glnA* gene expression at the level of transcription is less complex. The *glnA* gene in the *Bacilli* is preceded by a small gene, *glnR*, whose product is similar to regulatory proteins in that it contains a helix-turn-helix motif, common in DNA binding proteins. Extensive studies have been performed on the involvement of *glnR* in the regulation of the *B. subtilis glnA* gene and it has recently been suggested that the *glnR* gene product binds to the *glnRA* regulatory region to repress transcription.

There is no chemically defined minimal medium in which the industrial *C. acetobutylicum* P262 strain grows reliably and reproducibly, and it is not possible, at present, to study the regulation of the *glnA* gene by nitrogen in *C. acetobutylicum* P262. These studies were therefor carried out with the cloned *glnA* gene in *E. coli* YMC11 *glnALG*.

Expression of the *C. acetobutylicum glnA* gene was regulated by nitrogen in *E. coli* from its own regulatory region and there was no evidence for the presence of *ntr* genes nor was the *C. acetobutylicum* GS adenylylated in *E. coli* cell extracts. Sequence analysis of the *glnA* upstream region revealed no NR_I sites in the *glnA* upstream region and in addition, no upstream regulatory gene could be detected, in spite of the close evolutionary relationship between the GS enzymes of *C. acetobutylicum* and *Bacilli* spp.

The *C. acetobutylicum glnA* gene was studied in *E. coli* by several means, which included *glnA* transcript mapping, construction and analysis of *glnA* upstream deletion mutants, the construction of *glnA::lacZ* fusions, and site-directed mutagenesis. Initiation sites of *glnA* transcription were determined and it was demonstrated that two promoters p_1 and p_2 were utilized in both *E. coli* and *C. acetobutylicum*. Transcription of the *C. acetobutylicum glnA* gene was regulated by nitrogen levels in *E. coli*, and it was shown that upstream sequences, including a complex palindromic sequence overlapping p_1, played an important role in the regulation of *glnA* gene expression.

Deletion of a 3 kb *glnA* downstream region affected GS production levels and caused deregulation of *glnA* transcription. However, a striking finding in this study was that the regulation of the *glnA* gene on the transcriptional level did not correlate with the levels of GS produced. It is proposed that a small antisense RNA, which is transcribed from promoter p_3 located immediately downstream of the *glnA* gene, sequesters a region of the *glnA* mRNA containing the SD and AUG elements and thereby inhibits translation efficiency. The transcription initiation site for the downstream promoter p_3 was determined in *E. coli* and *C. acetobutylicum* cells, and it was shown that p_3 transcription levels were not affected by nitrogen levels, although expression levels of this promoter, fused to an indicator gene, varied with the growthphase.

An attempt was made to abolish promoter activity of the *glnA* downstream promoter p_3 by site directed mutagenesis whereby the -35 region of this promoter was changed from TTGCTA to AAGCTA. Surprisingly, this resulted in an increase of promoter activity for the mutant p_3 promoter and resulted in lower levels of GS production. Although this is in agreement with the hypothetical inhibitory effect of antisense RNA on GS production, it would be of great interest to isolate and characterize a p_3 promoter *down* mutant and to use this mutant in regulatory studies. Suggestions concerning feature investigations are made in Chapter 6.

Chapter 1: General Introduction and Literature Review.

Chapter 1

General Introduction and Literature Review

1.1. The ABE fermentation.

At the beginning of the twentieth century a great demand for organic solvents such as acetone, butanol, and ethanol occured, especially with the advent of World War I. The chemical production of these solvents was not efficient and was at that time expensive . Consequently, a search for microorganisms with the ability to produce organic solvents by fermentation was undertaken. A bacterium, *Bacillus granulobacter pectinovorum*, later named *Clostridium acetobutylicum*, which efficiently produced acetone and butanol from starch, was isolated by Weizmann (reviewed by Gabriel, 1928). The acetone-butanol (AB) fermentation by *C. acetobutylicum* has been widely documented and Rogers (1986), McNeil and Kristiansen (1986), and Jones and Woods (1986a) have comprehensively reviewed the process.

Biological limitations (e.g. butanol toxicity, bacteriophage infections) and various economic factors (e.g. cost of solvent recovery, high substrate cost) have contributed to the decline of the industrial AB fermentation and made it unprofitable in the face of competition from the petrochemical industry. In an era of diminishing oil reserves, however, the production of chemical feedstocks from renewable resources is an important long term goal for biotechnology. Recent research has therefore been focused on the molecular biology of nitrogen metabolism (Usdin *et al.*, 1986; Janssen *et al.*, 1988), solvent pathways (Carry *et al.*, 1988; Youngleson *et al.*, 1988; Youngleson *et al.*, 1989), and substrate utilization (Zappe *et al.*, 1986; Zappe *et al.*, 1987; Zappe *et al.*, 1988; Verhasselt *et al.*, 1989) with a view to the genetic engineering of the industrial strain *C. acetobutylicum* P262. A comprehensive review of the molecular genetics of *Clostridium* strains used in industry has been presented recently by Young *et al.* (1989).

1.1.1. The physiology of the AB fermentation.

A wide variety of polysaccharides are utilized by *C. acetobutylicum* via a complex set of biochemical pathways (Rogers, 1986; Jones and Woods, 1986a). Carbon dioxide, molecular hydrogen, acetic and butyric acid, as well as acetone, butanol and a small amount of ethanol are the major fermentation end-products.

The batch fermentation may be divided into two distinct physiological phases (Gottschal and Morris, 1981a; Jones *et al.*, 1982) (Fig. 1.1.):

- **Acidogenic phase.** The first phase of the fermentation coincides with the active growth period of the culture. Carbohydrates are catabolised to acetate and butyrate and the pH of the medium decreases. The relative amounts of acetate and butyrate present during the early fermentation phase appear to determine the ratio of acetone to butanol and the solvent yield during the final fermentation phase (Martin *et al., 1983)*.

- **Solventogenic phase.** The second phase of the fermentation involves the switch in the carbon flow from the acid-producing pathways to the solvent-producing pathways while there is a further consumption of the substrate. The pH of the medium rises during this phase.

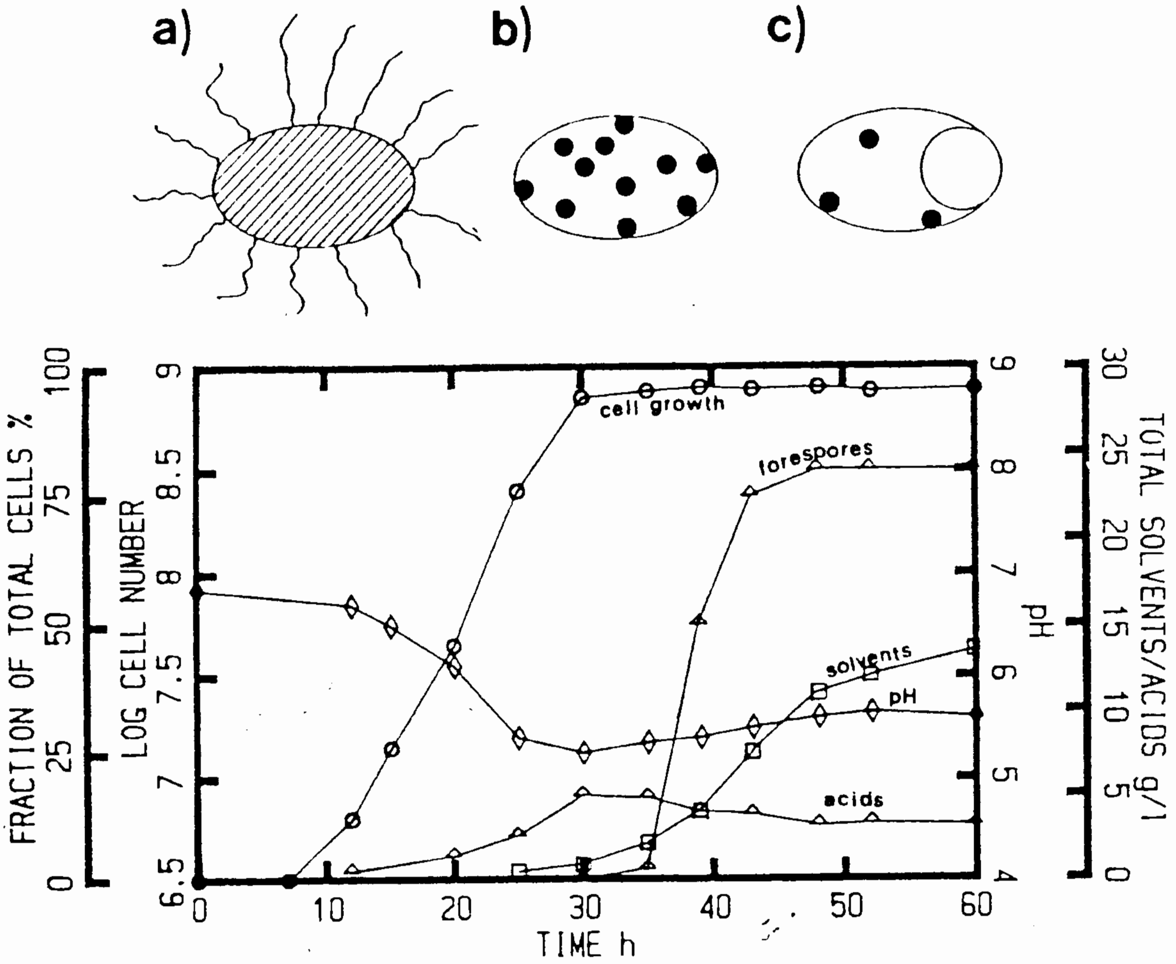


Fig. 1.1.: The different growthstages of C. *acetobutylicum* in a typical batch fermentation (Reysenbach *et al.,* 1986). Log cell number, pH, and percentage of forespore formation are indicated. Morphological changes are schematically represented in the different growth phases: **a.** vegetative cells with high motility, **b.** accumulation of granulose, **c.** consumation of granulose and formation of endospores.

The occurrence of clostridial forms and changes in morphology such as shape, size and motility, associated with the two fermentation phases, are routinely used to monitor the progress of industrial and laboratory fermentations (Jones *et al.*, 1982) (Fig.1.1). Granulose accumulation in C. acetobutylicum occurs only at the end of exponentional growth and is associated with the cessation of cell division, the onset of solvent production, and spore formation (Fig. 1.1; Reysenbach *et al.*, 1986).

C. acetobutylicum ferments sugars via the fructose biphosphate pathway (Embden-Meyerhof pathway) to pyruvate which is then converted by a phosphoroclastic reaction to acetyl-CoA. During the acidogenic phase, some acetyl-CoA is converted to acetate, while most enters a cyclic pathway for the production of butyrate via acetoacetate and butyryl-CoA. During the solventogenic phase, the metabolic flux is diverted and acetyl-CoA and butyryl-CoA function as key intermediates for the production of ethanol and butanol, while part of the acetoacetyl-CoA used to produce butyrate during the acid phase is converted to acetone by decarboxylation of acetoacetate. The commercial ABE fermentation yields an approximate 6:3:1 ratio of butanol, acetone and ethanol respectively (Spivey, 1978).

1.1.2. The onset of solventogenesis.

Enzymes required for the production of acetone, butanol, and ethanol are induced in *C. acetobutylicum* in batch cultures just after completion of exponentional growth and at the onset of solvent production (Andersch *et al.*, 1983; Rogers, 1986; Terracciano and Kashket, 1986) (fig. 1.1). How this transition from the acidogenic to the solventogenic phase is accomplished is not entirely clear. Apparently, threshold concentrations of glucose, nitrogen and acid end-products, and a narrow pH range are required for the initiation of solventogenesis (Long *et al.*, 1984). However, it has been demonstrated that pH itself is not the trigger (Gottschal and Morris, 1981a; Long *et al.*, 1984). It is more likely that a predetermined threshold pH may be required for metabolic functions such as the uptake of carbohydrates and ammonia (Roos *et al.*, 1984; Bowles and Ellefson, 1985), indicating that ammonia/glucose ratios may affect the induction of solvent production. Gottschal and Morris (1981b) failed to observe solvent production in glucose- or ammonia-limited chemostats. Nutrient limitation appeared to select for a population of cells which were unable to produce solvents and had lost the ability to sporulate under these conditions. However, glucose-limited continuous cultures supplemented with fairly high levels of butyrate (Bahl *et al.*, 1982) or nitrogen-limited continuous cultures (Andersch *et al.*, 1982; Monot and Engasser, 1983) were shown to produce solvents. Similar fermentation kinetics and solvent production occurred in a nitrogen-limited chemostat when the dilution rate, and hence the specific growth rate, was low and the pH maintained at pH 5.0 (Monot and Engasser, 1983). Thus it seems that it is not nutrition limitation *per se*, but rather poor growth conditions that are responsible for the onset of solvent production. In accordance with this, other growth-limiting factors, such as low phosphate levels, short exposures to oxygen, temperature shifts and toxic end-products, appear to induce solvent production (O'Brien and Morris., 1971; Bahl *et al.*, 1982; Herrero, 1983; McNeil and Kristiansen, 1985). It has also been suggested that the onset of solventogenesis is linked to the inhibition of cell division (Long, 1984, Ph.D. Thesis, University of Cape Town, South Africa). The cessation of cell division may be followed by the rapid onset of sporulation, and in batch culture (under appropriate conditions) greater than 90% of the cells may develop septa during solventogenesis.

1.1.3. Induction of sporulation.

Although nutrient depletion has been identified as the main factor in the induction of sporulation in *Bacillus* species (Losick *et al.*, 1986; Sonenshein, 1989), among the *Clostridium* spp. there is little evidence to suggest that nutrient limitation is involved (reviewed by Woods and Jones, 1986). In *C. acetobutylicum* the initiation of sporulation only occurred under conditions in which growth was limited in the presence of excess glucose and ammonia (Long *et al.*, 1984). Sporulation was initiated by the same factors which induce solvent formation, including a decrease in pH and the accumulation of acid end-products, suggesting a parallel switch for both solventogenesis and sporulation. In support to this are the findings of Gotschal and Morris (1981b), who have shown that loss of spore forming capability during continous cultivation of *C. acetobutylicum* is associated with a loss of ability to make solvents.

The central position held by glutamine synthetase (GS; EC 6.3.1.2.) in nitrogen and carbon metabolism suggests that it may be involved in the regulation of metabolism and sporulation in *Bacillus* spp.. Reysset and colleagues (1978) investigated the role of GS in the proces of sporulation in *Bacillus megaterium* and *B. subtilis*. They isolated glutamine auxotrophs with reduced GS production which were often deficient in their ability to sporulate. The majority of glutamine-independent revertants regained normal sporulation properties and GS activity. However, the pattern was not always consistent and some *B. megaterium* revertants, selected primarily for their ability to sporulate, still required glutamine for growth.

In all, the involvement of GS in the initiation of sporulation is highly controversial. Neither enzymes or metabolites of the nitrogen assimilation pathway in *B. licheniformis* appeared to be part of the signal which controls sporulation (Schreier *et al.*, 1981). Analogous results were obtained in *B. subtilis* where GS activity was unaltered during the onset of sporulation when bacteria were induced to sporulate by carbon or nitrogen limitation (Pan and Coote, 1979). However, investigation of the involvement of GS in the initiation of sporulation in *Bacillus polymyxa*, a nitrogen fixing facultative anaerobe, led to two important observations (Kantengwa and Ojha, 1985; Kantengwa and Ojha, 1986): *i*) *B. polymyxa* cells, committed to sporulation, showed a decline in the intracellular glutamate pool and had decreased GS activity and *ii*) the specific inhibition of GS activity by L-methionine sulfoximine resulted in the induction of sporulation (it is interesting that decoyinine, which is widely used to derepress sporulation under a variety of conditions [Mitani *et al.*, 1977], also inhibited GS activity). The GS from *B. polymyxa* has recently been purified and characterized (Ojha and Kantengwa, 1989).

1.2. Prospects for the genetic manipulation of *C. acetobutylicum*.

The recent developments in biotechnology have again focussed attention on the fermentation route for the production of ethanol and butanol and has led to a marked increase in research in these areas (reviewed by Gibbs, 1983; Jones and Woods, 1986b; Jones and Woods, 1988; Young *et al.*, 1989). The future of the ABE fermentation as an industrial process will depend on both the improvement of the process technology and on the improvement of solvent producing strains by genetic manipulation.

The development of efficient gene transfer systems for *C. acetobutylicum* is essential for the genetic manipulation of the bacterium. However, the construction of stable shuttle- vectors containing origins of replication from *C. acetobutylicum* has not been reported. In addition, transformation procedures have not yet been fully developed for *C. acetobutylicum* and chemical treatments or a natural process to induce competence have not been identified for this organism. However, protoplast transformation, similar to systems described for *B. subtilis*, has been demonstrated in *C. acetobutylicum* P262 (Allcock *et al.*, 1982), *C. acetobutylicum* N1-4080 (Reysset *et al.*, 1987), and in the autolysin defective strain *C. acetobutylicum* N-4081 (Reysset *et al.*, 1988)

High levels of extracellular and cell-bound deoxyribonuclease (DNase) activity are associated with the majority of *C. acetobutylicum* strains (Jones and Woods, 1986). These features may hamper the isolation of plasmids and chromosomal DNA and interfere with transformation and transfection. Various ways to overcome the extracellular nuclease problem have been tried, though with limited success (reviewed in Jones and Woods, 1986b). Plasmid transfer by conjugation from *Streptococcus* spp. to *C. acetobutylicum* has been reported (Oultram *et al.*, 1985; Reysset and Sebald, 1985; Oultram *et al.*, 1987) and it will be of great interest to determine the usefulness of integrative plasmids and plasmids which can act as transposon delivery systems. Few transposons of Gram-positive origin have been characterized. The successful expression of the streptococcal transposon Tn*917* in *C. acetobutylicum* was reported by Yu and Pearce (1986). However, no data were provided as to whether Tn*917* was incorporated into the host chromosome or still remained on the plasmid used for transfer. Only very recently, Bertram and Dürre (1989) clearly demonstrated that the streptococcal transposons Tn*916* and Tn*925* can be transferred from

Streptococcus faecalis to the chromosome of *C. acetobutylicum* in which the *tetM* resistance determinant proved to be a valuable selection marker. Moreover, using Tn*916*, these authors were able to obtain *C. acetobutylicum* mutants defective in a regulatory locus for solventogenesis. It is very interesting that Tn*916* can undergo precise deletion in *E. coli* (Gawron-Burke and Clewell, 1984). Tn*916* could thus be a powerful tool to clone specific genes (targetted by Tn*916* directed mutagenesis) from *C. acetobutylicum* into *E. coli*.

Bacteriophages, which have caused problems in industrial AB fermentations, are also potential cloning vectors. The occurrence of lysogeny (Hongo *et al.*, 1968) and phage induction (Ogata and Hongo, 1979) has been demonstrated although transduction has not been reported. It may be feasible to develop bacteriophage vectors analogous to the ϕ105 *B. subtilus* vectors (Errington, 1984).

During the last few years, an alternative method for the uptake of DNA by bacterial cells has been used. This method is based on the disruption of membrane structures by an electrical current, applied to the cells (for review, see Fiedler and Wirth, 1988; Wirth *et al.*, 1989). In general, electroporation has yielded high transformation efficiencies with well-known microorganisms such as *E. coli* (up to $1.0\text{x}10^{10}$ transformants per μg DNA). Recently, plasmids, based on the broad host-range streptococcal plasmid pAMβ1, have been successfully introduced to *C. acetobutylicum* ATCC

8052 cells by electroporation (Oultram *et al.*, 1988). However, a rather low frequency of transformation (1.1×10^3 cfu/μg plasmid DNA) was observed, ruling out current protocols for the shotgun cloning of manipulated plasmid DNA directly into *C. acetobutylicum*.

In view of the difficulties of transferring DNA into *C. acetobutylicum*, the cloning and study of genes in *C. acetobutylicum* is impeded at present. An alternative approach is to clone and study *C.acetobutylicum* genes in other bacteria. This facilitates the characterization of important structural and regulatory genes and allows the study of gene function by site-directed mutagenesis. An advantage of cloning genes from genetically poorly studied strains into well-characterized bacteria such as *E. coli* and *B. subtilus* is the availability of many defined mutants useful for the analysis of gene function and regulation by complementation.

1.3. Regulation of gene-expression.

Synthesis of a functional protein depends upon transcription of the appropriate gene, translation of the mRNA and, in many cases, post-translational processing and compartmentalization of the nascent polypeptide. Other important factors involved in the expression of a gene are mRNA stability, codon usage and protein stability.

1.3.1. Some facts about transcription.

This process is initiated by binding of the RNA polymerase to a regulatory region within the DNA, termed promoter sequence. One strand of the DNA template is transcribed which results in a messenger RNA (mRNA) molecule. The RNA polymerases core enzyme in prokaryotes is a complex protein consisting of a number of different subunits. Although the core enzyme ($\alpha_2\beta\beta'$) is transcriptionally active, it does not initiate specifically at promoters. Promoter specificity is mediated by a variety of additional factors termed sigma (σ) factors (for review, see Doi and Wang, 1986; Ishihama, 1988) which play an important role in the adaptation of the bacterial cell to environmental changes.

1.3.1.1. Structure of *E. coli* Eσ^{70} promoters.

E. coli promoters recognized by RNA polymerase containing the major sigma factor (σ^{70}), consist of two relatively conserved elements, the entry region TTGACA and the Pribnow box TATAAT, respectively about 35 and 10 bp away from the initiation nucleotide from which RNA synthesis starts (Rosenberg and Court, 1979; Siebenlist *et al.*, 1980; Hawley and McClure, 1983; Harley and Reynolds, 1987). The distance between these two elements plays an important role in promoter function (Ayers *et al.*, 1989) and structural differences within this spacer region also affect promoter activity (Auble and deHaseth, 1988; Lozinski *et al.*, 1989). In addition, DNA sequences upstream of the -35 region affect the level of transcription and are associated with DNA curvature (Plaskon and Wartell, 1987). It also has been shown that the early transcribed region (between +1 and +20, +1 being the first nucleotide transcribed) can contribute significantly to promoter activity (Kammerer *et al.* 1986). Consequently, these authors have defined a promoter as a sequence extending from +20 to -50 and subdivide it into a 'core', an 'upstream' (USR) and a 'downstream' (DSR) region.

1.3.1.2. Structure of *B. subtilis* $E\sigma^{43}$ promoters.

B. subtilus genes expressed during growth ($E\sigma^{43}$ dependent transcription) are usually transcribed from promoters showing structural similarities with *E. coli* $E\sigma^{70}$ promoters. However, it seems that besides the traditional core regions, up- and downstream regions (around positions -45 and -5 respectively) also display conserved sequences. This appears to be the case in most other gram-positive organisms (Graves and Rabinowitz, 1986). In addition, a survey of *B. subtilis* vegetative promoters showed a strong conservation of the sequence RTRTG (R=purine) at positions -18 to -14 upstream from the site of transcription initiation (Moran *et al.*, 1982; Graves and Rabinowitz, 1986). Of special interest is the TG pair adjacent to the TATAAT Pribnow box which is higly conserved (83% in Gram-positives and 28% in *E. coli*; Fig. 1.2). It has been suggested that such a TG pair is sufficient to compensate for the absence of specific -35 region sequences (Ponnambalam *et al.*, 1986; Keilty and Rosenberg, 1987).

A compilation of promoter sequences from 112 gram-negative promoters (Hawley and McClure, 1983; McClure, 1985) and 29 Gram-positive promoters (Graves and Rabinowitz, 1986) is presented in Fig. 1.2. Oliphant and Struhl (1988) developed a method in wich functional -35 and -10 promoter regions were obtained with random selection. They showed that the derived consensus sequence of the core region of *E. coli* $E\sigma^{70}$ promoters was in close agreement with that obtained by wild-type sequence comparisons and mutational analysis.

	-45 -35 -10 -5
$E\sigma^{A}$	Ta-AAAAA----$\underline{\text{TT}}$GACA---a--A---a-T-$\underline{\text{T}}$G-$\underline{\text{TA}}$TA$\underline{\text{AT}}$-AAtAt
$E\sigma^{70}$	---a------t-$\underline{\text{TTG}}$ACAt--t--------t-tg-$\underline{\text{TA}}$TAA$\underline{\text{T}}$------

Fig. 1.2.: Consensus sequences of *E. coli* $E\sigma^{70}$ and *B. subtilis* (and of other Gram-positive bacteria) $E\sigma^{A}$ promoters (Hawley and McClure, 1983; McClure, 1985; Graves and Rabinowitz, 1986). Bases that appear in any given position in more than 41% of the examples are in small letters, those more than 50% in capital letters, and those more than 75% are in underlined capital letters.

1.3.1.3. Transcription termination.

The termination of transcription by RNA-polymerase is a key event in gene expression. In *E. coli*, termination sites at both 5′ and 3′ ends of the RNA transcript have been described (for review, see Yager and von Hippel, 1987). Factor independent termination of transcription in *E. coli* is determined by self complementary DNA sequences leading to a RNA hairpin structure in the nascent transcript which can abort RNA chain elongation and results in release of the transcript (Rosenberg and Court,1979; von Hippel *et al.*, 1984; Platt, 1986). Release of the transcript and the RNA polymerase from the elongation complex can also be modulated by bacterial proteins such as the rho-factor (Roberts, 1969) and NusA (Greenblatt *et al.*, 1981) or tau proteins (Briat *et al.*, 1984). No clear specific DNA structures have been associated with factor dependent transcription terminators (Morgan *et al.*, 1985) although in some cases the presence of stem-loop structures facilitates rho and NusA mediated transcription termination (Platt, 1986).

Whereas the features of transcription terminators for *E. coli* and other enterobacteria is well documented, very little is known of their presence in gram-positive organisms. A sequence determining transcription termination has been characterized for the *Bacillus pumilus cat-86* gene (Mongkolsuk *et al.*, 1985). A variety of genes from both *Streptococcus sanquis* and *Staphylococcus aureus* contain potential stem-loop structures which seem to be implicated in the processes of inducibility (Byeon *et al.*, 1984) and attenuation (Horinouchi *et al.*, 1983). A functional terminator of transcription has been reported for the *Streptomyces fradiae aph* gene (Pulido and Jiménez, 1987). Shimotsu *et al.* (1986) have proposed that control of the *B. subtilus trp* operon was mediated by transcription attenuation in wich a stem-loop structure provides termination of transcription in response to the availability of tryptophan.

1.3.2. The importance of efficient translation.

This process in which the mRNA nucleotide sequence is translated into an amino acid sequence by the ribosome complex can be subdivided into three phases : i) translation initiation by interaction between the 16 S ribosomal RNA and a specific sequence on the mRNA (ribosome binding site, RBS), ii) elongation of the polypeptide chain and iii) dissociation of the ribosome-mRNA-polypeptide complex.

The start codon is usually an AUG codon, but prokaryotic ribosomes start occasionally at GUG or UUG triplets (reviewed by Kozak, 1983). *B. subtilis* and other gram-positive organisms have a much higher frequency for non-AUG initiation codons than *E. coli* (Hager and Rabinowitz, 1985).

In general, *E. coli* and *B. subtilis* RBSs consist of the start codon and a purine rich sequence 4 to 15 nt separated from the start codon. The consensus sequence 5′-AGGAGGU-′3 (Shine-Dalgarno (SD) sequence, Shine and Dalgarno, 1974) is complementary to the 3′ end of 16S rRNA at the positions 42 to 48 (Gold *et al.*, 1981; Gold *et al.*, 1988). In addition, Petersen (1988) has provided evidence that base pairing between mRNA and 16S rRNA in *E. coli* occurs over an extended region. The distance between the initiation codon and the SD is important for translation efficiency and statistical analysis has revealed a preferential spacing of 5 nt between the SD region and the startcodon (Scherer *et al.*, 1980). Other factors affecting translation are: (i) the sequence between AUG and SD (Hui *et al.*, 1984), (ii) the sequence of the 5′ untranslated region in front of the RBS (Stanssens *et al.*, 1985) and (iii) sequence modifications in the amino terminal end of the coding region (Zabeau and Stanley, 1982; Looman *et al.*, 1987; Petersen *et al.*, 1987). The discrimination between 'false' and 'real' starts for translation has been investigated by Dreyfus (1988) who has shown that, besides the SD element and start codon, the RBS of 'real genes' contain additional primary sequence information that is essential for efficient initiation of translation. Evidence was presented that a high content of As, especially downstream from the SD element, is a favourable factor. Recently, Thanaraj and Pandit (1989) have proposed an additional ribosome recognition site for highly expressed genes in *E. coli*. The site consists of a subsequence of 5′-UGAUCC-′3 which is complementary with the sequence 5′-GGAUCA-′3 starting at position 1529 in the 16 S rRNA. Apparantly, the average Gibbs free energy (ΔG) of -5.1 kcal/mol arising out of the binding of 16 S rRNA with the additional site can lead up to a thousand-fold increase in the association strength of ribosome binding and Thanaraj and Pandit (1989) have termed this site as the Translation-initiation Promoting Site (TPS). Recently, Schauder and McCarthy (1989) have confined the stretch of mRNA controlling the efficiency of initiation to the 'translational initiation region' (TIR), which extends beyond the limits of a classically defined RBS and involves the formation of a secondary structure in the mRNA.

Typical translation initiation sites of *E. coli* involve only a few bases of the SD sequence with an average Gibbs free energy (ΔG) of -11.7 kcal/mol whereas initiation in *B. subtilus* usually involves at least 5 bases paired with the 16 S rRNA (Hager and Rabinowitz, 1985). This translation-initiation specificity may explain why *E. coli* genes are only occasionally expressed in *B. subtilus* (Kreft *et al.*, 1983; Peschke *et al.*, 1985).

There are reports of translational control in *E. coli* mediated by the formation of secondary structures with complementary RNA(s) (reviewed by Simons, 1988; Inouye and Delihas, 1988). In addition, intramolecular long distance interaction in the mRNA between the RBS and a bulge loop structure has also been suggested as a feature of regulation at the level of translation (J. Botterman, Ph.D. Thesis, State University of Ghent, Belgium; Carter-Münschau *et al.*, 1989; Petersen, 1989).

The correlation between codon preference and relative amounts of iso-accepting tRNAs (codon bias) can contribute to differences in gene expression (Grosjean and Fiers, 1982; Gouy and Gautier, 1982; Ikemura and Ozeki, 1983) and it is well established that there is preferential codon usage in both *E. coli* and *S. cervisiae* (reviewed by Ernst, 1988). However, experiments testing the effects of codon usage on translation elongation have led to controversial results because, along with codon usage, other parameters like mRNA structure and degradation may be changed at the same time. A detailed study on the effect of codon usage on translation elongation has been presented recently by Sorensen *et al.* (1989). These authors inserted rare codons in the *E. coli lacZ* gene and measured the change in translation rate. They also introduced secondary structures in the *lacZ* mRNA and found, to their surprise, that this didn't cause any translation delay, although the message was more susceptable to mRNA-degrading enzymes. From this work it was concluded that *i*) codon usage determines translation rate in *E. coli*, and *ii*) the translation rate is indifferent to the presence of secondary structures.

An unusual codon usage in *B. subtilis* has been reported by Ogasawara (1985). Apparently, biased codon usage only exists for a limited number of genes and the number of rare codons is small. This may be related to the unique organisation of tRNA genes in *B. subtilis* (Vold , 1985). It has been suggested that, because tRNA genes in this organism are transcribed frequently and the resultant tRNAs can recognize a wide range of codons, codon usage is more equal in *B. subtilis* than in *E. coli* or *S. cerevisiae* (Ogasawara, 1985).

1.3.3. mRNA Stability and RNA turnover.

Cellular mRNA concentration is determined by the balance between the rates of mRNA synthesis and decay, and mRNA lifetime has consequently been described in terms of susceptibility to endonucleolytic or exonucleolytic digestion. The more vulnerable to one or both of these enzymatic reactions, the more 'unstable' the messenger. Belasco and Higgins (1988) recently reviewed the different aspects of RNA stability. Features that cause resistance or sensitivity to RNAse attack include sequence, secondary stem-loop structures, and the association of mRNA with ribosomes, proteins and antisense RNA.

Differences in mRNA stability can contribute to differential expression of genes within polycistronic operons (Newbury *et al.*, 1987; Chen *et al.*, 1988) and can modulate gene expression in response to environmental stimuli (Nilsson *et al.*, 1984).

1.3.4. Regulation of gene-expression by anti-sense RNA

Naturally occuring anti-sense RNAs which are able to control gene-expression have been discovered only recently. Their biological activities are diverse and are found in many regulatory circuits including osmoregulation (Aiba *et al.*, 1987; Andersen *et al.*, 1987; Andersen *et al.*, 1989), temperal phage development (Belin *et al.*, 1987; Wu *et al.*, 1988), plasmid replication and compatibility (Nordström *et al.*, 1988), transposition of IS elements (Simons and Kleckner, 1983), and catabolite repression (Okamoto and Freundlich, 1986). Basically, specific antisense RNA either affects RNA processing, interferes with the elongation of the target transcript, or inhibits translation of target mRNA(s).

Three major classes of antisense RNAs emerge from the reviews (Green *et al.*, 1986; Pines and Inouye, 1986; Simons, 1988). Class I antisense RNAs are complementary to the functional region of their target mRNA. They form a RNA:RNA hybrid by binding to the SD sequence and/or coding region, thereby directly inhibiting translation, or may destabilize the target message by rendering the duplex hybrid to increased suspectibility of RNases. Control on the level of translation also includes Class II antisense RNAs which are complementary to non-coding regions upstream of the SD sequence and cause refolding of the mRNA, thereby blocking ribosome binding. A third class of antisense RNAs exert their action on the level of transcription by binding just upstream of an A-U rich region of the mRNA. The RNA:RNA hybrid, immediately followed by such an A-U rich region in the mRNA, resembles a factor-independent transcription termination site, causing the RNA-polymerase to pause. Due to instability of the A-U rich sequence, premature release of the transcript occurs and transcription terminates.

1.3.5. Cellular sensory systems.

Prokaryotes are capable of sophisticated sensory behavior. Their sensory systems mediate a wide variety of adaptive responses involving changes in gene expression or cell movement. There is much evidence that adaptation to a change in the environment is accomplished by a network of signal transduction systems each consisting of two interacting proteins with conserved domains (Nixon *et al.*, 1986; Ronson *et al.*, 1987; Magasanik, 1988; Nohno *et al.*, 1989). It appears that modulator proteins process sensory information by phosphorylation or dephosporylation of an active or inactive effector protein (Ninfa *et al.*, 1988; Forst *et al.*, 1989; reviewed by Magasanik, 1988; Stock *et al.*, 1988; Bourret *et al.*, 1989). In addition, Kofoid and Parkinson (1988) have identified so-called "transmitter" and "receiver" (T/R) modules within modulators and effectors. Apparently, modulators all contain a homologous carboxy terminal domain of about 200 residues, whereas effectors share a homologous amino terminal domain of about 130 residues (summarized recently by Gross *et al.*, 1989). From the literature cited above it seems likely that the homologies between modulator proteins and between effector proteins reflect conserved protein kinase function and conserved phosphoacceptor activities respectively.

A large diversity of transcription activation mechanisms, in which a sensory component activates a 'dormant' transcriptional activator component in response to an inducer, has been reported in various bacterial species (Ninfa *et al.*, 1986; Magasanik, 1988; Ninfa *et al.*, 1989 Nohno *et al.*, 1989; Deretic *et al.*, 1989; Miller *et al.*, 1989; Forst *et al.*, 1989; reviewed by Bourret *et al.*, 1989; Gross *et al.*, 1989). An example of such a two- component regulatory system is shown in Fig. 1.3.

Unlike other canonical sensors, NtrB receives its sensory input from other signaling proteins rather than directly from the environment (keeping in mind that the true sensor for nitrogen availability is GS). The amino terminus of NtrB contains a receiver module that may be responsible for regulating its transmitter activity, either in response to sensory signals or as part of a feedback circuit (Kofoid and Parkinson, 1988). These authors suggested that a phosphate acceptor site is located within the receiver domain, which is in agreement with the findings of Keener

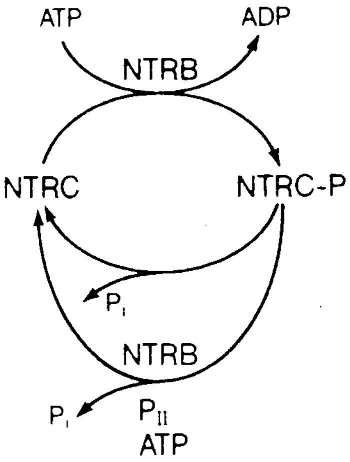


Fig. 1.3.: Model of sensory transduction by kinase and phosphotase activities of NtrB and NtrC (nitrogen regulatory proteins NR_I and NR_{II}, respectively). Activation of transcription (e.g. from the *glnALG* operon) only occurs when NtrC becomes phosphorylated by NtrB. NtrC-P dephosphorylates spontaneously (autophosphotase activity), but is also subject to regulated dephosphorylation that depends on NtrB, the *glnD* gene product (P_{II}), and ATP (from Keener an Kustu, 1988).

and Kustu (1988) who demonstrated that NtrB is capable of autophosphorylation. In addition, it has been reported that some modulators are able to phosphorylate effectors outside their native signal transduction system. Ninfa *et al.* (1988) showed that NR_{II} (gp*glnG*/gp*ntrB*) can phosphorylate CheY, and CheA can also phosphorylate NR_{I} (gp*glnL*/gp*ntrC*). Crosstalk between heterologous modulator/effector pairs provides an explanation for the complex phenotypes often associated with modulator mutations. These include *phoM*-dependent expression of *phoA* in *phoR* mutants (Wanner *et al.*, 1988) and the residual regulation of *glnA* in *glnB* (*ntrC*) mutants lacking NR_{II} (Backman *et al.*, 1983).

1.4. Bacterial systems involved in ammonia assimilation.

Ammonia is the nitrogen source of choice for most microorgansims. In many cases, however, inorganic compounds such as dinitrogen, nitrate or urea and complex organic compounds such as histidine, arginine and proline can be converted to ammonia via enzymatic processes. The exact set of pathways used to accomplish this conversion is characteristic for each organism. Of central importance to the biosynthesis of nitrogen compounds are the metabolic intermediates glutamine and glutamate. They provide cellular nitrogen for the synthesis of purines, pyrimidines, amino sugars, most of the amino acids and metabolites such as NAD and p-aminobenzoate.

The three main reactions, generally used in bacteria for ammonia assimilation and the synthesis of glutamate and glutamine, are shown below (gln = glutamine; glt = glutamate; α-kgt = α-ketoglutarate) :

(1) NH_3 + glt + ATP ------> gln + ADP + P_i

(2) gln + α-kgt + NADPH ------> 2 glt + $NADP^+$

(3) NH_3 + α-kgt + NADPH ------> glt + $NADP^+$

These three reactions are catalyzed by the action of GS (EC 6.3.1.2), glutamate synthase (GOGAT, glutamine-oxoglutarate amido transferase, EC 1.4.1.13) and glutamate dehydrogenase (GDH, EC 1.4.1.4.) respectively.

The regulation of the metabolic enzymes involved in bacterial ammonia assimilation has been reviewed comprehensively by several authors (Tyler, 1978; Magasanik, 1982; Schreier *et al.*, 1981; Magasanik and Neidhardt, 1987; Reitzer and Magasanik, 1987; Magasanik, 1988).

1.4.1. GDH.

GDH usually has a high K_m for ammonia and thus functions in most enteric bacteria only when an excess of ammonia is present in the medium (more than 1 mM) (Magasanik, 1982). Moreover, it is thought to be a completely dispensable enzyme as strains deficient in GDH have no unusual growth phenotype. This situation is quite different for several nitrogen-fixing bacteria such as *B. polymyxa* (Kanamori *et al.*, 1987a) and *Bacillis macerans* (Kanamori *et al.*, 1987b) which prefer the GDH pathway for ammonia assimilation in both ammonia grown and nitrogen-fixing cells. However, this is not the case for all nitrogen fixers: *Clostridium kluyveri* GDH is derepressed only in ammonia-grown cells and in *Clostridium butyricum*, GDH activity was only detectable when it was grown in a complex medium, but was virtually absent when cells were grown on synthetic glucose-salt media with NH_4^+ or N_2 as the nitrogen source (Kanamori *et al.*, 1989). In addition, *Klebsiella pneumoniae* (Nagatani *et al.*, 1971), *Clostridium pasteurianum* (Nagatani *et al.*, 1981), and *Bacillus azotofixans* (Kanamori *et al.*, 1988) have barely detectable GDH activities (even when grown in excess ammonia) or have a high K_m for NH_4^+. This diversity among the nitrogen fixers may be due to the fact that some organisms, such as *C. butyricum*, utilise high-energy fermentation pathways (with the generation of ATP) which may be more favorable for the utilisation of the ATP-requiring GS-GOGAT pathway (Kanamori *et al.*, 1989) and thus omit the GDH route for ammonia assimilation. In contrast, the assimilation of ammonia through the GDH pathway without further expenditure of energy may be necessary in nitrogen-fixing species such as *B. polymyxa*, which utilise low energy-yielding fermentation pathways.

Bacteria of the family *Rhizobaceae*, including *Bradyrhizobium japonicum* and *Agrobacterium tumefaciens*, are distinct from the *Enterobacteriaceae* in that most members contain two GS enzymes, GSI and GSII, encoded by *glnA* and *glnII*, respectively (Darrow *et al.*, 1977; Fuchs and Keister, 1980). No GDH activity has been detected in *Rhizobium phaseoli* (Bravo and Mora, 1988) and it was shown that when the *E. coli gdhA* gene was introduced into *R. phaseoli*, plants were preferentially nodulated by cells which had lost the plasmid born *gdhA* gene, indicating that the presence of GDH is incompatible with N_2 fixation.

It is thought that the obligate anaerobe *C. acetobutylicum* P262 is capable of nitrogen fixation; moreover, genes showing moderate amino acid similarities with *nifI* and *nifJ* genes of *K. pneumoniae* have been cloned in *E. coli* (Santangelo and Woods, personal communication). Although the study of NAD(P)H- dependent GDH in *Clostridium* species has been hampered by the presence of NADH oxidase and NAD(P)H-ferredoxin reductase in cell extracts (which cause very rapid oxidation of NAD(P)H during the spectrophotometric assay), it would nevertheless be of great interest to investigate the importance of the GDH pathway for ammonia assimilation in *C. acetobutylicum* as it is known that in this organism ammonia levels are important for solventogenesis and the initiation of sporulation (Long *et al.*, 1984; Jones and Woods, 1986). Recently, the *gdh* gene of *C. acetobutylicum* P262 has been cloned into *E. coli* strain PA340 (*gdh-1 gltB31*) (Fierro-Monti and Woods, personal communication).

In *B. subtilus*, only GOGAT is implicated in the synthesis of glutamate (Desphande *et al.*, 1980). Apparently, GDH, if present at all, plays a minor role in ammonia assimilation. Recently, Fisher (1989a) showed that both GOGAT and GDH activities are present in *Streptomyces coelicolor*, and showed that GDH activity is derepressed up to ten-fold when cells were grown with excess ammonia.

An unusual scheme of glutamine metabolism, namely the uptake and transport of glutamine with the subsequent conversion of glutamine to pyro-glutamate and ammonia, has been reported in the ruminal bacterium *Streptococcus bovis* (Chen and Russel, 1989). A high GDH activity was observed but GOGAT activity could not be detected. In addition, growth of *S. bovis* as dependent on the presence of glutamine in the medium, indicating a non-functional, if not absent, GS enzyme. It appeared that glutamate was produced entirely at the expense of glucose via the tricarboxylic acid (TCA) cycle intermediates isocitrate and α-ketoglutarate because (1) the carbon skeleton of glutamine was lost as pyro-glutamate, and (2) GOGAT to recycle glutamine into glutamate was absent. Consequently, GDH was thought to be the sole enzyme involved in ammonia assimilation and glutamate production in *S. bovis*.

The nucleotide sequence of *gdh* genes has been determined for many organisms including *E. coli* (McPherson and Wootton, 1983; Valle *et al.*, 1984), *S. typhimurium* (Miller and Brenchley, 1984), *Klebsiella aerogenes* (Mountain *et al.*, 1985), *Neurospora crassa* (Kinnaird and Fincham, 1983), *Saccharomyces cerevisiae* (Nagusa and Hall., 1985), and *Aspergilus nidulans* (Hawkins *et al.*, 1989). GDHs show a relatively high degree of sequence conservation, as demonstrated by amino acid sequence comparisons and alignment studies (Valle *et al.*, 1984; Hawkins *et al.*, 1989).

1.4.2. GOGAT.

Because bacterial GDH has a relatively high K_m for ammonia, reductive amination of α-ketoglutarate by the GDH pathway becomes unfavourable in many procaryotes when ammonium levels drop below 0.1 mM. Under these conditions, glutamate is produced mainly by the GOGAT pathway. GOGAT was first discovered in ammonia limited cultures of *K. aerogenes* (Tempest and Meers, 1970) and is now known to exist in many bacterial species, yeasts and plants. The enzyme has been purified to homogeneity from *E. coli* (Miller and Stadtman, 1972), *K. aerogenes* (Trotta *et al.*, 1974), *Thiobacillus thioparus* (Adachi and Suzuki, 1977), *B. megaterium* (Hemmila and Mantsala, 1978), *B. licheniformis* (Schreier and Bernlohr, 1984), *Azospirillum brasilense* (Ratti *et al.*, 1985), *Saccharomyces cerevisiae* (Masters and Meister, 1982), and *Nocardia medeterranei* (Mei and Jiao, 1989). The bacterial enzyme is an iron-sulphur flavoprotein which exists as a heterodimer of glutaminase (large subunit) and glutamate dehydrogenase (small subunit) activities. Apparently, NAD(P)H binds to the small subunit and transfers electrons to the large subunit, which reduces the flavin.

Subsequently, α-ketoglutarate binds to the small subunit, and glutamine binds to the large subunit. The glutamine amide is then transferred to α-ketoglutarate and the reduced flavin reduces in turn a proposed iminoglutarate intermediate to glutamate (Miller and Stadtman, 1972; Rendina and Orme-Johnson, 1978; Reitzer and Magasanik, 1987). The oligomeric structure seems to differ among the various species. The enzymes from *E. coli*, *B. megaterium*, and *A. brasilense* show a quaternary structure of the type $\alpha_4\beta_4$, whereas those from *K. aerogenes*, *B. licheniformis*, *T. thioparus* and *N. mediterranei* are of the $\alpha\beta$ type. In addition, the M_r of the two subunits may differ from organism to organism.

In both *E. coli* and *S. typhimurium*, the GOGAT subunit genes are genetically linked and cotranscribed (Lozoya *et al.*, 1980; Madonna *et al.*, 1985). The large and small subunits are encoded by the *gltB* and *gltD* genes, respectively. In *E. coli*, a third gene *gltF*, located immediately downstream of *gltD*, encodes a regulatory peptide thought to be involved in *Ntr* regulation (Castaño *et al.*, 1988). Evidence has been presented that *gltB* mutations can impose the Ntr$^-$ phenotype, which is the inability to utilize substrates such as arginine and proline, whose degradation yields either ammonia or glutamate (Pahel *et al.*, 1978). A relationship between the *gltBDF* locus and the Ntr network is further suggested by the finding that the Ntr phenotype of *gltB* mutants can be supressed by mutations in *ntrB* (*glnL*)

(Pahel *et al.*, 1978), a regulatory gene in the GS *glnALG* operon. *ntrB* mutations in turn, cause constitutive expression of the *glnA* gene (GlnC phenotype) because information about nitrogen availability is not processed. The exact mechanism of 'cross-talk' between the *gltBDF* and *glnALG* operons in *E. coli* is poorly understood, but it is likely that intracellular levels of glutamate play a role. It is known that glutamate has two opposite effects on the expression of the *gltBDF* and *glnALG* operons. Glutamate not only represses GOGAT activity (Miller and Stadtman, 1972), but is also able to derepress completely GS activity (Magasanik and Neidhardt, 1987; Castaño *et al.*, 1988). The latter effect is not caused solely by nitrogen limitation, but seems to involve the *gltF* gene product: *gltF* mutants are unable to fully derepress genes which respond to NtrA (GlnF) and NtrC (GlnG) under nitrogen limiting conditions (Gsd$^-$ phenotype) (Castaño *et al.*, 1988). A similar Gsd$^-$ phenotype has been observed previously for *K. aerogenes nac* (nitrogen assimilation control) mutants (Bender *et al.*, 1983) although the regulation of *glnALG* was not affected. Instead, strains carrying the *nac* allele were unable to utilize proline, histidine, or serine, amino acids whose biosynthesis is Ntr-dependent. This *nac* gene has not been identified in *E. coli*.

The *glt* locus of *Azorhizobium sesbaniae* ORS571 has recently been cloned into *E. coli* (Hilgert *et al.*,1987; Donald *et al.*, 1988). *A. sesbaniae*, which induces stem and root nodules on its host plant, *Sesbania rostrata*, is not only able to fix nitrogen during symbiosis (Fix phenotype), but is also capable of dinitrogen-dependent growth in the free-living state (Nif phenotype) (Dreyfus *et al.*, 1983). It was found that the *glt* locus of *A. sesbaniae* was organised in a similar fashion as for the *gltB* and *gltD* genes in enteric bacteria, although the presence of a *gltF* gene was not established. It was also suggested that *gltB* and *gltD* may not be co-transcribed, as is the case in enterics, but rather constitute two independent cistrons.

The GOGAT genes in *B. subtilis*, *gltA* and *gltB*, are controlled by the product of *gltC*, which is situated immediately upstream of the *gltA* gene. It has been shown that *gltA* and *gltB* define two cistrons which are independently transcribed (Deshpande and Kane *et al.*, 1980). The third gene, *gltC*, is transcribed in the opposite direction (Bohannon *et al.*, 1985) and encodes a positive regulator which is thought to bind the *gltCA* intergenic region when levels of glutamate are low (Bohannon *et al.*, 1989). In addition, it has been suggested that GltC is a negative regulator of its own synthesis. GltC would undergo a conformational change under conditions of limiting glutamate, such that it acts to stimulate the association of RNA polymerase with the *gltA* promoter, increasing *gltA* transcription. This conformational change, however, would not alter its ability to bind DNA or to repress *gltC* transcription. How this duality in function is accomplished remains unanswered. It may be possible that GltC exists as a mixture of an active and an inactive form in an equilibrium which would shift towards the more active form when expression of *gltAB* is required. A similar situation has been reported for the *E. coli* nitrogen regulator I (NR_I) wich displays an analogous pleiotropic function in the control of the *glnALG* operon (Weglenski *et al.*, 1989). The complete DNA sequence of the *gltC* gene has been determined (Bohannon *et al.*, 1989). The derived amino acid sequence for GltC shows a region with extensive similarity to a region in the DNA-binding domain of *B. subtilis* sigma factor σ^E. Comparison of the GltC sequence to that of other proteins revealed that it is very similar to members of the "LysR family" of DNA binding regulatory proteins, *LysR, IlvY* and *CysB* of *E. coli* and *NodD* of *Rhizobium trifolii* (Bohannon *et al.*, 1989). The *gltCA* intergenic region contained reiterated sequences with consensus sequence 5′-ATATTGTTT-3′. It is proposed that these sequences represent potential regulatory sequences for GltC binding. As described by Bohannon *et al.* (1989), similar sequences have been found in the target DNAs of several other LysR family proteins. Further study is needed to establish the significance of these sites. It is interesting that the *B. subtilis citB* gene, which codes for aconitase, has several copies of this sequence located in the proximity of its promoter (Dingman et al., 1987a; personal observation). It would be of interest to investigate the importance of these sites because aconitase levels in *B. subtilis* have been associated with changes in the intracellular concentration of α-ketoglutarate (Fisher and Magasanik, 1984) and sporulation (Dingman *et al.*, 1987b).

1.4.3. GS

The importance of this enzyme in metabolic pathways is reflected in the vast amount of literature available. It has a central role in neurotransmitter metabolism in astrocytes of the human brain (Sc. Am.,pp 44-52, april 1989), is the target enzyme in weed control due to its susceptebility to specific inhibitors, such as L-phosphinothricin, present in commercially available herbicides (De Block *et al.*, 1987) and has a key function in nitrogen metabolism and/or ammonia assimilation of numerous organisms, ranging from mammals and higher plants to fungi and bacteria. The GS enzyme is subject to a variety of control systems, which may include feedback inhibition by metabolites, regulation via adenylylation-deadenylyation, oxidation, proteolytic degradation, temperature, and the kinetic parameters intrinsic to the enzyme. In addition, transcription of its structural gene is regulated by nitrogen- and carbon-catabolite control mechanisms and, for eukaryotic genes, hormone and fat-specific regulatory elements have been identified upstream of the transcription start. Although the emphasis of this review is on the structure, function and properties of bacterial GS enzymes, some aspects of eukaryotic GS enzymes are included.

Table 1.1. : Cloned GS genes (updated nov.'89).

Organism	Nu/Ad/Se[a]	Reference
E. coli	12/+/+	Backman *et al.*, 1981 Colombo and Villafranca, 1986
S. thypimurium	12/+/+	Koduri *et al.*, 1980 Janson *et al.*, 1986
Vibrio alginolyticus	12/+/+	Bodasing *et al.*,1986 Maharaj *et al.*, 1989
Thiobacillus ferrooxidans	12/+/+	Barros *et al.*, 1985;1986 Rawlings *et al.*, 1987
Bacteroides fragilis	6/-/+	Southern, 1986;1987 Hill *et al.*, in press.
Flavobacterium	(b)	Goodman and Woods (pers.com.)
K. pneumonia	12/*/-	Espin *et al.*, 1982 Alvarez-Morales *et al.*, 1984
B. subtilis	12/-/+	Gardner and Aronson, 1984 Fisher *et al.*, 1984 Strauch *et al.*, 1988 Nakano *et al.*, 1988
Bacillus cereus	12/*/+	Nakano *et al.*, 1989
C. acetobutylicum	12/-/+	Usdin *et al.*, 1986 Janssen *et al.*, 1988
A. tumefaciens	*/*/-	Rossbach *et al.*, 1988
S. coelicolor[c]	12/+/+	Wray and Fisher, 1988 Fisher and Wray, 1989
Anabaena spp.7120	12/*/+	Fisher *et al.*, 1981 Tumer *et al.*, 1983
R. leguminosarum	12/*/+	Filser *et al.*, 1986 Colonna-Romano *et al.*, 1987
R. meliloti	12/*/-	Somerville and Kahn, 1983
B. japonicum	12/*/+	Carlson and Chelm, 1985 Carlson and Chelm, 1986
Azotobacter vinelandii	*/*/-	Toukdarian and Kennedy, 1986
Spirulina platensis	*/*/-	Riccardi *et al.*, 1985
Azospirillum brasilense	*/*/+	Bozouklian and Elmerich, 1986
Rhodopseudomonas capsulata	12/+/-	Scolnik *et al.*, 1983 Caballero *et al.*, 1985

(a) Specifications: Nu, number of subunits; Ad, regulated by adenylylation (+) or not (-); Se, amino acid sequence available (+) or not (-): (*), data not available
(b) work in progres
(c)putatively regulated by adenylylation/deadenylylation

1.4.3.1. Structure and function of GS.

Two structural forms of the prokaryotic GS enzyme, which are termed GSI and GSII, have been described so far.

Most bacteria investigated have only the GSI enzyme but members of the *Rhizobiaceae* contain both GSI and GSII enzymes (Fuchs and Keister, 1980a). These GS enzymes differ in a number of physical properties including molecular size, heat stability and isoelectric pH, and show different post-translational control (Darrow and Knotts, 1977; Fuchs and Keister, 1980b). They also differ markedly in the mechanism in which their respective genes, *glnA* and *glnII*, are regulated (Howitt and Gresshoff, 1983; Rossbach *et al.*, 1987; Carlson *et al.*, 1987; Martin *et al.*, 1988; Rossi *et al.*, 1989). Two isoforms of the GS enzyme have been reported for the extreme thermophile *Bacillis caldolyticus* (Wedler *et al.*, 1980). Both enzymes are composed of 12 subunits, with a total M_r of approximately 620 kilodaltons, but differ in their physical and kinetic properties.

Procaryotic GSI enzymes are dodecamers with a M_r of approximately 600 kilodaltons, composed of a single type of subunit (Streicher and Tyler, 1980; Bhatnagar *et al.*, 1986). However, prokaryotic GS enzymes which do not have the dodemeric structure have been reported (Krishnan *et al.*, 1986; see also table 1.1). Dodemeric enzymes have a typical toroid structure consisting of two rings of 6 subunits. An unusual feature of the GS enzyme is the formation of 12 active sites at heterologous interfaces between adjacent subunits. According to Almassy *et al.* (1986), each active site is formed by six β-strands of the large C-domain (containing the adenylylation site Tyr-397) of one subunit and two β-strands from the Trp-57 loop of the smaller N-terminal domain of the next subunit. An atomic model for the dodemeric GS enzyme of *S. typhimurium* has been determined by algorithm predictions (Janson *et al.*, 1986) and X-ray crystallography (Almassy *et al.*, 1986; Yamashita *et al.*, 1989) (Fig. 1.4). Furthermore, a detailed study of the chemical mechanisms for a typical GS reaction with ATP, glutamate and ammonia as substrates, has been presented by Colanduoni *et al.* (1987).

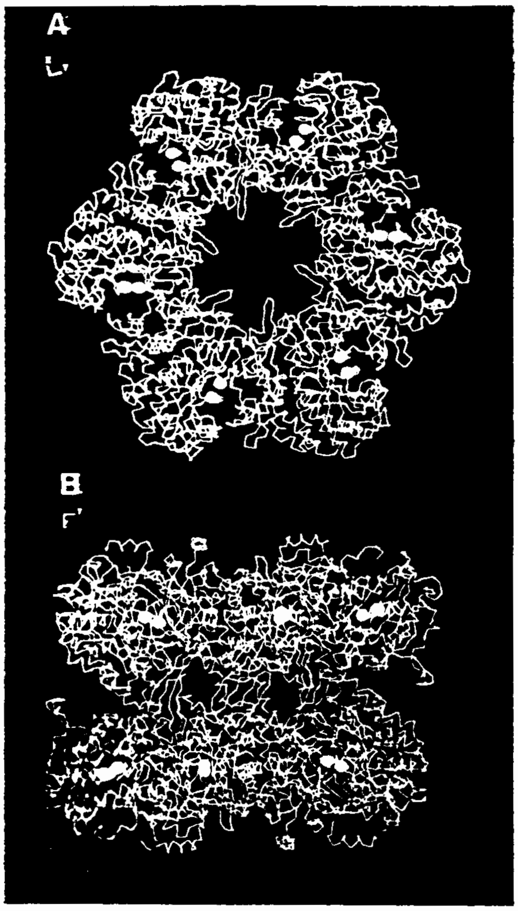


Fig. 1.4.: Presentation of the quaternary structure of GS. Each active site is indicated by a pair of spherical Mn^{2+} ions. The GS molecule can be seen projected down the 6-fold axis (A) or is viewed from the side (B), in which only the six nearest subunits are shown.

Whereas GSI subunits vary in length from 444 to 474 amino acids, GSII is similar to eukaryotic octameric GS enzymes which contain only 355 to 373 amino acids per subunit and thus lack the carboxy-terminal portion including the adenylylation site (Carlson and Chelm, 1986). Based on the fact that *B. japonicum* GSII shares homology with GS proteins of higher plants, these authors have suggested that GSII was aquired from plants by transfer of genetic material. Furtheron, they suggested that GSII from *Rhizobiaceae* could possibly play a special role in plantbacterial interactions. However, it was shown recently by Rossbach *et al.* (*1988*) that neither GSI nor GSII are important for virulence or opine metabolism in *A. tumefaciens*.

Besides the *glnA* and *glnII* loci, a third novel locus *(glnT)*, which *is* able to direct glutamine biosynthesis in the *E. coli* mutant YMC11 *(glnALG)*, has been cloned from *A. tumefaciens* (Rossbach *et al.*, 1988) *and R. meliloti* (de Bruijn *et al.*, 1985). In both cases no GS transferase or biosynthetic acivity could be detected in YMC11(GlnT$^+$) cells, in spite of the fact that they could grow (albeit extremely slowly) on minimal plates without glutamine. No other enzymatic activity capable of catalyzing glutamine biosynthesis in *E. coli* except GS has been identified thus far. It is possible that GS activity directed by the *glnT* locus is the result of an enzymatic reaction normally coupled to another biosynthetic pathway, such as purine biosynthesis.

Recently, Hill *et al.* (1989) have determined the nucleotide sequence of the *Bacteroides fragilis glnA* gene encoding a novel GS enzyme which exists as a hexamer of identical subunits with an apparant Mr of 75 000 (Southern *et al.*, 1987) and which is not regulated by adenylylation. The *B. fragilis* GS subunit is approximately 270 and 400 amino acids longer than the GSI and GSII subunits, respectively. Five β-strands, associated with the active site (Almassy *et al.*, 1986), represent highly conserved regions in the amino acid sequence of GSI and GSII enzymes (Janssen *et al.*, 1988). Only four such regions with moderate similarity were present in the amino acid sequence of the *B. fragilis* GS. Furthermore, it differs from all the other prokaryotic GS 's in that the active site Trp residue in region I is replaced by a Val residue which shows no functional similarity with Trp (the Trp residue in the GS of the gram-positive bacteria *B. subtilus*, *S. coelicolor* and *C. acetobutylicum* is replaced by a Phe residue, which is funtionally similar to Trp; Strauch *et al.*, 1988; Wray and Fisher, 1988; Janssen *et al.*, 1988). In view to these differences with GSI and GSII enzymes, the *B. fragilis* GS enzyme has been classified as being the first member of a new class of enzymes, GSIII (Hill *et al.*, 1989).

1.4.3.2. Cascade control of GS.

Because of the importance of glutamine in cellular metabolism and because of the multiple functions it performs, it is not surprising that GS is subject to rigorous control, both in its catalytic activity and its biosynthesis. During the last few years, it has become apparent that these two regulatory mechanisms are linked in the 'bicyclic cascade', a complex process which not only regulates GS enzyme activity, but also regulates the biological function of the *ntrB* (*glnL*) and *ntrC* (*glnG*) gene products (called NR_{II} and NR_I respectively). The cascade control of GS activity has been reviewed in great detail (Magasanik, 1982; Rhee *et al.*, 1985; Reitzer and Magasanik, 1987; Magasanik and Neidhardt, 1987; Magasanik, 1988) and has been shown to be a fast and sensitive process in which GS activity is adjusted accurately in response to nitrogen availabillity.

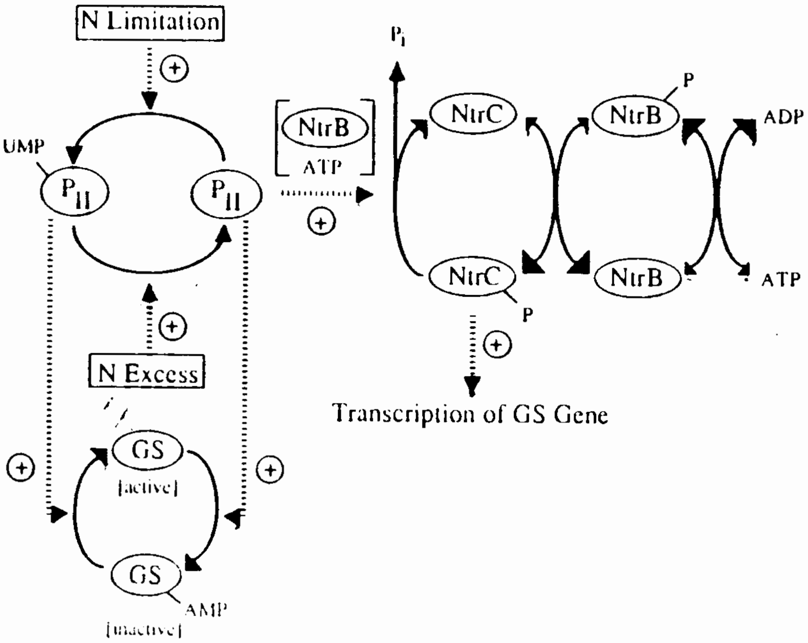


Fig. 1. 5: Cascade control of GS (for explanation, see text) (from Bourret *et al.*, 1989)

Both adenylylation and deadenylylation of GS is carried out by the same enzyme, adenylyltransferase (ATase) (Fig. 1.5.). However, the direction of this reaction is dictated by P_{II} (encoded by *glnB*), a tetrameric protein that exists in two possible forms. In nitrogen excess, when the glutamine to α-ketoglutarate ratio is high, P_{II} is deuridylylated and promotes adenylylation of GS. The average number of adenylylated subunits per dodecamer is inversely proportional to the specific activity of the GS enzyme. Conversely, when nitrogen is scarce (low glutamine to α-ketoglutarate ratio), P_{II} becomes uridylylated and allows deadenylylation of GS by ATase (GlnE) resulting in GS with increasing enzymatic activity.

The different steps in the cascade have been studied extensively (Adler *et al.*, 1975; Garcia and Rhee, 1983; Bueno *et al.*, 1985; Son and Rhee, 1987) and the *E. coli*, *R. leguminosarum* and *K. pneumoniae glnB* genes have been cloned and characterised recently (Son and Rhee, 1987; Colonna-Romano *et al.*, 1987; Holtel and Merrick, 1988). In addition, the regulation of the *B. japonicum glnB* gene has been investigated very recently (Martin *et al.*, 1989). These authors have shown that the *B. japonicum glnB* gene was regulated by NtrC at tandem promoters, one of which had the $E\sigma^{54}$ consensus sequence.

The P_{II} protein is also involved in the transcriptional control of Ntr regulated promoters, including those of the *glnALG* operon (Fig. 1.5). Apparently, P_{II} communicates with

NR_{II} under nitrogen excess conditions, causing a rapid dephosphorylation of NR_I (Ninfa and Magasanik, 1986; Keener and Kustu, 1988) (the precise mechanism of dephosphorylation is unknown to date). The presence of P_{II} will thus result in the cessation of Ntr gene expression because NR_I must be phosphorylated by NR_{II} before it can act as a transcriptional activator of Ntr controlled genes (see below; Hunt and Magasanik, 1985; Ninfa and Magasanik, 1986). A reversal of these events occurs when cells are grown under nitrogen limiting conditions. When UTase is activated by the intracellular excess of α-ketoglutarate over glutamine and P_{II} is removed by conversion to P_{II}-UMP, NR_{II} will again phosphorylate NR_I, and the latter regains its ability to activate Ntr regulated genes.

1.4.3.3. Cumulative feedback inhibition of GS activity.

In *E. coli*, GS activity is subject to feedback inhibition by the products of glutamine metabolism. These metabolites all contain an amide group derived directly or indirectly from glutamine via highly branched pathways in which GS represents the first common enzyme. Such a feedback inhibition of GS is cumulative because a number of different metabolites, which generally are independent in their action, may decrease GS activity progressively provided that each inhibitor is present at a physiological concentration that produces only partial inhibition by itself (Woolfolk and Stadtman, 1964; Shapiro and Stadtman, 1970). This cumulative effect is readily explainable by the existence of separate non-interacting binding sites on the

GS enzyme for different feedback inhibitors (reviewed by Rhee *et al.*, 1985) and provides an effective means of regulation of GS activity. In addition, it appears that feedback inhibition of biosynthetic activity for adenylylated GS enzyme differs from that for the deadenylylated form (Stadtman and Ginsburg, 1974). This may be explained by the fact that GS has a dual function. When cells are grown in an ammonia-containing medium, GS is partially adenylylated and its primary function is the formation of glutamine. Under these conditions, GS is not required for ammonia assimilation (GDH becomes active) and consequently, is subject to feedback inhibition. However, when ammonia levels drop significantly, GS regains it superior role in ammonia assimilation, is deadenylylated and shows an altered susceptebility to feedback inhibition.

1.4.3.4. A novel sigma-factor essential for Ntr-regulation. Hunt and Magasanik (1985) and Hirschman *et al.* (1985) have identified the regulatory *glnF*(*ntrA*) gene product as a new sigma (σ^{54}) factor, specifically required for the transcription of nitrogen-regulated (*ntr*) promoters. These promoters do not contain canonical -35 and -10 sequences but instead have the consensus (-26) CTGGYAYR-N_4-TTGCA (-10) (Beynon *et al.*, 1983; Ausubel, 1984; Gussin *et al.*, 1986). The *ntrA* gene has been studied extensively in *E. coli* (Castaño and Bastarrachea, 1984), *K. pneumoniae* (Merrick and Stewart, 1985; Merrick and Gibbins, 1985), *Azetobacter vinelandii* (Toukdarian and Kennedy, 1986), *R. meliloti* (Ronson *et al.*,1987b), and *Pseudomonas putida*

(Köhler *et al.*, 1989). A *ntrA* gene from *Thiobacillus ferooxidans*, which was recently cloned and sequenced (Berger and Rawlings, personal communication), showed extensive amino acid homology with the *ntrA* genes from *K. pneumonia* (49%) and *R. meliloti* (44%) (Dave Berger, Ph.D. thesis, University of Cape Town). In addition, the molecular cloning of the *ntrA* gene from *Rhizobium* spp. NGR234 has been described (Stanley *et al.*, 1989), although no DNA sequence data were provided. An NtrA analogue, the product of the nifR4 gene, has been reported by Jones and Haselkorn (1989) in the nitrogen-fxing photosynthetic bacterium R. capsulatus, based on the comparison between the amino acid sequence of the nifR4 gene product and the amino acid sequence of NtrA from other nitrogen-fixing organisms.

The *ntrA* gene product regulates other loci involved in the utilization of non-preferred nitrogen sources, such as arginine, proline and histidine (Magasanik, 1982) and has been implicated in the transcription of genes encoding amino acid transport components (Schmitz *et al.*, 1988). In addition, *nif* genes from nitrogen-fixing bacteria such as *K. pneumoniae* (for review, see Gussin *et al.*, 1986) and *fix* genes of some root/stem nodule bacteria such as *R. meliloti* (Gubler and Hennecke, 1988), are regulated by NtrA.

In *E. coli*, NtrA has been implicated in the regulation of two anaerobically inducible enzymes, formate dehydrogenase (FDH_H) and hydrogenase isoenzyme 3 (Birkmann *et al.*, 1987), both part of the formate hydrogenlyase (FHL) pathway. Furthermore, it was shown that the gene encoding the selenopolypeptide (*fdhF*) of FDH_H has a *ntr/nif* consensus promoter which is dependent upon the presence of a functional *ntrA* gene product for the activation of transcription. Two genes, involved in hydrogenase activity in *E. coli*, have been cloned and sequenced (Stoker *et al.*, 1989). Their gene products, HydH and HydG, show considerable similarity in amino acid sequence with the two components of sensor/ activator pairs of cellular sensory systems. However, the physiological significance of these similarities has not yet been elucidated.

Recently, Ninfa *et al.* (1989) demonstrated that *E. coli* $E\sigma^{54}$ is able to recognize *Caulobacter crescentus flbG* and *flaN* flagellar gene promoters *in vitro*. These promoters carry highly conserved sequence elements at -24 and -12 that are very similar to the consensus sequence of *ntr* and *nif* gene promoters of the enteric bacteria and *Rhizobium* spp. Subsequently, it has been shown by Mullin and Newton (1989) that these conserved sequences at -12 and -24 in *flbG* and *flaN* are required for transcription of these genes *in vivo*. Similar promoter sequences have been found upstream from the xylene-catabolic genes on the TOL plasmid of *Pseudomonas putida* (Dixon, 1986), the gene encoding the C4-dicarboxylate transport protein (*dctA*) of *R. meliloti* (Ronson *et al.*, 1984; Ronson et al., 1987c; Jiang *et al.*, 1989), and the *mvaA* gene of *Pseudomonas mevalonii* (Wang *et al.*, 1989).

Kustu *et al.* (1989) have reviewed the physiological importance of σ^{54} in eubacterial cells comprehensively, and have suggested that this sigmafactor, although not essential for bacterial viability under all conditions, may represent a class on its own in which it resembles the most abundant sigma factor (σ^{70} and its homologs) rather than other alternative sigma factors because, unlike alternative σ factors, σ^{54} and its homologs seem to be required for the expression of a diverse set of genes in a variety of bacteria. In this regard, it is interesting that σ^{54} shows only limited amino acid sequence similarity to other sigma factors (Helmann and Chamberlain, 1988). Moreover, specific binding of $E\sigma^{54}$ to its promoter is not sufficient for open complex formation: initiation of transcription requires the action of a bacterial activator, which, in turn, can be subject to regulation in response to a distinct physiological signal (section 1.3.5; see further). For example, transcription of the *glnA* gene in enteric bacteria requires the activator protein NtrC. Similarly, transcription of *nif*-genes in a variety of bacteria is dependent on NifA, transcription of the *xylCAB* and *xylS* genes requires XylR, and expression of the *dctA* gene is DctD dependent.

1.4.3.5. The glnALG operon in enteric bacteria.

The various genes involved in the control of the *glnALG* operon in *E. coli* and related enterics are given in Table 1. 2 . The *E. coli glnA* operon was cloned and characterized (Backman *et al.*, 1981; Pahel *et al.*, 1982; MacNeil *et al.*, 1982a; Tuli *et al.*, 1982) and the DNA sequence of the entire region was subsequently determined (Miranda-Rios *et al.*, 1987). Other genes mentioned in Table 2.1 are scattered throughout the bacterial chromosome. The *glnALG* operon contains three promoters, *glnAp1* and *glnAp2*, located upstream of the *glnA* gene (Reitzer and Magasanik, 1985), and *glnLp*, situated within the *glnAL* intergenic region (Ueno-Nishio *et al.*, 1983; Ueno-Nishio *et al.*, 1984) (Fig. 1.6). *glnAp1* and *glnLp* are both $E\sigma^{70}$ promoters repressed by NR_I and serve to maintain basal levels of GS, whereas the Ntr dependent promoter *glnAp2* is activated by phosphorylated NR_I when nitrogen becomes limiting. Several mutations have been introduced in the *E. coli glnA* promoter region and these mutants were subsequently characterized by DNA sequence analysis (Osorio *et al.*, 1984; Leon *et al.*, 1985; Reitzer *et al.*, 1987). This led to a better understanding of the complex organisation and function of the two *glnA* promoters in *E. coli*.

Table 1.2. : The *glnALG* operon and related genes.

Gene	Product	Target of action (*)	Role
glnA	GS	-	sensor
glnG	NR_I (NtrC)	T	effector
glnL	NR_{II} (NtrB)	T	modulator
glnB	P_{II}	CA, T	signal
glnD	UTase	CA, T	transducer
glnE	ATase	CA	"
glnF	σ^{54} (NtrA)	T	sigmafactor

(*) CA, catalytic activity; T, transcription

An unusual feature of the *glnALG* regulatory region is the presence of five binding sites for NR_I, all upstream of *glnAp2* (Reitzer and Magasanik, 1986) and a NR_I binding site which overlaps the transcriptional start and the TATAAT sequence of *glnLp* (Reitzer and Magasanik, 1983; Rocha *et al.*, 1985). Two sites, NR-1 and NR-2, for which NR_I has a high affinity, have been identified at *glnAp1*. Site 1 overlaps the portion of the *glnAp1* promoter homologous to the canonical TTGACA sequence and site 2 covers the transcriptional start site (Reitzer and Magasanik, 1985). This abolishes expression from *glnAp1* due to binding of NR_I and, at the same time, activates the *glnAp2* promoter.

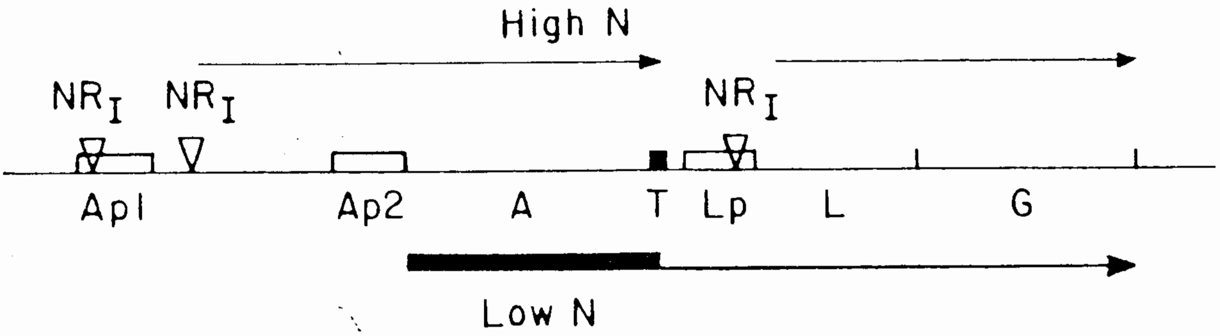


Fig. 1.6: Transcription of the *E. coli* *glnALG* operon in nitrogen-excess ("High N") and nitrogen-limiting ("Low N") conditions.▭ , promoters; ▼ NRI-binding sites; ■ terminator of transcription. The solid bar represents enhanced expression of the *E. coli* *glnA* gene under nitrogen-limiting conditions. From Magasanik (1988).

The NR_I site at *glnLp* did not appear to contribute to the activation of *glnAp2* (Reitzer and Magasanik, 1986), although it was shown recently that the NR_I binding site of the *glnL* promoter can replace completely the high affinity site 1 of the *glnA* promoter (Reitzer *et al.*, 1989). Most interesting is the fact that NR-1 and NR-2, which are originally located 140 and 110 bp upstream of *glnAp2*, respectively, resemble eukaryotic enhancer sequences in their ability to activate transcription irrespective of distance and orientation (Reitzer and Magasanik, 1986). Several bacterial systems, including the *lac*, *gal*, and *ara* regulons of *E. coli*, contain distant regulatory sites for regulation of gene expression (reviewed by Gralla, 1989). It has been suggested that gene regulation from a distance involves DNA looping, a sugestion strengthened by the observations of Ptashne *et al.* (1986) who have studied the action of purified repressor directly using electron microsopy. Reitzer *et al.* (1989) have recently studied the mechanism in which NR_I-phosphate interacts with the RNA-polymerase-σ^{54} complex in great detail using insertions of integral or half integral turns of the DNA-helix either between the two high affinity NR_I sites 1 and 2, or between the binding sites for NR_I-phosphate and RNA-polymerase, and presented evidence for a direct contact between NR_I-phosphate and the RNA-polymerase-σ^{54} complex in which DNA loop formation is involved. Similar experiments and results have been presented by Maeda *et al.* (1988) who have studied OmpR-mediated regulation of *ompC* gene expression in *E. coli* and by Buck *et al.* (1987) who have given evidence for DNA loop formation in a study of NifA-mediated activation of transcription at the *K. pneumonia nifH* promoter. In the latter case, the integration host factor (IHF) has been implicated in DNA bending between the NifA binding site and the *nifH* promoter, thereby facilitating contact between NifA and $E\sigma^{54}$ holoenzyme (Santero *et al.*, 1989). Although the IHF is mainly known for its involvement in site-directed recombination, plasmid replication, and transcription termination/anti-termination (Friedman, 1988), a more general function of the IHF has been suggested (Kur *et al.*, 1989). These authors also presented a consensus *ihf* binding site for the IHF, based on hydroxyl-radical foorprinting. Finally, Minchin *et al.* (1989) have presented evidence for DNA loop formation by studies on the NtrC-mediated activation of transcription at the *K. pneumoniae nifLA* promoter.

It has been shown that NtrA levels are independent of the nitrogen source (de Bruijn and Ausubel, 1983; Castaño and Bastarrachea, 1984), suggesting that the cells always contain a subpopulation of core polymerase bound with NtrA. This situation is reflected by the findings of Reitzer *et al.* (1987) and Sasse-Dwight and Gralla (1988) that the *E. coli* Ntr-regulated promoter *glnAp2* is occupied by $E\sigma^{54}$ irrespective of nitrogen availability. Transition from a closed to an open complex is only stimulated by NR_I-phosphate, which binds on either site of the DNA double helix at least 30 bp from the upstream boundary of the *glnAp2* promoter (Reitzer *et al.*, 1989) and which is not required for subsequent maintenance of open complexes or initiation of transcription (Popham *et al.*, 1989). Although it was shown earlier by Sasse-Dwight and Gralla (1988) that NR_I-phosphate binds a weak NR_I binding site at the opposite site of the DNA-helix, this is not a prerequisite for NR_I-phosphate bound to the high affinity NR_I binding sites 1 and 2 (Reitzer *et al.*, 1989).

GlnG mutants with greater ability to catalyze the isomerisation from closed to open $E\sigma^{54}$-*glnAp2* complexes have been identified recently (Weglenski *et al.*, 1989). It was shown that this increase was not caused by improved kinetics of the mutant because the rate of phosphorylation, phosphotase activity, and affinity for NR_I-sites or for the σ^{54}-*glnAp2* complex, was unaffected. Moreover, one of the mutants, NR_I316, was capable of transcription activation at the closed complex, whether phosphorylated or not. Most interesting, this mutation (Ser160 to Phe160) was mapped within the central domain of NR_I, which is homologous to the primary structure of central domains of NifA and DctD, both of which are transcription activators at NtrA-specific promoters (Drummond *et al.*, 1986; Gussin *et al.*, 1986; Ronson *et al.*, 1987c). It thus appeared that the central domain of NR_I was implicated in the specific binding of NR_I-phosphate to the closed NtrA-*glnAp2* complex. It has now been suggested that phosphorylation of NR_I would cause a conformational change within this domain which would enable NRI-phosphate to stimulate open complex formation more efficiently (Popham *et al.,* 1989; Weglenski *et al.,* 1989). The first direct indication that the central domain is essential for NtrC function was reported by Martin *et al.* (1988) who have shown that a truncated NtrC of *B. japonicum,* lacking its C- terminal domain, still remained functional as an activator of *glnII* expression. A recent report of Huala and Ausubel (1989) is consistent with the above findings. These authors showed that a truncated NifA protein, lacking both the N- and C-terminal domains, still had approximately 50% activity in stimulating transcription from the *nifH* promoter. And finally, it was shown that the *R. leguminosarum* NifA protein was fully functional as an activator despite the fact that the NH_2-terminal domain, normally present in all other documented NifA proteins, was missing (Lismaa and Watson, 1989).

In *B. subtilis,* there is no evidence for the presence of Ntr genes nor for post-translational modification by adenylylation (Deuel *et al.,* 1970; Fisher and Sonenshein,

1984) although feedback inhibition and possibly sulfhydryl modification may play roles in regulating catalytic activity (Deuel and Stadtman, 1970; Deuel and Prusiner, 1974). In addition, it has been shown that GS activity is regulated efficiently by the nitrogen source (Pan and Coote, 1979). During the last few years it has become evident that the *B. subtilis gln* locus has a genetic organisation which differs markedly from that of the enteric bacteria. There are strong indications that GS itself plays an important regulatory role in its own synthesis.

1.4.3.6. The *B. subtilis glnRA* operon.

The *B. subtilis* structural gene for GS , *glnA*, was isolated using the Charon 4A lambda vector (Ferrari *et al.*, 1981) and subsequential mapping and regulatory studies were performed (Gardner and Aronson, 1984; Fisher *et al.*, 1984). It became clear from these preliminary experiments that the regulation of GS activity in response to nitrogen levels was exerted through changes in *glnA* gene expression (Fisher *et al.*, 1984). At this stage, it was suggested that a region of approximately 500 bp upstream of the *glnA* gene played an important role in the regulation of *glnA* at the transcriptional level. This was confirmed by experiments of Schreier *et al.* (1985) and Schreier and Sonenshein (1986) who showed that GS itself was involved in the control of *glnA* transcription; *glnA::lacZ* fusions were negatively regulated only in the presence of the GS enzyme. Moreover, in contrast with the *glnALG* operon in enteric bacteria, downstream sequences (up to 130 bp downstream from the end of the *glnA* gene) were found not to be required for regulation.

Additional information was obtained when the entire coding region of the *B. subtilis gln* locus was sequenced (Strauch *et al.*, 1988; Nakano *et al.*, 1989). It was established that the *glnA* upstream region contained an open reading frame (ORF), termed *glnR*, which codes for a polypeptide with a M_r of 16.000. Striking amino acid similarities between GlnR and diverse regulatory proteins (Cro, TrpR) were observed, and a region with a helix-turn-helix structure indicated that the protein had DNA-binding properties . In support of this, it has been shown that GlnR is able to interact with the *glnRA* intergenic region (Brown and Sonenshein, 1989; cited in Schreier *et al.*, 1989). It has been proposed that GS interacts with GlnR in response to nitrogen excess conditions, forming a large complex, which would repress *glnA* expression (Schreier *et al.*, 1989). A second model includes the modification of GlnR by GS, causing it to bind to the *glnRA* intergenic region and repress transcription. Recently, Zhang *et al.* (1989) have investigated the exact nature of GS overproducing *glnA* mutants of *B. subtilis* which all mapped within a region associated with the active site of GS. This is in agreement with the earlier observation that these mutants had altered catalytic properties (Dean *et al.*, 1977). Regulatory properties of the *B. subtilis* GS is therefore likely to be, at least partially, dependent on its intrinsic catalytic parameters.

Nakano *et al.* (1989) have recently described the cloning of the *Bacillus cereus glnA* gene in *E. coli* and have determined the entire nucleotide sequence of the *glnA* coding region as well its upstream region. These authors detected a *glnA* upstream gene (ORF129) with considerable homology with the *B. subtilis glnR* gene. Although no regulatory studies have been undertaken to investigate the role of ORF129 in the regulation of the *B. cereus glnA* gene, it seems likely that the *B. cereus glnA* operon resembles the genetic organisation of the *B. subtilis glnRA* operon and that it might be regulated in a similar way.

1.5. Aim of this thesis.

C. acetobutylicum is an important industrial organism with great potential in the fermentation technology for the production of solvents from a variety of carbohydrates. Although *C. acetobutylicum* has been used for industrial solvent production during seven decades, genetic studies on this organism have only been initiated recently, mainly due to the advent of new recombinant DNA techniques, but also as a result of the renewed interest in the fermentative use of cheap substrates. Although much of the current interest centers around the complex multi-branched pathways of substrate utilization, it is also essential to study the genes involved in nitrogen metabolism, as it was earlier shown that nitrogen levels were important for solventogenesis and sporulation in *C. acetobutylicum*.

The expression of the cloned *C. acetobutylicum glnA* gene was studied in this work, not only to gain a better understanding of gene regulation in *C. acetobutylicum*, and in particular, the role of GS in nitrogen metabolism, but also to investigate wether the *C. acetobutylicum glnA* gene was regulated in a similar fashion as the *glnA* genes in the respective *E. coli glnALG* or *B. subtilis glnRA* operons. These two operons have served as model systems for the regulation of GS activity in bacterial cells for many years and are well documented. In this study it is shown that regulation of the *C. acetobutylicum glnA* gene differs markedly from the regulation of its counterparts in *E. coli* and *B. subtilis* because upstream sequences of the C. acetobutylicum *glnA* gene are able to modulate *glnA* gene expression without the action of ntr-genes (as in *E. coli*) or a *glnA* upstream regulatory gene *glnR* (as in *B. subtilis*). In addition, a short anti sense RNA, which is transcribed from the *glnA* downstream region and binds to the 5′ end of the *glnA* mRNA, is involved in the regulation of the *C. acetobutylicum glnA* gene in *E. coli*.

Chapter 2: Structure of the *C. acetobutylicum glnA* Gene and comparison with other GS Genes.

Chapter 2

Structure of the *C. acetobutylicum glnA* Gene and comparison with other GS Genes.

2.1. Summary

The nucleotide sequence of a 3.0 kb segment containing the *C. acetobutylicum glnA* gene was determined. Comparisons of the deduced amino acid sequence of the *C. acetobutylicum* GS with the amino acid sequence of GS of other organisms revealed some interesting features. A stretch of residues corresponding to the central loop, which protrudes into the central space of the GS dodecamer, was missing. Furthermore, the presumptive adenylylation site lacked a Tyr residue at the correct position and showed no homology with the adenylylation sites of the GS enzymes from other organisms. Five regions, which show a high degree of homology between the various GS enzymes, and are associated with the active site of the GS enzyme, were conserved in the *C. acetobutylicum* GS.

Two extended consensus promoter sequences (p_1 and p_2), characteristic of gram-positive bacteria, were identified in the *glnA* upstream region. The organisation of these promoters appeared to be complex in that p_1 was overlapped by a complex palindromic sequence (CPS) and p_2 showed features characteristic to promoters sensitive to the stringent response. A putative ribosome binding site, consisting of an ATG initiation codon and a SD sequence 5′-AAGGGGGA-3′, was situated eight bp upstream from the initiation codon.

The *C. acetobutylicum glnA* upstream region contained several regions of potential Z-DNA and various long stretches of adenine or thymine tracts which were similar to other nucleotide sequences involved in DNA curvature. Also present within this region were direct repeats displaying similarities with sequences present in the target DNAs of amino acid biosynthesis genes.

A long stretch of inverted repeat sequences, which have the potential to form alternative structures in the *glnA* mRNA, were found downstream of the *glnA* coding region. Further downstream of the *glnA* gene, a third putative extended promoter consensus sequence (p_3), oriented towards the *glnA* gene, was observed. It is proposed that transcription of p_3 would result in an antisense RNA,

capable of forming a RNA:RNA hybrid with the 5′-end of the *glnA* mRNA.

An open reading frame (ORF), showing homology in its amino terminal region with bacterial aspartokinases and homoserine dehydrogenases, was detected approximately 550 bp upstream from the *glnA* gene. This ORF, wich extends in the opposite direction of the *glnA* gene, contained a 5′-TTGGGGGC-3′ SD element eight bp from the ATG initiation codon, and was preceded by two putative promoter sequences.

2.2. Introduction.

As has been described in Chapter 1, GS plays a central role in bacterial nitrogen metabolism and its structural gene (*glnA*) has been extensively studied in various organisms (see Table 1.1). In enteric bacteria, GS is regulated at the level of enzyme activity by adenylylation-deadenylylation and feedback control and at the transcriptional level by a complex mechanism involving a number of gene products (Reitzer and Magasanik, 1987; Magasanik and Neidhardt, 1987). The *ntrC* (*glnG*) product, along with the *ntrA* (*glnF*) product, controls the *glnALG* operon. The *ntrA* and *ntrC* genes also regulate the expression of other operons that are involved in nitrogen metabolism. However, *B. subtilis* lacks an adenylylation-deadenylylation system and there is no evidence for global-regulatory *ntr* genes. The expression of the *B. subtilis* *glnA* gene appears to be controlled by a regulatory protein, encoded by the *glnR* gene (Schreier *et al.*, 1989). The regulation of GS activity in the enteric bacterium *E. coli* and the gram-positive aerobic endospore former *B. subtilis* is well documented (see Chapter 1) and both the *E. coli* *glnALG* and the *B. subtilis* *glnRA* operons serve as model systems for GS regulation in prokaryotes. In contrast, very little is known about the nature and regulation of the GS enzyme in the obligate anaerobic endospore former *C. acetobutylicum*, which produces solvents from a variety of carbohydrate substrates.

Several gene banks for the isolation of *C. acetobutylicum* genes in *E. coli* has been reported (Zappe *et al.*, 1986; Cary *et al.*, 1988; Verhasselt *et al.*, 1989) and the cloning of the *C. acetobutylicum glnA* gene has been reported previously (Usdin *et al.*, 1986). These authors have shown that the *glnA* gene was expressed and regulated by levels of nitrogen from its own regulatory region in *E. coli*. There was no evidence for a global regulatory *ntr* system, and the enzyme was not adenylylated. To investigate the molecular structure and regulation of the *C. acetobutylicum glnA* gene, the nucleotide sequence of the cloned *glnA* gene was determined (Janssen *et al.*, 1988).

2.3. Materials and methods.

2.3.1. Bacteria and plasmids.

Plasmids pHZ200 and pHZ202 (Usdin *et al.*, 1986) were used as the primary source of DNA carrying the coding region of the *C. acetobutylicum* GS. The general purpose cloning vector pLK34 (Appendix D; Botterman and Zabeau, 1987) was used in the construction of various subclones and deletions. Plasmid vectors pUC8 (Vieira and Messing, 1982), pUC13 (Messing, 1983), pUC18 (Norrander *et al.*, 1983), and pUC19 (Norrander *et al.*, 1983), were used for all subcloning for sequencing except a 330 bp *Taq*I fragment and a 545 bp *Alu*I fragment wich were cloned into the *Acc*I and *Sma*I sites of M13mp18 (Messing *et al.*, 1983; Yanisch-Perron *et al.*, 1985), respectively. In principle, wildtype *E. coli* strain K514 (*thr-1*, *leu-6*, *thi-1*, *supE*44, *lacY*1, *tonA*21, *hsr*, *hsm*) (Zabeau and Stanley, 1982) was utilized for all constructions. *E. coli* strain LK111 ($lacI^-$, *lacZ*ΔM15, $lacY^+$) (Zabeau and Stanley, 1982) was used as a host for pUC plasmids and *E. coli* JM105 (Δ*lac-pro*, *thi*, *strA*, *endA*, *sbc*15, *hsdR*4, F′ *traD*36, *proAB*, $lacI^q$*Z*ΔM15) (Yannisch-Perron *et al.*, 1985) when M13 vectors were utilized. When analysis of the GlnA phenotype was required, *E. coli* strains YMC11(Δ*glnALG*) (Backman *et al.*, 1981) or ET8051 (Δ[*rha-glnA*] *hutC*) (Pahel and Tyler, 1970) were used as host. All plasmids and strains used in this study are listed in Appendix A. Vector restriction maps of relevant plasmids are listed in Appendix D.

2.3.2. Media and buffers.

All media and buffers not described in the text are listed in Appendix B.

2.3.3. Growth conditions.

E. coli LK111 harbouring pUC derived recombinant plasmids were plated on 2 x Yeast/Tryptone (YT) agar containing isopropyl-β-D-thio-galactopyranoside (IPTG), 5-bromo-4-chloro-3-indolyl-β-galactoside (X-gal), and ampicillin (Ap) (Appendix B). *E. coli* JM105 was maintained on Cold Spring Harbor (CSH) minimal A medium or M9 medium (Appendix B). All strains were grown on either 2x YT or on

Luria-Bertani (LB) medium at 37°C for the preparation of plasmid (pLK34 and pUC vectors and derivatives) or replicative form (RF) (M13 vectors and derivatives) DNA. To analyse plasmids for their GlnA phenotype, CSH minimal medium, supplemented with 8 mM $(NH_4)_2SO_4$, was used. This medium was also used as a biological selection for certain constructs.

2.3.4. Plasmid constructions.

Plasmid pHZ202 was a EcoRI deletion derivative of pHZ200 (Usdin *et al.*, 1986) (Fig. 2.1). pJP1, pJP2, and pJP-S are *Xba*I, *Hind*II, and *Sst*I deletions of pHZ200 with the $GlnA^+$, $GlnA^-$, and $GlnA^+$ phenotype, respectively. pGln1300 (formerly called pJP3; Janssen *et al.*, 1988) was constructed by subcloning the 3.2-kilobase (kb) *Xba*I-*Bgl*II fragment containig the entire *C. acetobutylicum glnA* gene into plasmid pLK34, linearized with *Xba*I and *Bam*HI restriction endonucleases (Fig. 2.1). Plasmids used for sequencing are outlined in the sequencing strategy (Fig 2.2). Plasmid pGlnUR was constructed by the subcloning of the 1300 bp SmaI-EcoV fragment of pGln1300 into the *SmaI*-site of pUC18. After linearisation of pGlnUR with either *EcoRI* or *PstI*, undirectional shortening of insert DNA (from either side) was accomplished with *Bal31* endonuclease (Appendix C) as was described by Misra *et al.* (1985). All constructs were verified by restriction endonuclease digestion of recombinant DNA and subsequent electrophoresis in 0.8 to 1.2% agarose Tris-acetate gels (Appendix C).

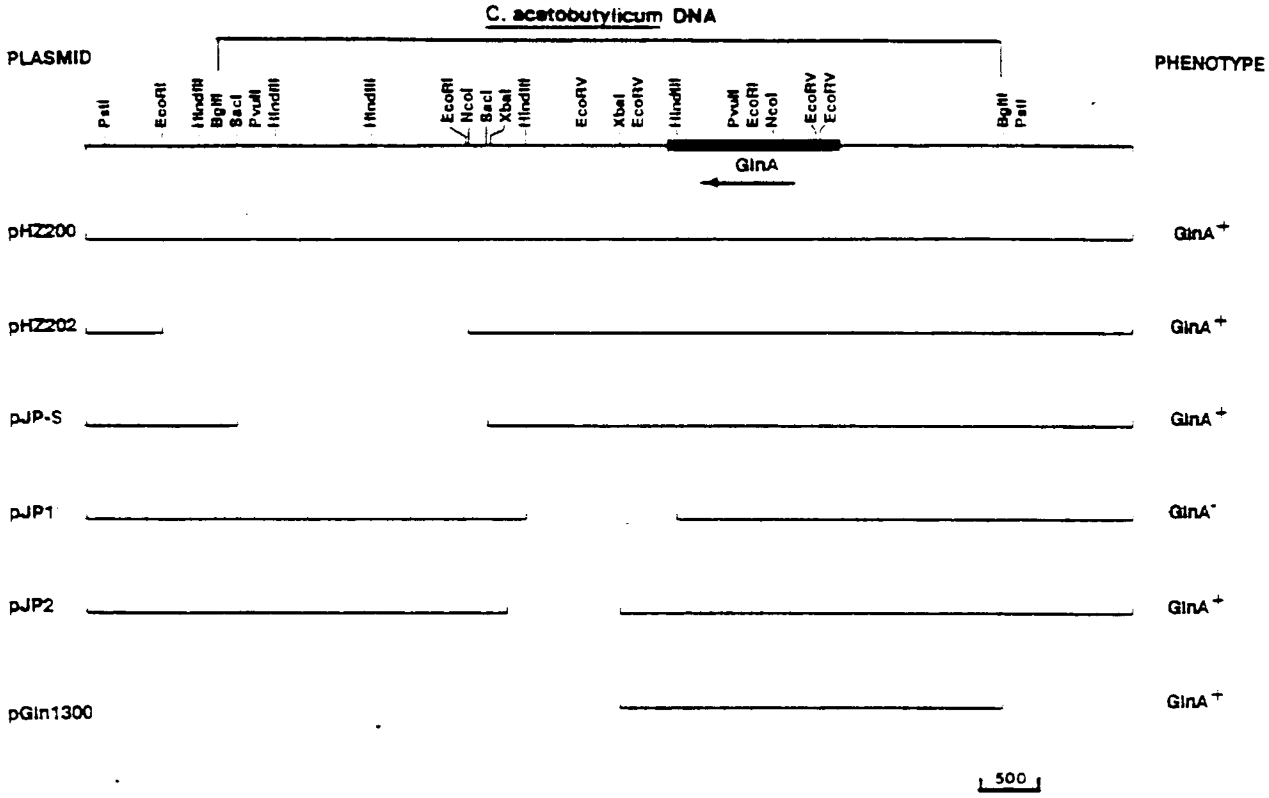


Fig. 2.1.: Restriction map of the original plasmid pHZ200 (Usdin *et al.*, 1986) containing the *C. acetobutylicum* *glnA* gene and approximately 1.3 and 3.5 kb up- and downstream regions, respectively. Various deletion plasmids and subclones, constructed in order to locate the *glnA* gene, are presented.

2.3.5. Transformation of *E. coli* strains K514, LK111, and JM105.

Competent *E. coli* cells were prepared and transformed as described in Appendix C. When transformed *E. coli* JM105 cells were plated, 100 μL of transformed cells were combined with 200μL exponentional phase JM105 indicator cells, and 20 μL IPTG (100 mM in distilled H_2O), 40 μL X-gal (2%, v/w in dimethylformamide), and 3 mL of H top agar were added. This mixture was plated on H agar plates (Appendix C) and incubated overnight at 37°C. Exponential *E. coli* JM105 cells were infected with M13 phage originating from white plaques and incubated for 270 min at 37°C with aeration. The cells were collected by centrifugation and M13 RF DNA was prepared by the miniprep method (Appendix C). DNA restriction analysis and subsequent gel electrophoresis were used to confirm the correct construction of recombinant plasmids as described in section 2.3.4. The culture medium containing the recombinant M13 phage was retained and used to prepare SS DNA templates.

2.3.6. Sequencing strategy.

Restriction mapping, deletion analysis, and subcloning, showed that the GlnA$^+$ phenotype was located on a *Xba*I-*Bgl*II fragment which was subcloned from pHZ200 into pLK34, which resulted in plasmid pGln1300 (Fig. 2.1). Plasmids used for sequencing were obtained by cloning various restriction fragments from pJP3 into pUC or M13 vectors or were *Bal31*-digested pGlnUR derivatives (section 2.3.4). Every construct was verified by restriction endonuclease analysis. The entire *Xba*I-*Bgl*II fragment was sequenced in both directions using overlapping clones (Fig. 2.2). Parts of this work has been described in Janssen *et al.* (1988) and Janssen *et al.* (1989).

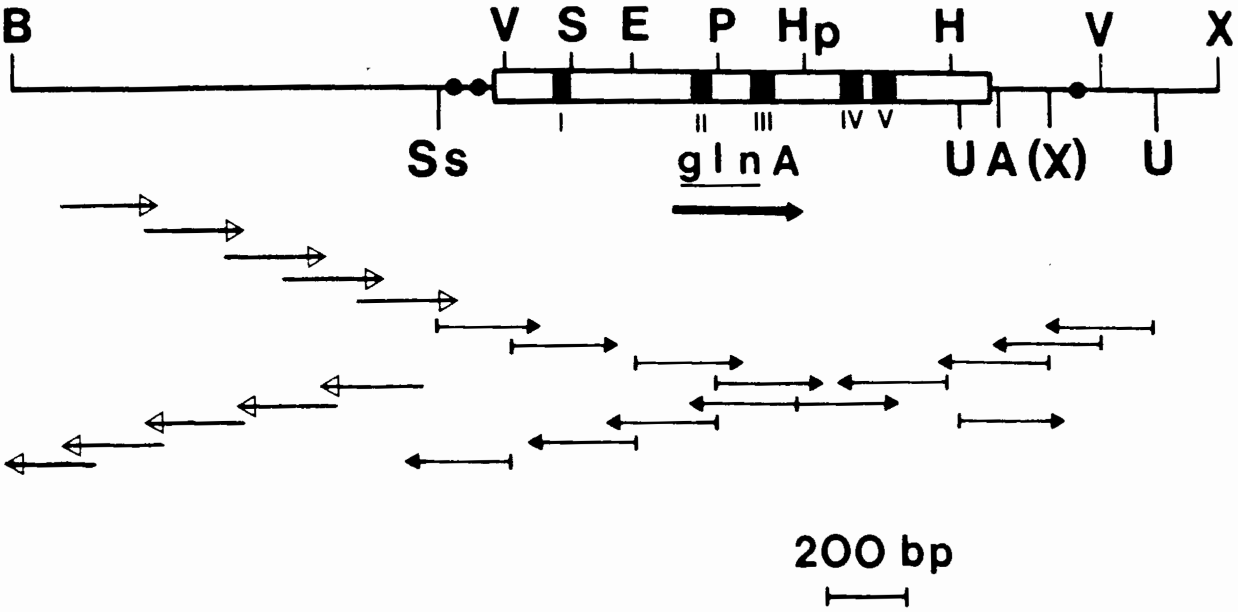


Fig. 2.2.: Sequencing strategy for the *C. acetobutylicum glnA* gene. The arrows represent templates and direction of sequencing. The length of each arrow corresponds with 220 bp of sequence, although routinely up to 300 bp of sequence was analysed per priming. Solid circles represent the positions of the three extended promoter sequences. The five regions (I, II, III, IV, and V) coding for regions of high amino acid homology between GS enzymes from procaryotes and eucaryotes are shown within the *glnA* gene. Restriction endonuclease abbrevations: B, *BglII*; V, *EcoRV*; S, *StyI*; E, *EcoRI*; P, *PvuII*; Hp, *HpaI*; H, *HindIII*; U, *Sau3AI*; A, *AluI*, X, *XbaI*; Tq, *TaqI*.

2.3.7. Preparation of template DNA.

Plasmid DNA from *E. coli* strains for the construction of subclones and subsequent nucleotide sequence analysis was isolated as described in Appendix C. SS M13mp DNA was isolated and prepared to serve as the SS template for sequencing reactions as described in the Bethesda Research Laboratory (BRL) manual supplied with their sequencing kit. SS template derived from pUC plasmids (or from pGln1300) was prepared as follows: pUC (or pGln1300) plasmid DNA (4-8 μg in 20 μL deionized H_2O) was denatured by adding 2 μL of NaOH (2 M) and leaving at room temperature for 5 min. The reaction was neutralised by adding 5 μL of sodium acetate (3 M), diluted with 25 μL H_2O and the DNA precipitated by adding 150 μL cold (-70°C) ethanol and holding at -70°C for at least 5 min. The DNA was collected by centrifugation (30 min, 10.000 rpm, 4°C) in a microfuge. The DNA pellet was washed with 70% cold (-20°C) ethanol, air dried and resuspended in 7 μL of H_2O. At this stage, SS DNA template was stored at -20°C for up to 24 h, or used directly for the priming annealing reaction.

2.3.8. Primers and primer annealing reaction.

Primers for reverse and forward sequencing on both pUC and M13 SS templates were obtained from BRL. Two unique primers were a gift from Prof. D. Botes (Dept. of Biochemistry, University of Cape Town) and were synthesised in an Applied Biosystems 381A DNA synthesiser (Forster City, CA, USA) using their proprietary reagents and solvents. Primer#1 had the sequence 5'-TAC.GCC.GTT.TTC.TTT.TAC-3', complementary with nt 37 to 54 within the *glnA* coding region (Fig. 2.3.). Primer#2 had the sequence 5'-TTT.AGT.TAA.GGC.GGG.GAT-3', complementary with a region 178 to 195 bp downstream of the *glnA* gene (Fig. 2.3). Primers for pUC plasmids or for pGln1300 were annealed to template SS DNA by adding 2 μL Sequenase™ buffer (United States Biochemistry [USB] Corporation, Sequenase™ kit version 2.0) and 2 μL primer (2 to 4 ng/μL) to the 7 μL of SS DNA, and incubating for 30 min at 40°C. The primer-template annealing-mix was allowed to cool down to room temperature (RT) and was used directly for sequencing reactions.

2.3.9. Sequencing reactions.

DNA sequences were determined using the dideoxy chain termination procedure (Sanger *et al.*, 1977). Reagents for the DNA sequencing reactions were obtained in kit form (BRL kit M13 C/SS I or USB Sequenase™ kit 2.0) and used according to the manufacturer's specifications. To avoid secondary structures and contractions in the template, dimethylsulfoxamide (DMSO) was added according to the method of Winship (1989). The DNA chain was labelled with [α^{35}S]dATP (> 37 TBq/mmole) obtained from Amersham Int., UK (Amersham product SJ1304).

2.3.10. Polyacrylamide gel electrophoresis.

The sequencing reactions were analyzed on denaturing polyacrylamide gels. Polyacrylamide (8%) (Appendix C) gels (0.4 mm spacers) were cast in a BRL apparatus (BRL model SO) (34 x 40 cm). A sharktooth comb was used to form the wells (BRL 1045 SC, 24 teeth, 0.4 mm thick, 6 mm point to point). The composition of the gels and the running conditions were as described in the manual supplied with the kit. All glass plates were silanized (2% dichloro-dimethyl-silane / 98% chloroform), rinsed with distilled H_2O and cleaned with ethanol (96%) prior to use. After electrophoresis, gels were dried onto filter paper (Whatman No. 3) using a Dual

Temperature Slab Gel Dryer (Model 1125B) (Hoefer Scientific Instruments, San Fransisco, USA).

2.3.11. Autoradiography.

Gels containing [^{35}S]-labelled DNA were placed under Cronex 4 film (Röntgen autoradiographic film) in an X-ray cassette and exposed for 24-76 h. The autoradiographs were developed using Kodak GBX X-ray developer and fixer.

2.3.12. DNA sequence analysis.

Manipulation of DNA sequences was done using an IBM XT computer and the MicroGenie^{TM} (Beckman version 299) software package developed by Queen and Korn (1984). Single letter codes were used for amino acids. A table of three- and one-letter codes for amino acids is given in Appendix F.

2.4. Results and discussion.

2.4.1. Nucleotide sequence of *glnA* gene.

The GlnA$^+$ phenotype was located on a 3.2 kb XbaI-BglII fragment (Fig. 2.1.) and the entire nucleotide sequence was determined (Appendix E). The DNA sequence contained an ORF which from the presumptive start codon (ATG) to the stop codon (TAA) contained 1332 nucleotides encoding 444 amino acid residues (Fig. 2.3.). The M_r of the predicted polypeptide coded for by this region was 49,630. This is lower than the apparant M_r of 59,000 estimated by gel electrophoresis of the *C. acetobutylicum* GS polypeptide encoded by plasmid pHZ200 (Usdin *et al.*, 1986).

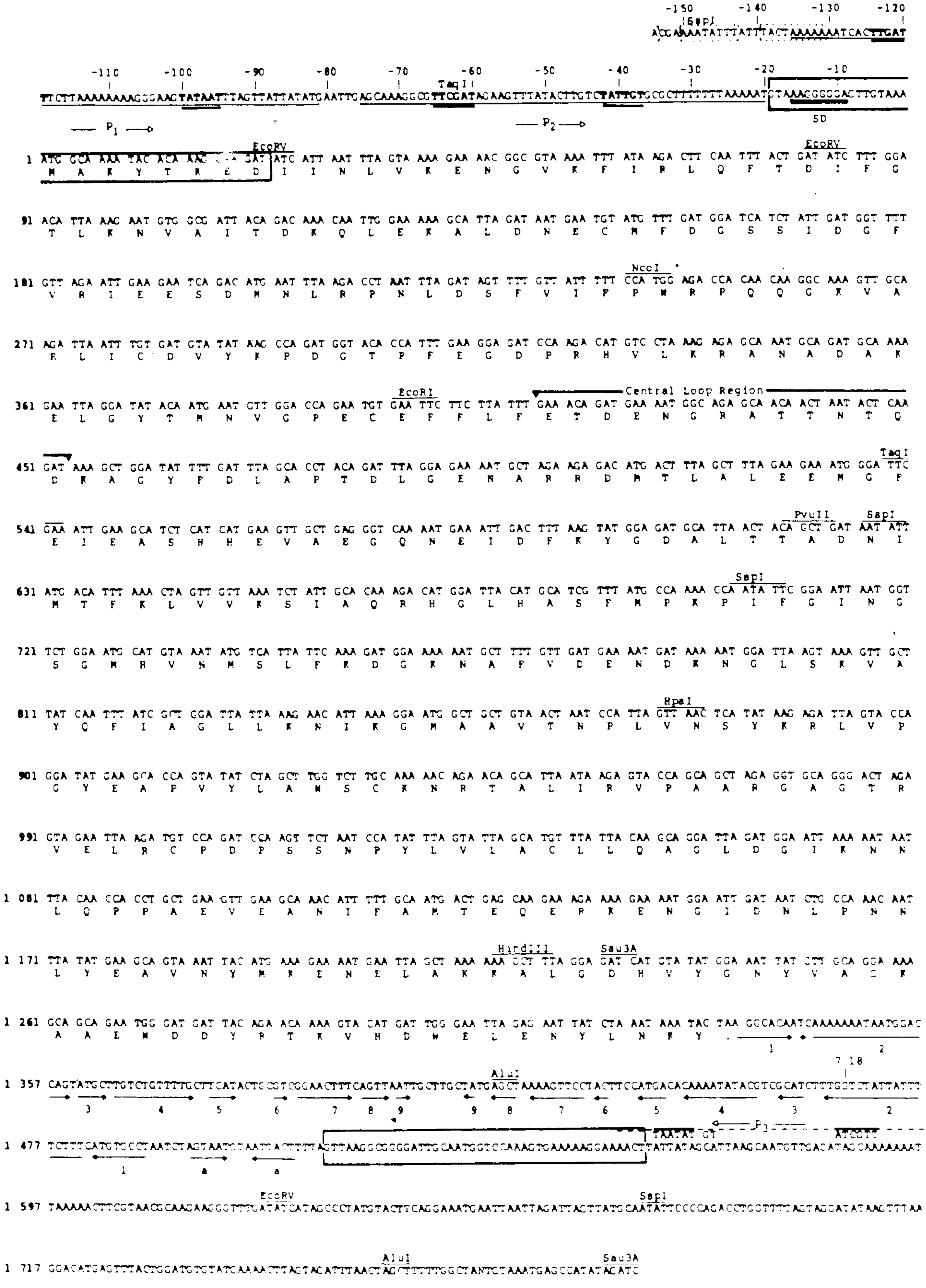


Fig. 2.3.: Complete sequence of the *glnA* gene from *C. acetobutylicum* with up- and downstream regions. The derived

amino acid sequence is given in the one-letter code (Appendix F) from positions 1 to 1332 (444 residues). Sense promoters p_1 and p_2 and antisense promoter p_3 are underlined. Classical -10 and -35 regions are in boldface type and underlined twice. The putative Shine-Dalgarno (SD) sequence AAGGGGGA starts at position -9. The region forming the central loop in the GS polypeptide (Almassy *et al.*, 1986; Yamashita *et al.*, 1989) (residues 137 to 152) is indicated. Inverted repeat sequences are indicated by arrows 1 to 9 and a. Boxed sequences (-19 to +24 and +1516 to +1559) show homology. Anti sense transcription of region 1516 to 1559 could lead to an RNA transcript able to bind the 5′ end of the glnA mRNA. Deletion 7.18 is indicated at position 1466.

The (G + C) content of the nucleotides within the ORF (31.8%) was higher than the (G + C) content of the nucleotides in the regulatory regions immediately upstream (25%) or downstream (21%) of the *glnA* gene. The average (G +C) ratio for the whole genome of *C. acetobutylicum* is approximately 28% (Cummings and Johnson, 1971). Although the *glnA* gene was the first *C. acetobutylicum* gene cloned and sequenced (Usdin *et al.*, 1986; Janssen *et al.*, 1988), our laboratory has determined the nt sequence of four other *C. acetobutylicum* genes: an endo-β-1,4-glucanase gene (*eng*) (Zappe *et al.*, 1988), a butanol dehydrogenase gene (*adh1*) (Youngleson *et al.*, 1989a), a β-hydroxybutyryl-CoA dehydrogenase gene (*hbd*) (Youngleson *et al.*, 1989b) and a xylanase gene (*xyn*) (H. Zappe, personal communication), with (G+C) contents of 33.8, 32.6, 34.0, and 37.7%, respectively.

2.4.1.1. Codon usage. The codon usage of the five *C. acetobutylicum* genes, *glnA*, *eng*, *adh1*, *bdh*, and *xyn*, was analyzed and all five genes showed a strong bias towards the use of codons in which A and U predominate :

Table 2.1.: Comparison of (G + C)-content and codon usage in *C. acetobutylicum* genes and *E. coli* genes.

aa	(Codon)[a,b]	*C. acetobutylicum* genes(mol%)[c]					*C.aceto*	*E.coli*[d]
		glnA	*eng*	*xyn*	*adh1*	*hbd*[e]	(mean%)	(mol%)
Ala	(GCU)	3.4	4.7	1.2	4.6	7.5	4.26	3.11
	(GCC)	0.0	0.5	0.4	0.3	0.0	0.22	2.09
	(GCA)	5.4	1.1	3.1	3.1	3.2	3.23	2.55
	(GCG)	0.2	0.0	0.4	1.0	0.0	0.33	3.40
Cys	(UGU)	1.1	0.5	0.8	1.0	0.7	0.82	0.32
	(UGC)	0.2	0.5	0.4	0.3	0.4	0.34	0.45
Asp	(GAU)	5.9	4.7	1.9	4.6	5.3	4.49	2.69
	(GAC)	0.9	1.2	1.5	0.5	0.4	0.89	2.58
Glu	(GAA)	7.0	2.6	2.7	7.7	7.8	5.56	5.04
	(GAG)	0.7	0.5	0.8	0.5	0.0	0.49	1.83
Phe	(UUU)	3.6	2.9	3.1	3.9	2.5	3.17	1.60
	(UUC)	1.1	0.2	0.8	1.0	1.4	0.92	2.08
Gly	(GGU)	1.1	1.9	1.9	2.1	1.4	1.68	3.55
	(GGC)	0.7	1.4	3.5	0.8	0.4	1.33	3.03
	(**GGA**)*	**5.2**	**3.5**	**3.5**	**3.6**	**7.1**	**4.57**	**0.34**
	(GGG)*	0.2	0.0	0.4	0.3	0.0	0.17	0.51
His	(CAU)	1.8	1.2	0.4	2.6	0.4	1.13	0.65
	(CAC)	0.0	0.2	0.4	0.3	0.4	0.24	1.02
Ile	(AUU)	3.4	1.9	2.7	4.1	5.7	3.55	2.32
	(AUC)	0.7	0.2	1.2	0.5	2.8	1.08	3.88
	(**AUA**)*	0.7	**5.4**	**2.3**	**2.3**	**1.4**	**2.42**	**0.03**
Met	(AUG)	2.9	2.6	2.3	4.1	3.9	3.17	2.91
Lys	(AAA)	5.6	4.5	5.0	5.2	6.0	5.25	4.55
	(AAG)	1.6	1.9	1.9	3.6	2.5	2.29	1.38
Leu	(**UUA**)*	**7.2**	**2.6**	**2.3**	**5.9**	**4.3**	**4.46**	**0.55**
	(UUG)*	0.2	0.5	1.2	0.3	0.4	0.49	0.78
	(CUU)*	0.2	0.9	1.5	0.5	1.8	0.98	0.83
	(CUC)*	0.0	0.5	0.0	0.0	0.0	0.10	0.63
	(CUA)*	0.9	0.0	0.0	1.6	0.4	0.56	0.17
	(CUG)	0.2	0.2	0.0	0.0	0.0	0.08	6.64
Asn	(**AAU**)*	**6.6**	**6.1**	**5.0**	**3.4**	1.4	**4.48**	**0.88**
	(AAC)	1.4	1.0	2.7	0.3	1.4	1.33	2.75
Pro	(CCU)*	0.7	2.1	1.2	0.7	0.8	1.08	0.37
	(CCC)*	0.0	0.0	0.0	0.0	0.0	0.00	0.25
	(CCA)*	3.8	1.4	1.5	2.1	3.9	2.56	0.82
	(CCG)	0.0	0.0	0.0	0.3	0.0	0.06	2.68

Gln	(CAA)	2.5	2.6	1.5	2.3	2.1	2.21	1.15
	(CAG)	0.0	1.4	2.3	0.0	0.0	0.72	3.18
Arg	(CGU)	0.0	0.5	0.4	0.0	0.0	0.18	3.09
	(CGC)	0.0	0.0	0.4	0.0	0.0	0.08	1.86
	(CGA)*	0.0	0.0	0.8	0.0	0.0	0.16	0.12
	(CGG)*	0.0	0.0	0.0	0.0	0.0	0.00	0.17
	(**AGA**)*	**4.3**	**0.7**	**1.5**	**2.3**	**3.2**	**2.41**	**0.06**
	(AGG)*	0.0	0.0	0.0	0.0	0.4	0.08	0.02
Ser	(UCU)	1.4	3.1	0.4	0.5	0.4	1.13	1.32
	(UCC)	0.0	0.3	0.4	0.0	0.0	0.14	1.28
	(**UCA**)*	0.9	**4.2**	**2.3**	**3.4**	**4.3**	**3.00**	**0.42**
	(UCG)*	0.2	0.2	0.0	0.3	0.0	0.14	0.57
	(**AGU**)*	0.7	**2.8**	**5.0**	0.5	0.0	**1.80**	0.32
	(AGC)	0.0	0.9	1.9	0.0	1.1	0.78	1.08
Thr	(ACU)	1.8	6.1	3.8	2.1	2.8	3.33	1.17
	(ACC)	0.0	0.5	1.5	0.0	0.0	0.40	2.49
	(**ACA**)*	**2.7**	**4.7**	**4.6**	**4.1**	**1.8**	**3.58**	**0.29**
	(ACG)	0.0	0.0	1.5	0.0	0.0	0.31	0.97
Val	(GUU)	2.7	2.6	2.3	1.6	5.0	2.82	2.80
	(GUC)	0.2	0.5	1.2	0.3	0.0	0.42	0.95
	(GUA)	3.0	2.1	1.5	3.1	2.8	2.50	1.71
	(GUG)	0.2	0.0	0.4	0.5	0.0	0.23	2.00
Trp	(UGG)	0.9	2.4	2.7	0.5	0.4	1.36	0.74
Tyr	(UAU)	2.9	3.8	5.0	3.6	2.1	3.48	1.06
	(UAC)	0.9	1.4	1.2	0.3	0.4	0.81	1.55
G+C	(%)	31.8	33.8	37.7	32.6	34.0	34.0	51.0

a *C. acetobutylicum* preferred codons are underlined.
b *E. coli* rare codons are marked by an asterisk.
c Expressed as a molar percentage of codon usage.
d 52 E. coli proteins combined (Allf-Steinberger, 1984).
e *C. acetobutylicum* gene designations:
glnA : glutamine synthetase (Janssen *et al.*, 1988)
eng : endo-β-1,4- glucanase (Zappe *et al.*, 1988)
xyn : xylanase (H. Zappe, pers. comm.)
adh1 : butanoldehydrogenase (Youngleson *et al.*, 1989a)
bdh : β-hydroxybutyryl- CoA dehydrogenase (Youngleson *et al.*, 1989b)

Codons CCC and CCG, both coding for Pro, were not used in any of the five *C. acetobutylicum* genes, and a further eight codons (GGG, CAC, CGC, CUG, CCG, AGG, UCC, UCG) were not used more than once in any of these genes. In *E. coli*, the following codons are regarded as minor codons and are rarely utilized: GGA (Gly), UUA (Leu), AAU (Asn), CCA (Pro), AGA (Arg), UCA (Ser) and ACA (Thr) (Koningsberg and Godson, 1983; Allf-Steinberger, 1984). This codon bias can contribute to differences in gene expression (Chapter 1). In all five *C. acetobutylicum* genes these minor codons are utilized preferentially and at high frequencies in comparison to the other codons.

In the *glnA* gene, the AGA codon was used in 100% of the codons coding for Arg (Table 2.1), and UUA was used 82% of the codons used to code for Leu. Both Arg and Leu are amino acids encoded by six possible codons. The AGA codon for Arg is extremely rare in *E. coli* genes (0.06 mol%) (Table 2.1). It is remarkable that in spite of the utilization of a high percentage of *E. coli* rare codons, high levels of enzymes were produced from the above five *C. acetobutylicum* genes cloned in *E. coli* (Janssen *et al.*, 1988; Zappe *et al.*, 1988; Youngleson *et al.*, 1989 ; H. Zappe and J. S. Youngleson, personal communication).

<u>2.4.1.2. The central loop region.</u> Studies on the crystalized GS from *S. typhimurium* (Almassy *et al.*, 1986) indicated that each subunit contained a central loop between amino acid residues 154 and 177 which extends into the central space of the GS dodecamer (Fig. 2.4.a). This region is cleaved by several proteases with different specificities (Fig. 2.4.b) (Dautry-Varsat *et al.*, 1979; Monroe *et al.*, 1984; Monroe *et al.*, 1985).

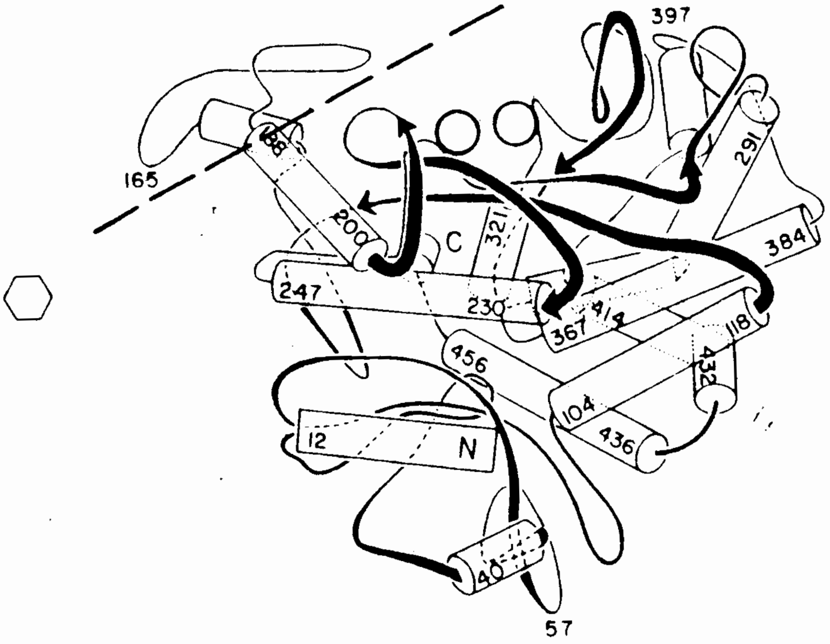


Fig. 2.4a.: A schematic drawing of the GS polypeptide chain (Almassy *et al.*, 1986). The 6-fold axis is indicated by a hexagon at the far left. The central loop at the upper left is not present in *C. acetobutylicum* GS.

```
                                  b c  d        d   b
                                  | |  |        |   |
Ec  FDDIRFGSSISGSHVAIDDIEGAWNSSTQYEGGNKGHRPAV--KGGYFPVPPVDSAQD
St  FDDIRFGASISGSHVAIDDIEGAWNSSTKYEGGNKGHRPGV--KGGYFPVPPVDSAQD
Va  FDDVKFATDMSGSFFKIDDVEAAWNTGSDYEEGNKGHRPGV--KGGYFPVAPVDSSQD
Rl  FDDVKYKADPYNTGFKLDSTELPSNDDTDYETGNLGHRPRV--KGGYFPVPPVDSAQD
Tf  FDSVTWNIDMSGCAYLVDAEEAAWNSGKEYEDGNMGHRLGV--KGGYFPVPPVDSAQD
Ab  FDDVKFKVEMNKVSYEFDSEEGPYTSDKDYEDGNLGHRPGV--KGGYFPVAPVDSGSD
An  FDSARFAQNANEGYYFLDSVEGAWNSGKEGTADKPNLAYKPRFKEGYFPVSPTDSFQD
Sc  FDSVRFATRENESFYHIDSEAGAWNTGAL-ED-NRE--YKVRYKGGYFPVPPVDHFAD
Bs  FKLDEKGEPTLELN---D-------------------------KGGYFDLAPTDLGEN
Bc  FKVDEKGNPTLELN---D-------------------------NGGYFDLAPMDLGEN
Ca  FETDENGRATTNTQ---D-------------------------KAGYFDLAPTDLGEN
```

Fig. 2.4b.: Comparison of the central loop region of GS enzymes from 11 different organisms. This region in *E. coli* GS is known to be sensitive to several proteases (b= chymotrypsin, c= *Staphylococcus aureus* protease, d= trypsin) (see text)

Proteolytic nicking led to a severe reduction (97%) of enzymatic activity (as measured by transferase and forward asssays) due to an apparent decrease in the binding affinity of nicked GS (GS^*) for glutamate and to a loss of the synergism between ATP and glutamate binding (Lei *et al.*, 1979; Monroe *et al.*, 1984). Conversely, binding of glutamate tends to protect the central loop of intact GS from proteolysis (Lei *et al.*, 1979), although one proteolysis site (Glu165) is 35 Å from the active site (Almassy *et al.*, 1986). This could indicate a long-range communication between the central loop and the active site. The nicked enzyme GS^* retained its overall dodecameric structure, was still recognized by antibodies to the intact enzyme, and was still adenylylated (Dautry-Varsat *et al.*, 1979). No digestion occurred beyond the Pro-Val-Pro-Pro portion of the large fragment (Fig. 2.4.b) with any of the proteases used (Monroe *et al.*, 1984). These authors have suggested that the combination of three prolines in four residues might 'kink' the peptide chain and might serve to bury the peptide chain at the end of the exposed region of sequence while holding the susceptible region for exposure to the various proteases. The *C. acetobutylicum* GS enzyme lacks 26 amino acids involved in the formation of the central loop in *S. typhimurium*. In addition, the PVPP motif, which might help to expose the central loop region to proteases, is abscent (Janssen *et al.*, 1988; this study). The only other GS enzymes known to lack this region and the PVPP motif, are the GSs from *B. subtilus* (Strauch *et al.*, 1988) and *B. cereus* (Nakano *et al.*, 1989), which show a high similarity (approximately 60%) (see section 2.4.1.4.) in their amino acid sequence with the *C. acetobutylicum* GS. Although there is no evidence for a physiological significance for the absence of these structural features in the *B. subtilus*, *B. cereus*, and *C. acetobutylicum* GS enzymes, it is striking that all three organisms exhibit increased protease activity associated with stage II of endospore formation.

2.4.1.3. The adenylylation site.

Amino acid sequences around the *E. coli* Tyr residue 401 were of interest because it has been shown to be the site of GS adenylylation (Shapiro and Stadtman, 1968; Heinrikson and Kingdon, 1971). The amino acid sequences surrounding the Tyr residue in GS enzymes from eleven organisms, including the *C. acetobutylicum* GS, is shown in Fig. 2.5.

The *C. acetobutylicum* GS is the only enzyme in which a Tyr residue is absent. The GS enzymes which are subject to adenylylation showed strong homology among adjacent amino acids (GS below the line in Fig. 2.5). Although the *B. subtilis* and *Anabaena* GS enzymes have a Tyr residue in the correct position, there was no evidence that the cloned Anabaena and *B. subtilis* GS enzymes were regulated by adenylylation in *E. coli* (Fischer *et al.*, 1981, Fischer *et al.*, 1984). Furthermore, the *S. coelicolor* GS enzyme is adenylylated in its natural host, but was not adenylylated in *E. coli* cells (Wray and Fisher, 1988). This may be due to conformational differences of the three-dimensional structure surrounding the Tyr residue. The cloned *C. acetobutylicum* GS is not subject to adenylylation (Usdin *et al.*, 1986) and not only lacks the Tyr residue but shows very little homology to the *E. coli* GS.

						AMP							Score[b]
Ca[a]	Val	Glu	Ala	**Asn**	Ile	Phe	Ala	Met	Thr	Glu	Gln	**Glu**	2/12
Bs	Ile	**Asp**	Arg	**Asn**	Ile	Tyr	Val	Met	Ser	Lys	**Glu**	**Glu**	5/12
Bc	Val	**Asp**	Arg	**Asn**	Ile	Tyr	Val	Met	Thr	Lys	**Glu**	**Glu**	5/12
An	Leu	**Asp**	**Lys**	**Asn**	Ile	Tyr	Glu	Leu	Ser	**Pro**	**Glu**	**Glu**	7/12
Sc	Ile	**Asp**	**Lys**	Asp	**Leu**	Tyr	Glu	**Leu**	Ala	**Pro**	**Glu**	**Glu**	8/12
Rl	**Met**	**Asp**	**Lys**	Asp	**Leu**	Tyr	**Asp**	**Leu**	**Pro**	**Pro**	Lys	**Glu**	10/12
Va	**Met**	**Asp**	**Lys**	Asp	**Leu**	Tyr	**Asp**	**Leu**	**Pro**	Ala	**Glu**	**Glu**	10/12
Tf	**Met**	**Asp**	**Lys**	**Asn**	**Leu**	Tyr	**Asp**	**Leu**	**Pro**	Ala	**Glu**	**Glu**	11/12
Ab	**Met**	**Asp**	**Lys**	**Asn**	**Leu**	Tyr	**Asp**	**Leu**	**Pro**	Ala	**Glu**	**Glu**	11/12
St	**Met**	**Asp**	**Lys**	**Asn**	**Leu**	Tyr	**Asp**	**Leu**	**Pro**	**Pro**	**Glu**	**Glu**	12/12
Ec	**Met**	**Asp**	**Lys**	**Asn**	**Leu**	Tyr	**Asp**	**Leu**	**Pro**	**Pro**	**Glu**	**Glu**	12/12

Fig. 2.5.: Amino acid sequences arround the tyrosine residue of various GS enzymes. (a) Ca, *C. acetobutylicum*; Bs, *B. subtilis*; Bc, *Bacillus cereus*; An, *Anabaena* sp. strain 7120; Sc, *S. coelicolor*; Tf, *T. ferrooxidans*, Ab, *A. brasilense*; St, *S. typhimurium*; Ec, *E. coli*; Va, *V. alginolyticus*; Rl, *R. leguminosarum*. (b) Homology score with the *E. coli* adenylylation site in matches. Residues in bold letters are identical to the *E. coli* residues; underlined residues have similar function. References are as presented in Table 1.1.

2.4.1.4. GS amino acid homology. In a comparison of eucaryotic and procaryotic GS enzymes, it was shown that, although the amino acid similarity between the enzymes was only approximately 15%, the major part of this similarity was located in five regions (Janssen *et al.*, 1988; Fig. 2.6.). These regions, which were conserved in the *C. acetobutylicum* GS, are all associated with the proposed GS active site (Janson *et al.*, 1986; Almassy *et al.*, 1986).

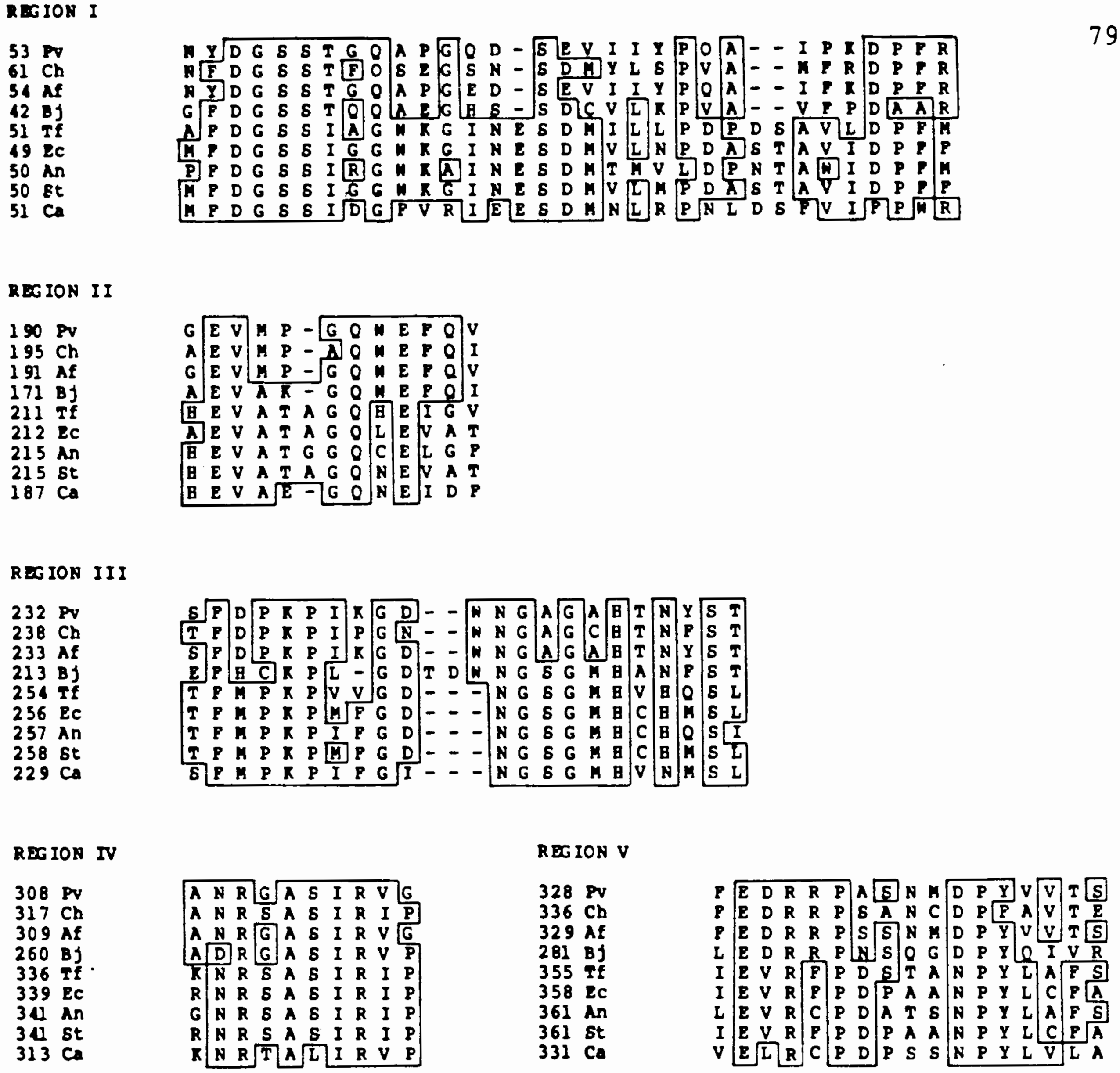


Fig. 2.6.: Comparison of amino acid sequences of five regions of homology of GS enzymes from *Phaseolis vulgaris* (Pv; Gebhardt *et al.*, 1986), Chinese hamster (Ch; Hayward *et al.*, 1986), alfalfa (Af; Tischer *et al.*, 1986) and *B. japonicum* (Bj; Carlson and Chelm, 1986) and various bacterial GS enzymes. Abbrevations are as in Fig. 2.5.

Regions II to V are β- strands closely associated with two Mn^{2+} cations of one subunit while region I contained the *S. typhimurium* Trp57 residue which is within 6 Å of the adenylylation site, Tyr397, of the adjacent GS subunit, and is thought to complete the active site. It is interesting that in all three gram-positive organisms, this Trp57 residue is replaced by a functionally similar phenylalanine. Region III is considered to be the ATP- binding site because its sequence resembles the amino acid sequence GXXGXGK (Fig. 2.6.), found in ATP-binding proteins (Walker *et al.*, 1982; Fry *et al.*, 1986). Region IV is thought to be the glutamate-binding site because of its similarity with the sequence DRGASIV of bovine and chicken GDH (positions 18 to 24). This sequence is close to to the Lys27 residue which is implicated in GDH substrate binding (Moon and Smith, 1973; Rasched *et al.*, 1974). Recently, Colanduoni *et al.* (1987) studied the kinetic mechanism of substrate binding of the *E. coli* GS enzyme and have associated a Lys residue in glutamate binding. Apparently, substrate binding requires an ordered mechanism in which ATP is bound first, glutamate second, and ammonia third. Each step requires a Lys residue with a specific pK_a value, depending on the state of proton transfer. Binding of ATP seems to involve Lys47 (in the *E. coli* GS) (Pinkofsky *et al.*, 1984) which is part of the Trp57 loop that lies across the active-site channel (Almassy *et al.*, 1986). The two other Lys residues involved in substrate binding have not been identified yet.

The *C. acetobutylicum* GS showed the highest amino acid sequence similarity (approximately 60%) with the GS enzymes from Gram-positive Bacillus species but showed less similarity with the GS from another Gram-positive organism, *S. coelicolor* (37%) (Table 2.2).

Table 2.2.: Percentage of similarity in amino acid sequence of eleven bacterial GS enzymes, based on the alignment given in Appendix H. These data were converted in a distance matrix and used for the construction of a phylogenetic three. Abbrevations are as in Fig. 2.5.

%	Ec	St	Va	Rl	Tf	Ab	An	Sc	Bs	Bc	Ca
Ec	-	96.0	77.6	64.2	62.5	61.9	54.2	51.8	42.3	40.0	38.7
St	96.0	-	77.9	64.7	63.1	62.9	55.8	52.5	42.7	40.2	39.1
Va	77.6	77.9	-	62.6	65.7	61.1	53.5	53.2	39.8	40.1	37.3
Rl	64.2	64.7	62.6	-	58.0	69.1	58.1	53.0	40.5	39.4	37.1
Tf	62.5	63.1	65.7	58.0	-	59.7	53.8	50.5	40.4	40.3	36.4
Ab	61.9	62.9	61.1	69.1	59.7	-	57.6	53.7	40.0	39.2	36.4
An	54.2	55.8	53.5	58.1	53.8	57.6	-	56.4	36.8	35.0	36.3
Sc	51.8	52.5	53.2	53.0	50.5	53.7	56.4	-	40.9	39.7	36.9
Bs	42.3	42.7	39.8	40.5	40.4	40.0	36.8	40.9	-	87.3	59.9
Bc	40.0	40.2	40.1	39.4	40.3	39.2	35.0	39.7	87.3	-	59.0
Ca	38.7	39.1	37.1	37.1	36.4	36.4	36.3	36.9	59.9	59.0	-

The percentage of similarity for all 11 GS enzymes is 21.8 %

Evolutionary relationships between the various bacterial enzymes were estimated using Felsenstein's PHYLIP (Phylogeny Inference Package) program (version 2.7), which constructs phylogenetic trees based on the distance matrix method of Fitch and Margoliash (1967).

The phylogenetic tree presented in Fig. 2.7, performed on GS sequences from 10 different organisms, is in close agreement with the results of Wray and Fisher (1988) who have placed *S. coelicolor* in the same phylum as Gram-negative organisms. These authors have also suggested that evolutionary restrictions imposed on GS in *B. subtilus*, *B.*

cereus and *C. acetobutylicum* might be different than for GS enzymes which do contain a central loop region and which are regulated by an adenylylation system.

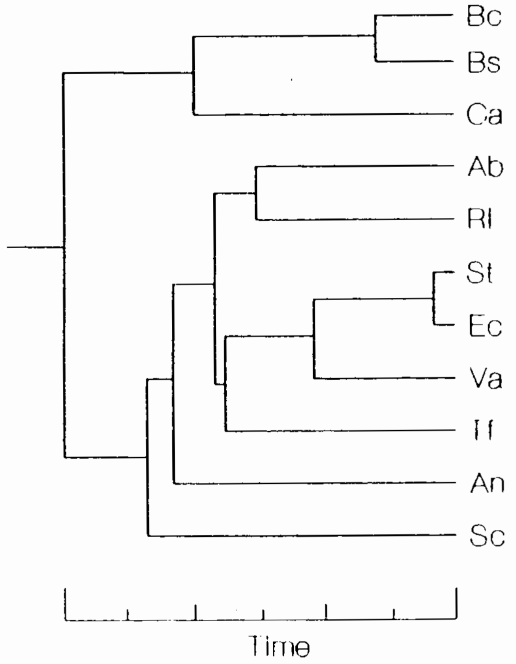


Fig. 2.7.: An estimate of evolutionary relationships between various bacterial GS enzymes. Abbrevations are as in Fig. 2.5 and references are as in Table 1.1.

2.4.2. Analysis of the regions surrounding the *glnA* coding region.

2.4.2.1. The ribosome binding site. The ORF of the *C. acetobutylicum glnA* gene was preceded by a putative ribosome binding site (RBS) consisting of a AAGGGGGA Shine and Dalgarno (SD) sequence, situated between the -9 and -16 nucleotides from the initiation codon AUG (Shine and Dalgarno, 1974). This sequence was identical to the SD sequence reported for the endoglucanase D gene of *C. thermocellum* (Jollif *et al.*, 1986) and the α-amylase gene of *Bacillus stearothermophilus* (Nakajima *et al.*, 1985). Ribosome binding sites of 27 clostridial genes, concerned with toxin production, substrate utilization and fermentative metabolism, have been summarized by Young *et al.* (1989). These authors have shown that the avarege free energy of SD-pairing in clostridial genes amounts to -16.7 kcal/mol. This is much higher than the -9 kcal/mol free energy for the prototype *E. coli* binding site AGGA, and is in agreement to the idea that SD elements of Gram-positive mRNAs are chraracteristically able to form highly stable complexes with 16 sRNA (Hager and Rabinowitz, 1985; see Chapter I). The 8 nt spacing between the SD sequence and the AUG initiation codon of the *C. acetobutylicum glnA* gene is very similar to the optimal spacing (7 ± 2 nt) deduced for *E. coli* RBS, which suggests that the *C. acetobutylicum glnA* RBS would function normally in *E. coli*.

Based on the comparison of RBS regions from 251 *E. coli* genes, an additional site of interaction between *E. coli* mRNA and 16 sRNA has been proposed by Petersen *et al.* (1988) who have suggested that 16 nt at the 5′ terminus of 16 sRNA might be involved in mRNA recognition. The complementary sequence in the mRNA falls within the first 24 nt coding for the genetic message and contains at least three consecutive nt from the string TCAAACTCTTCAATTT. The *C. acetobutylicum glnA* gene exhibits from nt +4 to +19 a similar sequence GCAAAATACACAAAGG. Duplication of the CAA sequence has been compared with the advantageous duplication of the -AGG- triplet in the SD sequence (Petersen *et al.*, 1988). This duplication is present in the extended RBS of the *C. acetobutylicum glnA* gene (CAAA and CAA, respectively).

An analysis of the known sequences of a large number of *E. coli* genes shows that the use of codons immediately after the AUG start codon deviates strongly from the overall codon usage. It is thought that the effect of the second codon is mainly exerted by its influence on the structural composition of the RBS; secondary structures in the RBS have been shown to decrease gene expression by shielding the SD element and the translation initiation codon (Looman *et al.*, 1986). The effect of the codon following the AUG triplet on the expression of a modified *lacZ* gene has been studied in *E. coli* (Looman *et al.*, 1987). A secondary structure in the RBS of the *C. acetobutylicum glnA* gene is not apparent and the codon in the second position (GCA) within the *glnA* gene does not belong to a group of triplets which are known to have a strong influence on gene expression.

2.4.2.2. *glnA* transcription initiation. The region upstream of the *glnA* gene was characterized by two putative promoter sequences (p_1 and p_2) containing classic -10 and -35 RNA polymerase binding consensus sequences equivalent to the *E. coli* $E\sigma^{70}$ and *B. subtilis* $E\sigma^{A}$ promoter sequences (Fig. 2.3.; Fig. 2.8). In addition, both putative promoters showed extended promoter consensus sequences involving the -5 and -45 regions which have been reported to be a feature of the promoter sequence (recognized by the major form of RNA polymerase) of gram-positive bacteria (Graves and

Rabinowitz, 1986). So far only a limited number of clostridial transcription initiation sites have been determined by mRNA transcript mapping (Graves and Rabinowitz, 1986; Beguin *et al.*, 1986). Determination of transcription initiation sites for the *C. acetobutylicum glnA* gene has been described recently (Janssen *et al.*, 1989) (this study, see Chapter 3). These results confirmed the above proposed *glnA* promoter sequences, both in *C. acetobutylicum* strain P262 as well as in *E. coli* strain YMC11 (Appendix A), in which the *C. acetobutylicum glnA* gene was cloned.

The expression of the *glnA* gene in enteric bacteria such as *E. coli, is* regulated by a set of nitrogen regulatory (*ntr*) genes. One of these genes, *ntrA (glnF/rpoN)*, codes for a sigma factor, σ^{54} wich recognizes a specific promoter sequence with consensus (C/T)TGGCA(C/T)N_5-TTGCA between positions -27 to -11 from the transcription initiation site (Chapter 1). These promoter sequences have also been found in a wide variety of organisms in the regulatory regions of genes concerned with physiological processes such as anaerobic metabolism, nitrogen fixation, pillin and flagellum production, and the degradation of organic compounds (reviewed by Kustu *et al.*, 1989) (see Chapter 1). No such promoter sequences, recognized by $E\sigma^{54}$, were detected in the upstream region of the *C. acetobutylicum glnA* gene (Fig. 2.8). Although it is apparant that the *C. acetobutylicum glnA* gene is not expressed from a σ^{54} promoter, it is unclear at present wether or not a global *ntr*-system exists in *C. acetobutylicum*.

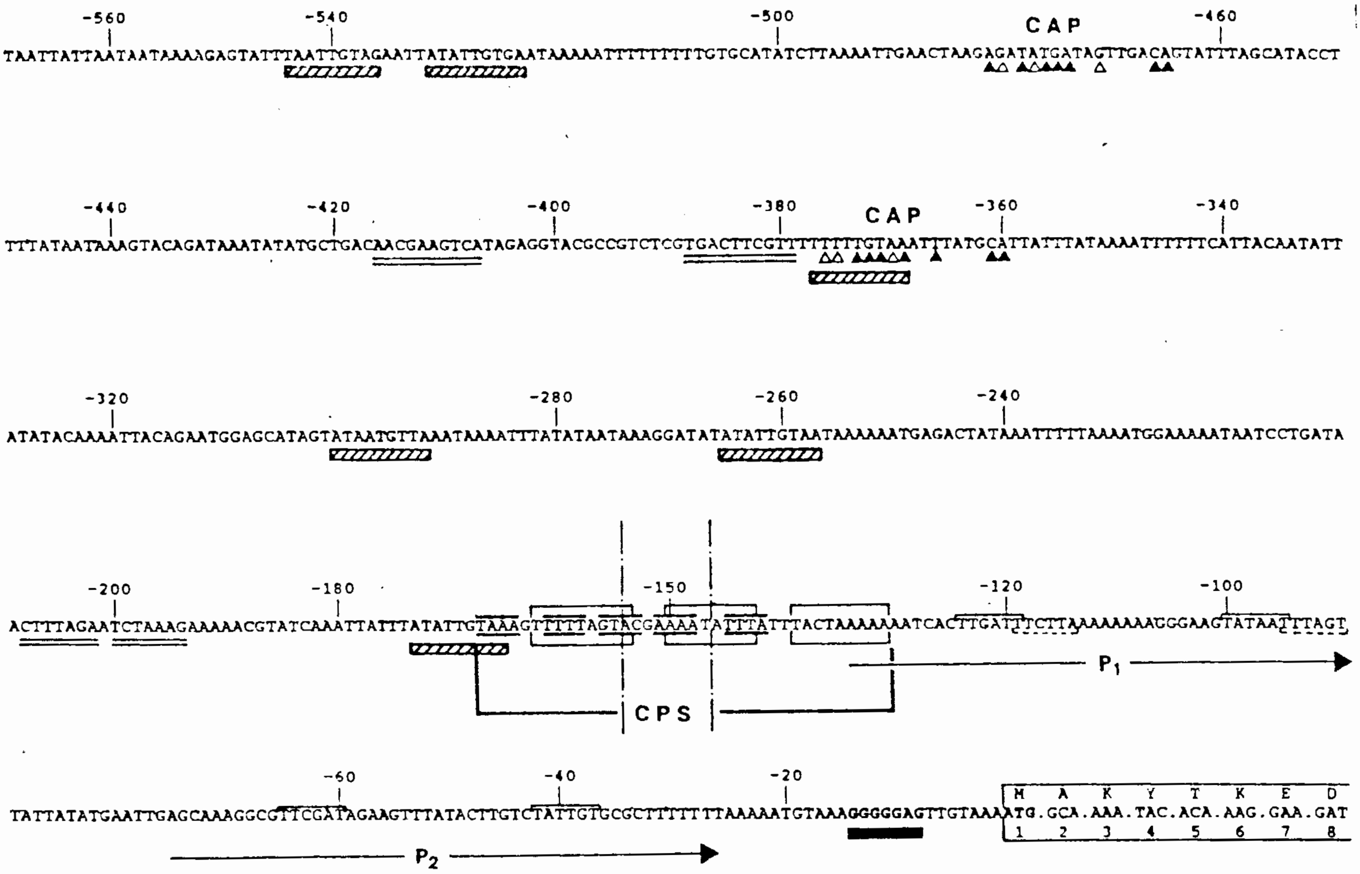


Fig. 2.8.: Nucleotide sequence of the upstream region of the **C. acetobutylicum glnA** gene. Putative catabolite activator protein (CAP) binding sites are underlined with triangles (▲,homology; △ , no homology to CAP consensus sequence. Perfect inverted repeats are underlined twice, the SD sequence is underlined with a solid bar, and promoters p_1 and p_2 are underlined with a full line. Direct repeats with consensus 5′-ATATTGTAA-3′ (see text) are shown with a dashed bar (▨) under the sequence. A complex palindromic sequence (CPS) shows two axes of symmetry.

<u>2.4.2.3. Regions of dyad symmetry and direct repeats.</u> Some aspects of transcription termination in prokaryotes have been described in Chapter 1. Inverted repeats followed by

oligo-T tracts have been detected downstream of the coding regions of several clostridial genes. For the *C. pasteurianum* ferredoxin gene (Graves and Rabinowitz, 1986) and the *C. thermocellum celA* gene (Béguin *et al.*, 1986), the 3′-termini of the corresponding mRNAs have been mapped within such dyad symmetry elements.

The immediate dowmstream region of the *C. acetobutylicum glnA* gene consisted of a 158-bp stretch of inverted repeat sequences (indicated in Fig. 2.3 by arrows 1 to 9). mRNA transcribed from this region would have the potential to form a number of stem-loop structures, including a long 59-bp stem-loop with a ΔG value (Salser, 1977) of -60.8 kcal/mol, followed by a short 6-bp stem-loop with a ΔG of -0.4 kcal/mol.

Other possible structures which could be formed include a shortened stem-loop of 38 bp with a ΔG value of -33.4 kcal/mol, followed by a 13-nucleotide U-rich region(UAUUAUUUUCUUU) (Fig. 2.9) or three shorter stem-and-loop structures with ΔG values of +3.4, -12.8 and -11.5 Kcal/mol, respectively. The nature of this downstream region suggests that these structures may be involved in the termination of transcription or alternatively, these structures may be important for mRNA stability (Chapter 1).

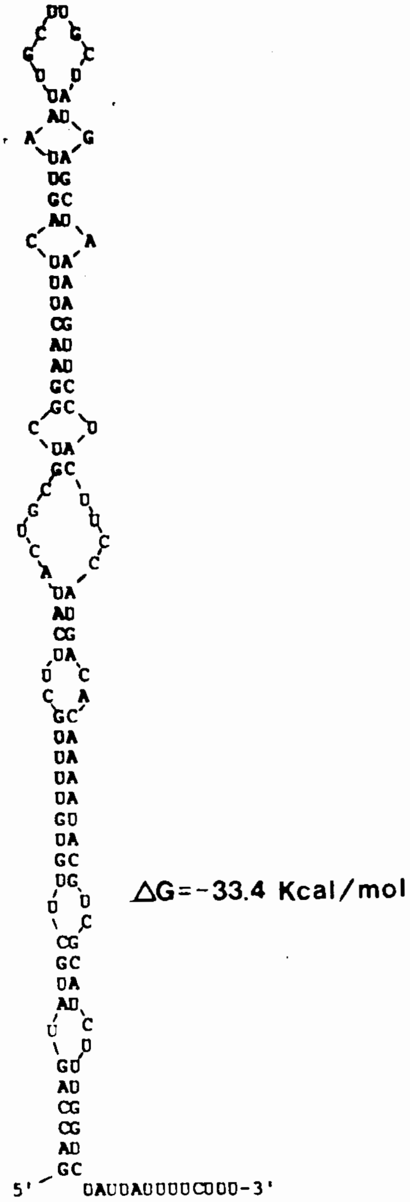


Fig. 2.9.: Inverted repeat sequences in the downstream region of the *C. acetobutylicum glnA* gene may result in secondary structures of the nascent *glnA* mRNA. One of these structures, which resembles a factor-independent terminator of transcription, is represented above (ΔG= -33.4 Kcal/mol).

Similar structures which resemble factor-independent transcription termination sites were detected in the *C. acetobutylicum glnA* upstream region (Fig. 2.8). One of these RNA structures, positioned on the DNA template between nucleotides -68 and -33, is of interest as this region overlaps the core sequence of the putative promoter p_2, including both the -10 and -35 regions. mRNA from this region transcribed from putative promoter p_1 would be capable of forming a stem-loop structure (ΔG= -11.0 kcal/mol), followed by a stretch of seven uridine residues, which suggests that it may act as a factor-independent transcription terminator (Fig.2.10a). The upstream regulatory region of the *B. subtilis trp* operon has a nucleotide sequence that generates a mRNA transcript with a similar secondary structure, wich acts as a transcriptional termination site in the presence of tryptophan (Shimotsu *et al.*, 1986).

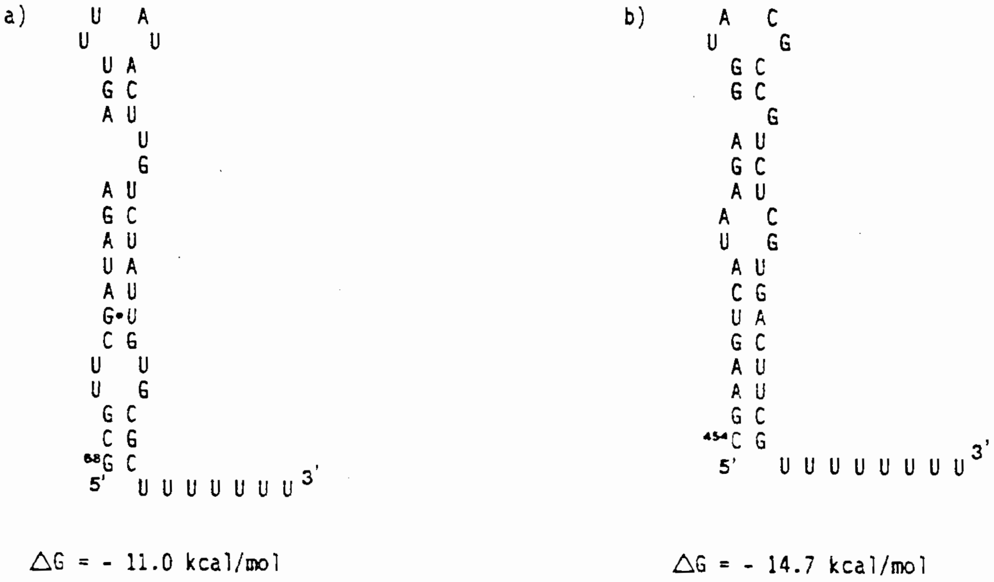


Fig. 2.10: Possible secondary structures present in the *glnA* upstream region: (a) Secondary structure which overlaps the *C. acetobutylicum glnA* promoter p_2 and is located at the 5'-end of *glnA* mRNA transcripts originated from the *C. acetobutylicum glnA* promoter p_1 ; (b) RNA transcribed from a region approximately 500 bp upstream of the *glnA* gene is capable of forming a similar secondary structure. Both structures resemble factor-independent terminators of transcription. The position within the *glnA* upstream sequence is indicated at the 5'-end.

Plaskon and Wartell (1987) have suggested that DNA curvature is associated with upstream activation of transcription and have investigated nucleotide distributions for forty-three prokaryotic promoter sequences. Analysis of a 390 bp upstream region of the *C. acetobutylicum glnA* gene (Fig. 2.8; nucleotides -150 to -540) showed an extremely high curvature score of 8.2. Only two *E. coli* rRNA P1 promoters, *rrnD* P1 and *rrnB* P1, exhibited higher curvature scores of 12.3 and 9.2, respectively. It is known that these promoters can be responsible for 36-50% of the total transcription initiated in *E. coli* (Lazzarini *et al.*, 1981). Several other promoters with significant curvature scores also had a high transcription initiation rate. Investigation of the A+T rich upstream region of the *spoVG* promoter (Plaskon and Wartell, 1987) revealed that this region, which is essential for efficient transcription by the σ^{32} and σ^{37} forms of the *B. subtilis* RNA polymerase, had a curvature score of 2.6, and the highly transcribed *hisR* promoter in *S. typhimurium* also had a correspondingly high curvature score of 2.8 (Bossi, 1983; Plaskon and Wartell, 1987). The involvement of DNA-looping in gene-activation has been furthermore proposed for the *K. pneumoniae nifH* promoter, activated by the *nifA* gene product (Buck *et al.*, 1987), and the *E. coli glnAp2* promoter, which is activated by NR_I (gp*ntrC*) (Reitzer and Magasanik, 1986). Recently, McAllister and Achberger (1989) performed insertional mutagenesis of the Alu156 promoter isolated from the *B. subtilis* bacteriophage SP82. This promoter is strongly dependent on the curved upstream DNA for efficient function.

In vivo and in *vitro* analysis of mutant promoters revealed that proper rotational orientation of the curved DNA was important. This is consistent with the idea that the curvature deflects the DNA helix back towards the promoter-bound RNA polymerase which then would allow the enzyme to interact directly with upstream DNA. Such interactions are proposed to favour structural changes in the DNA for open promoter complex formation (McAllister and Achberger, 1989). Furthermore, Gralla (1989) and Raibaud (1989) have recently reviewed some aspects of transcriptional activation including DNA looping

Within the same upstream region, several direct repeats were found with a high degree of similarity to a 9 bp sequence with consensus 5'-ATATTGTTT-3', present in various target DNAs of amino acid biosynthetic genes (Table 2.3). Such direct repeats have been proposed to be binding sites for regulatory proteins belonging to the *LysR* family (Kölling and Lother, 1985; Bohannon and Sonenshein, 1989). Although similar sequences have also been detected in the immediate upstream regions of the *E. coli*, *B. cereus* and *C. acetobutylicum glnA* genes and the *B. subtilis citB* gene (Table 2.3.), further studies involving DNA protection analysis and site directed mutagenesis are needed to establish the significance of these sites in the regulation of gene expression.

Table 2.3.: Direct repeat sequences found in the target DNAs of amino acid biosynthetic genes.

Sequence	A	B	location	organism*	reference+
ATATTGTAA	9	7	*glnA*	Ca	This study
ATATTGTAA	9	7	"	"	"
ATATTGTGA	8	7	"	"	"
ATAA**TGT**TA	7	7	"	"	"
T**T**T**TTGT**AA	7	5	"	"	"
TA**ATTGT**AG	6	5	"	"	"
GA**ATTGT**CA	6	5	*glnA*	Bc	(a)
ATG**TTG**CTT	5	7	"	"	"
AG**ATTG**CGA	6	5	"	"	"
TA**ATTG**AAA	6	4	"	"	"
T**TATTG**CAT	6	6	"	"	"
ATATTGGTG	6	7	*glnA*	Ec	(b)
T**TA**A**TGT**TT	5	8	"	"	"
ATATTGCAG	7	6	"	"	"
ATATTGTTT	7	9	*gltC/A*	Bs	(c)
ATATC**GT**TT	6	8	"	"	"
AG**ATT**CTTT	5	7	"	"	"
TTTTTTAAT					
T**TATTG**AAT	6	6	*asnC/A*	Ec	(d)
T**TATTG**CAT	6	6	"	"	"
AA**ATTG**ATT	5	7	*nodD/ABC*	Rt	(e)
AA**ATTG**ATT	5	7	*nodD/FE*	"	"
ATT**TTGT**TT	6	8	*citB*	Bs	(f)
T**TATTGT**GT	6	7	"	"	"
T**TATTGT**AT	7	7	"	"	"
AA**ATTGT**GA	6	6	"	"	"
AA**ATTGT**GA	6	6	"	"	"
AC**ATTG**ATA	6	5	"	"	"

Columns A and B are the similarity scores (out of 9) for the consensus sequences 5'-ATATTGTAA-3' (Janssen *et al.*, 1989) and 5'-ATATTGTTT-3' (Bohannon and Sonenshein, 1989), respectively.

(*) Abbrevations are used as follows: Ca, *C. acetobutylicum*; Bs, *B. subtilis*; Bc, *B. cereus*; Ec, *E. coli*; Rt, *Rhizobium trifolii*.

(+) References: (a), Nakano *et al.*, 1989; (b), Miranda-Rios *et al.*, 1987; (c),Bohannon and Sonenshein, 1989; (d), Kölling and Lother, 1985; (e), Schofield and Watson, 1986; (f), Dingman *et al.*, 1987a.

2.4.2.4. Putative antisense RNA. A sequence of 43 nt, which includes the RBS in addition to the first 24 nt of the glnA gene, showed considerable homology with a 44 nt sequence situated downstream of the *C. acetobutylicum glnA* gene (boxed in Fig. 2.3). Immediately in front of this 44-base downstream region, a third extended Gram-positive putative promoter sequence (p_3), oriented towards the 3'-end of the *glnA* gene, was detected (Fig. 2.3). Transcription from this putative promoter would result in an anti-sense RNA with the capability to form a strong RNA:RNA hybrid structure with the 5'-end of the *glnA* mRNA (see chapter 5). In those instances which have been analyzed so far, the production of antisense RNA has been reported to inhibit gene expression as a consequence of either transcriptional interference or specific RNA:RNA intermediates which effect translation efficiency or mRNA stability (reviewed in Chapter 1). Direct evidence for anti-sense RNA transcription from the *C. acetobutylicum glnA* downstream region has been established by primer extension reactions in both *E. coli and C. acetobutylicum* (Chapter 5). Furthermore, site-directed mutagenesis and mutagenesis by deletion was performed on this *glnA* downstream region. This led to interesting observations concerning the regulation of the *C. acetobutylicum glnA* gene in *E. coli* (Chapter 5).

2.4.2.5. Identification of an ORF showing homology with bacterial aspartokinases and homoserine dehydrogenases. An ORF with a suitable SD element (5'-TTGGGGGC-3'), eight nt from the presumptive startcodon AUG, was detected approximately 550 bp upstream of the *C. acetobutylicum glnA* gene. This ORF, oriented in the opposite direction from *glnA*, was preceded by two putative Gram-positive extended promoters, separated by a large hairpin structure (results not shown; see Appendix E). Although it was truncated at a *Sau3A* restiction site, originally used to isolate the *C. acetobutylicum glnA* gene (Zappe *et al.*, 1986), amino acid homology within its amino terminal region with various *E. coli* aspartokinase-homoserinedehydrogenases and one aspartokinase from *B. subtilus* was detected (Fig. 2.11).

C GlnX	**MN**T**I**I T**KFGG**S**SLAD**ANQFRK**VKD** **I**I YS**N**D**A**RKY **V**IP **SA**P**GK**RDS
E ThrA	**M** R**V**L **KFGGTSLA**NAERFL**RVAD** **IL** **E** **SN** **A**RQGQ**V**AT**VLSA**PA**K**I**TN**
E MetL	LH**KFGG**S**SLAD**VKCYL**RVA**G **I**MA**E**Y**S**QPDDMM**V**V **SA**AGST**TN**
E LysC	**M**SE**IVV**S**KFGGTS**V**AD**FDAMN**RSAD** **I**V L**S** D**AN**VRL**V** **VLSA**SAGI**TN**
B AKII	**M**GL**IV** Q**KFGGTS**VGSVEKIQNA**A**NRA**I** A**E**K Q KGH**VVVVV****SA**M**GK**STD

Fig. 2.11 : Comparison of the amino terminal region of the GlnX from *C. acetobutylicum* (C) with *E. coli* (E) aspartokinase-homoserine dehydrogenases I, II, and III (encoded by *thrA*, *metL*, and *lysC*, respectively) and *B. subtilis* (B) aspartokinase II (AKII) (This study; Katinka *et al.*, 1980; Zakin *et al.*, 1983; Cassan *et al.*, 1986; Chen *et al.*, 1987). Identical residues present in three or more of the above aa sequences are in bold type and residues of the *C. acetobutylicum* GlnX aa sequence wich match with any of the above *E. coli* aspartokinase-homoserine dehydrogenases or *B. subtilis* aspartokinase, is underlined .

Chapter 3: Identification and characterization of transcription initiation sites upstream of the *C. aceto butylicum glnA* gene

Chapter 3

Identification and characterization of transcription initiation sites upstream of the *C. acetobutylicum glnA* gene.

3.1. Summary

mRNA extracted from *C. acetobutylicum* cells was used for the determination of transcription start points upstream of the *glnA* gene. Two major transcripts were identified by Northern hybridisation using part of the *C. acetobutylicum glnA* coding region as a probe. Primer-extension on mRNA isolated from *E. coli* YMC11 (pHZ200) cells harbouring the *C. acetobutylicum glnA* gene, revealed four transcription start points t_{1a}, t_{1b}, t_{2a}, and t_{2b}, of which only t_{1a}, t_{2a}, and t_{2b} were present in mRNA isolated from *C. acetobutylicum*. Transcription start points t_{1a} and t_{1b} were separated by 8 nucleotides whereas t_{2a} and t_{2b} involved adjacent nucleotides. These results confirmed the previously proposed putative promoters p_1 and p_2 and suggested that transcript t_{1b} in *E. coli* might be the product of abberant transcription from a promoter which overlaps promoter p_1.

Promoter p_2 displayed several features common to stringent response controlled promoters. A *glnA::lacZ* fusion containing only the p_2 promoter and the first eight codons of the *glnA* structural gene was used in nutritional shift-down experiments designed to induce the stringent response in both rel^+ and *relA1* strains of *E. coli*. From these experiments, it was concluded that promoter p_2 was not effected by a RelA mediated factor.

3.2. Introduction.

Only a limited number of transcription initiation sites for *Clostridium* species have been determined by mRNA transcript mapping (Graves and Rabinowitz, 1986; Béguin *et al.*, 1986) and no data for the transcription starts of *C. acetobutylicum* genes are available. In all examples studied so far, *glnA* expression in prokaryotes is regulated at the level of transcription (Chapter 1). In enteric bacteria, this involves the combined action of the *ntr* gene products, in which NtrA (σ^{54}) and phosphorylated NtrC both bind to the *glnA* promoter region to activate transcription. In contrast, regulation of *glnA* gene expression in *B. subtilis*, and possibly also in *B. cereus*, is accomplished by the action of the *glnR* gene product, which binds to the *glnRA* promoter region and represses transcription. It was shown in a earlier report that the *C. acetobutylicum glnA* gene was regulated by nitrogen levels in *E. coli* from its own regulatory region (Usdin *et al.*, 1986). However, no evidence was found for a global Ntr system (Usdin *et al.*, 1986) nor for the presence for a regulatory gene upstream of the *C. acetobutylicum glnA* gene (Janssen *et al.*, 1988; this study, Chapter 2). Because the primary stucture of the *C. acetobutylicum* GS is very homologous to the GSs from *B. subtilis* and *B. cerues* (Chapter 2), it was of interest to establish if the *C. acetobutylicum glnA* gene was regulated in the same way as its counterparts in *Bacil*lus spp. . Therefor, the transcription initiation sites of the *C. acetobutylicum glnA* gene were determined as a first step towards more detailed regulatory studies (Chapter 4).

3.3. Materials and Methods.

2.3.1. Strains and plasmids. Plasmids pHZ200, pGln1300, and plasmids used for nucleotide sequencing, have been described in Chapter 2. Plasmid pMC1871 is a gene fusion vector (Casadaban *et al.*, 1980) (Appendix D) used to fuse 3′- end truncated genes with the nineth codon of the *lacZ* coding region. *E. coli* strains K514, LK111, YMC11, and JM105 were used as described in Chapter 2. *E. coli* strains JM105 (*lac*$^-$, *rel*$^+$) and JM107(*lac*$^-$, *relA1*) were used for the nutrient shift down experiments. The industrial *C. acetobutylicum* strain P262 (National Chemical Products, South Africa) was used to isolate mRNA. All plasmids and strains used in this study are listed in Appendix A.

3.3.2. Media and Buffers.

The Lac^+ phenotype of *glnA::lacZ* fusions was tested on either McConkey (Difco) or Eosin Methylene Blue (EMB; Oxoid) agar (Appendix B). *E. coli* YMC11(pHZ200) cells, used for primer-extension reactions, were grown in CSH minimal medium. *C. acetobutylicum* was grown in *Clostridium* Basal (CB) medium (O'Brien and Morris, 1971) or in Molasses Fermentation (MF) medium (Appendix B). Media for the growth of *Clostridium* were prepared under anaerobic conditions as described in Appendix B.

3.3.3. Growth conditions.

E. coli strains were grown as described in Chapter 2. Regulatory studies on the expression of the *C. acetobutylicum glnA* gene were performed in *E. coli* using CSH minimal medium supplemented with 15 mM glutamate, 15 mM glutamine or 15 mM glutamate, 0.15 mM glutamine in order to create nitrogen-excess or nitrogen-limiting conditions, respectively. These media have been used previously for regulatory studies on heterologous *glnA* gene expression in *E. coli* (Usdin *et al.*, 1986; R. Maharaj, Ph. D. Thesis 1988, University of Cape Town, South Africa). *C. acetobutylicum* cells were grown on buffered CB medium (Appendix B) as described by Allcock *et al.* (1982) or MF medium and incubated in an anaerobic glove box at 37°C. Anaerobic conditions were maintained throughout and special care was taken when the media were prepared to avoid traces of oxygen. *C.*

acetobutylicum spore stocks were prepared and maintained according to the method of Long *et al.* (1983) (Appendix A).

3.3.4. Plasmid constructions.

glnA::lacZ fusions were constructed as follows: a *SmaI-EcoRV* fragment from pGln1300, containing the entire *glnA* upstream region and the first eight codons of the *glnA* coding region, was cloned into the *SmaI* site of pMC1871. This resulted in the fusion plasmid pLac1300. Fusion between the *glnA* gene and *lacZ* was facilitated by the presence of an *EcoRV* restriction endonuclease site in the *glnA* gene (Fig. 2.3.) which was fused in-frame with the unique *SmaI* site of pMC1871. The *SmaI* site was consequently reconstructed at the 5'-end of the *SmaI-EcoRV* fragment and used for further manipulations. Plasmid pLac1300 was linearized with *SmaI* restriction endonuclease, followed by *Bal31* (0.5 u/µL; Boehringer Mannheim) digestion (different time points; Appendix C) and subsequent restriction with endonucleases *SacI* (located within the *lacZ* gene) and *PstI* (ten-fold excess; biochemical selection). Finally, these *Bal31*-truncated / *SacI*-digested fragments were cloned back into pMC1871, wich was opened with restriction endonucleases *SmaI* and *SacI* and purified over Low Melting Agarose (LMA) (Appendix C). This resulted in *glnA::lacZ* fusion plasmids with decreasing amounts of *glnA* upstream regions. All constructions were obtained by selection for the Lac$^+$ phenotype on McConkey plates after transformation of *E. coli lac* LK111 cells with the final ligation mixture. A variety of red, pink, and white colonies were obtained after incubation for 24 h at 37°C and colonies were randomly picked for the preparation of DNA (Appendix C). When analysed by restriction endonuclease digests, recombinants were classified according to the amount of deleted *glnA* upstream sequence, and correlated with their phenotype on McConkey plates. Class I (red) represented recombinants in which up to 1100 bp of the *glnA* upstream region was deleted. Class II (pink) contained recombinants with only approximately 150 bp *glnA* upstream left, and Class III (white) recombinants were either parental pMC1871 plasmids or *glnA* upstream deletions with the essential regulatory sequences removed. One of the *glnA::lacZ* fusion plasmids choosen for further study, pLac106 (Class II), contained only 106 bp of the upstream region as determined by nucleotide sequence analysis, and had a reconstructed *SmaI* site due to a GGG sequence at position -104 to -106, located within the core region of promoter p_1.

3.3.5. Transformation procedures. Transformation of *E. coli* cells was carried out as described in Appendix .

3.3.6. Enzyme activity assays. β-galactosidase levels were measured using the colorimetric assay of Miller (1972). Total protein was determined by the method of Lowry *et al.* (1957), using bovine serum albmumin as the standard (Appendix G). Samples were taken from cultures in mid-exponential growth, unless otherwise stated in the text. Cells were harvested and washed in CSH minimal medium. Cell pellets were quick frozen using liquid nitrogen and stored at -70^{o}C. Assays were performed on cell-free lysates, obtained by sonication, using the conditions recommended for each particular assay (Appendix C).

3.3.7. Large scale preparation of mRNA. RNA from *E. coli* cells was extracted according to the method of Aiba *et al.* (1981) (Appendix C). RNA extraction from *C. acetobutylicum* cells was performed by a modification of the method of Zappe *et al.* (1986), used originally for the preparation of chromosomal DNA from *C. acetobutylicum.* CB medium in Hungate tubes was inoculated with a diluted spore stock of *C. acetobutylicum* P262 (Appendix A) and allowed to germinate and grow to a cell density of $3x10^8$ cells/mL. These cells were then used as an 1:25 inoculum for the preparative culture (600 mL MF medium). Growth was monitored by either measuring the OD_{600} or by microscopic cell count, or both. Exponentionally growing cells (3 x 200 mL) (clostridial stage/acidogenic phase) were harvested after 10 h growth by centrifugation, each cell pellet was resuspended in 10 mL of 2% SDS, 0.1 M EDTA solution, and immediately passed trough a 20 K French^R Pressure Cell (SLM Instruments, Inc., USA) under 118.5 MPa of pressure. Lysed cells were collected directly into 10 mL of buffered phenol (Appendix B), which was kept ready at 65°C in a Sorvall SS34 polycarbonate tube. The two phases were mixed by carefully inverting the tube several times and then separated by centrifugation at 27,000 g for 10 min. The hot phenol extraction was repeated once, the aquaous phase was removed to 2.1 mL Eppendorf tubes and was subjected to repeated chloroform:isoamylalcohol (24:1 v/v) extraction to remove the phenol, followed by two water saturated ether extractions. Residual ether was allowed to evaporate after the second ether extraction and nucleic acids were precipitated by the addition of 2.5 volumes of ethanol (98%) and 1/10 volume of sodium acetate (3 M). Precipitation was facilitated by storage at -20^{o}C for at least 60 min. After centrifugation (10 min Eppendorf microfuge), RNA was resolved in DNase reaction buffer (Appendix B) and RNase free DNase (21 u/μL; Boehringer Mannheim) was added to remove unwanted chromosomal DNA. After incubation for 30 min at 37^{o}C, the reaction mixture was extracted with phenol and the aqueous phase treated with chloroform: isoamylalcohol. Nucleic acids were precipated as above and the RNA pellet stored in 70% ethanol at -70^{o}C. To test wether the RNA was intact, a small sample was loaded on a submarine agarose gel (Bethesda Research Laboratory) (Appendix C). In general, RNA of fairly high molecular weight was obtained using the above method. However, *C. acetobutylicum* grown on MF medium produced extensive capsular material, especially during the clostridial stage. Despite three phenol extractions and numerous chloroform:isoamylalcohol extractions and nucleic acid precipitations, RNA pellets retained their brown color and resuspension of these pellets was difficult. This caused concern as it was noted that RNA in 70% ethanol (-70oC) was considerably degraded, even when it was stored for a relatively short time of one to two weeks. This was in contrast to *E. coli* RNA, which was also stored at -70^{o}C in 70% ethanol and after a period of eighth months, little or no degradation was observed. The usual precautions were always observed when handling RNA (Maniatis, 1984). This included the treatment of all glassware, Pipetman tips, Eppendorf tubes and reagents with 0.1 % diethylpyrocarbonate (DEPC) and the use of rubber gloves to avoid contamination with RNases.

3.3.8. Probe preparation. The probe used for Northern blotting was obtained by digestion of 10 μg pGln1300 with resriction endonucleases *NcoI* and *HindIII* (Fig. 2.3.; Appendix D). After electrophoresis of the 975 bp *NcoI-HindIII* fragment over Low Melting Agarose, the DNA was recovered from the gel and purified by subsequent phenol extraction, chloroform:isoamylalcohol (24:1 v/v) treatment and precipitation with isopropanol and Na-perchlorate (Appendix C). The DNA probe was labelled with [α-^{32}P] dATP to a high specific activity by nick-translation (Appendix C; Rigby *et al.*, 1971). The progress of the reaction was monitored by Cherenkov counting of trichloro-acetic acid (TCA) precipitated material, recovered on glass filters (Whatman GFC) according to the techniques described in Maniatis *et al.* (1984) Contaminating nucleotides were removed from the radioactively labelled probe preparation using a Sephadex G50 spin column as described by Maniatis *et al.* (1982). Specific activities of approximately 1 x 107 counts/min/μg of DNA were routinely obtained. The radioactively labelled probe was stored in a lead container at -20oC and denatured by boiling in a fume hood just before use.

3.3.9. Northern blotting. Northern hybridisation was performed essentially as described by Maniatis *et al.* (1982) using a 1.2% agarose formaldehyde gel. Hybridisation and washing were done at 68°C under stringent conditions. Linear DNA fragments were used as molecular weight markers. Electophoresis proceeded for 5 h at 6V/cm (100V) until the two markers were separated symmetrically. All buffers, reagents, and equipment, used for the Northern blotting, are described in Appendices B and C.

3.3.10. Preparation of template DNA. Large scale preparation of DNA was as described in Appendix C. Template DNA for sequencing reactions was prepared as described in Chapter 2.

3.3.11. Primers and primer anealing reactions. Primer#1, described in section 2.3.8., was used for the synthesis of cDNA and sequence analysis in the primer extension reactions. Annealing of primer#1 to template DNA for sequence analysis was performed as described in Chapter 2. For cDNA synthesis, primer#1 was anealed to RNA as follows: [^{32}P]-5′ end labeled primer#1 (0.2 pmoles; Appendix C) was mixed with either 50 μg RNA obtained from *E. coli* YMC11(pHZ200) grown in nitrogen-limiting or nitrogen-excess medium, or 100 μg RNA from *C. acetobutylicum* P262 grown in MF medium. The mixture was denatured at 95°C in 10 μL start solution (10 mM Tris-HCl pH 8.3, 0.35 mM EDTA; Appendix C) for 5 min, chilled on dry ice, and incubated for 6-10 h at

45^{o}C. Calculation of the ideal hybridisation temperature, based on the GC-content of the used primer, was done as described by Mason and Williams (1985). At this temperature, only perfect hybrids should be stable. Sequence analysis of pLac constructs was performed using primer#3 (5'-CG.CCA.GGG.TTT.TCC.CAG-3'; D. Botes, Dept. of Biochemistry, University of Cape Town) which is complementary to a sequence located at positions +46 to +62 within the *lacZ* coding region. This primer was preferred to the commercially available primers because sequence analysis of the fusion point between the *glnA* and *lacZ* genes was required.

3.3.12. Primer extension reactions. After the annealing reaction of primer#1 to RNA was completed, nucleic acids were precipitated with 70 % ethanol and 0.3 M Na-acetate and the pellet was washed twice with 95% ethanol and resuspended in 18 μL reverse transcriptase reaction buffer (RTB; Appendix C). RNase inhibitor (1 μL) (RNasin; Boehringer Mannheim, 25 U/μL), together with 2 μL deoxynucleotides (dNTP; 10 mM) and 1 μL Reverse Transcriptase (RT; Boehringer Mannheim, 21 U/μL), were added and the reaction mixture was incubated for 45 min at 42^{o}C. The RNA was degraded by adding 2 μL RNase A (Sigma, 10 ng/μL), and the mixture was incubated for a further 10-30 min at 18-20^{o}C. After phenol extraction and ethanol precipitation, the cDNA was resuspended in 10 μL of 99 % formamide, 0.1 mg/mL bromophenol blue, 0.1 mg/mL xylene cyanol and electrophoresed in a 6% acrylamide-7 M urea sequencing gel (Chapter 2) together with the sequencing reactions (using primer#1), performed on plasmid pHZ200.

3.3.13. Sequencing reactions. All sequence analysis reactions were performed on plasmid DNA as has been described in Chapter 2.

3.4. Results and discussion.

3.4.1. Identification of *C. acetobutylicum glnA* mRNA transcripts.

Fig. 3.1: Northern blot analysis of *C. acetobutylicum glnA* RNA (50 µg). RNA extraction and Northern blot procedures are described in Materials and Methods. The *C. acetobutylicum* RNA blot was probed with ^{32}P labeled DNA obtained by nick-translation. Trailing hybridizing material is the likely result of RNA-degradation and aspecific hybridisation.

sequencing gel (Chapter 2) together with the sequencing reactions (using primer#1), performed on plasmid pHZ200.

3.3.13. Sequencing reactions. All sequence analysis reactions were performed on plasmid DNA as has been described in Chapter 2.

3.4. Results and discussion.

3.4.1. Identification of *C. acetobutylicum glnA* mRNA transcripts.

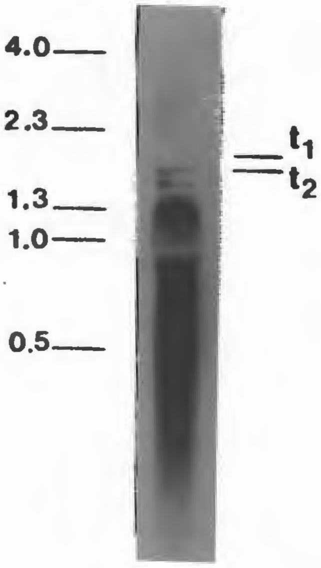


Fig. 3.1: Northern blot analysis of *C. acetobutylicum glnA* RNA (50 µg). RNA extraction and Northern blot procedures are described in Materials and Methods. The *C. acetobutylicum* RNA blot was probed with ^{32}P labeled DNA obtained by nick-translation. Trailing hybridizing material is the likely result of RNA-degradation and aspecific hybridisation.

Two major *glnA* mRNA transcripts were detected by Northern blot analysis on RNA extracted from clostridial phase cells of *C. acetobutylicum* P262 (Fig. 3.1.) using part of the *glnA* gene as a probe. The size of the two transcipts, t_1 and t_2, was estimated to be 1.80 and 1.66 nt, respectively. This would be sufficient to code for the GlnA polypeptide, which is 444 amino acids long and has a predicted M_r of 49,630 daltons.

As has been discussed in Chapter 2 (section 2.4.2.3.), the 3′-terminal of the *C. acetobutylicum glnA* mRNA contains several inverted repeat sequences within a 200 nt region downstream of the coding region. It has been suggested that one of these alternative secondary structures would be capable of forming a factor-independent transcription terminator (Janssen *et al.*, 1988). It is tempting to speculate that the two transcripts, t_1 and t_2, seen in Fig. 3.1., originate from the two putative promoters p_1 and p_2 (Fig. 2.3.), respectively, and that both would end within the 200 bp *glnA* downstream region. However, to elucidate the precise start points of the two transcripts, primer-extension experiments were undertaken.

3.4.2. Analysis of *C. acetobutylicum glnA* promoters.

3.4.2.1. Determination of *glnA* transcriptional start sites.

The start sites of *glnA* transcription were determined by primer-extension using a synthetic oligonucleotide primer that hybridises to *glnA* mRNA, 37 to 54 nt from the *glnA* start codon (Fig. 2.3.).

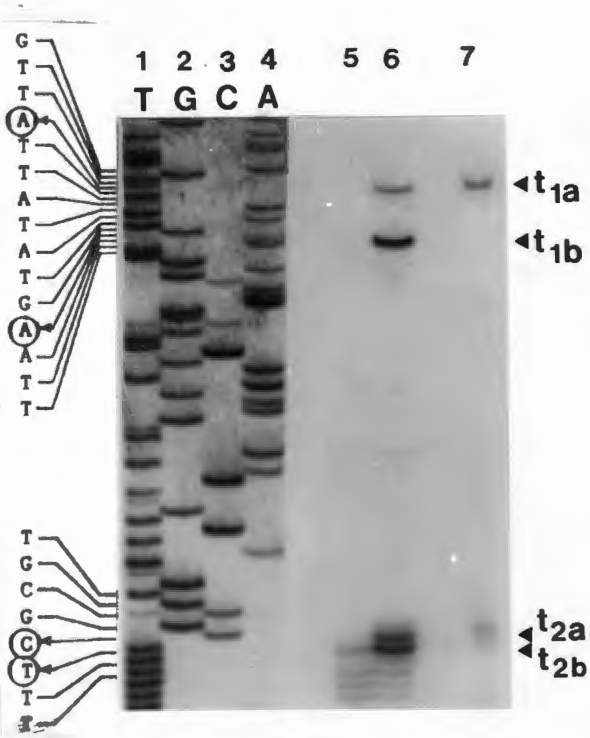


Fig. 3.2.: Identification of *C. acetobutylicum glnA* transcription start points. Primer-extension analysis was performed using a 18-mer synthetic oligonucleotide primer as described in Materials and Methods. Lanes 1-4: sequencing ladders, derived from sequence analysis on pHZ200 with the same primer, are shown with the A, C, G, and T reactions from left to right. Lanes 5-6: *E. coli* YMC11 *glnALG*(pHZ200) cells, grown in nitrogen-excess (Lane 5) and nitrogen-limiting (Lane 6) medium. Lane 7: *C. acetobutylicum* P262 cells grown in MF medium. Four major transcripts, t_{1a}, t_{1b}, t_{2a}, and t_{2b}, are indicated.

Primer-extension reactions were carried out on RNA extracted from exponential phase cells of *E. coli* YMC11 *glnALG* (pHZ200), grown in CSH minimal medium, or from exponential phase cells of *C. acetobutylicum* P262, grown in molasses (FM) medium. The autoradiogram showed four major transcription start points t_{1a}, t_{1b}, t_{2a}, and t_{2b} for *E. coli* YMC11 (pHZ200) cells grown in nitrogen-limiting medium (Fig. 3.2.; Lane 6). Transcription start points t_{1a} and t_{1b} were separated by 8 nucleotides whereas t_{2a} and t_{2b} involved adjacent nucleotides. It appears that the transcripts t_{2a} and t_{2b} were controlled by the consensus promoter p_2 (Fig. 3.3.), because these startpoints, located at positions -33 and -32, respectively, were preceded at position -37 by a consensus -35 TATAAT box belonging to the p_2 promoter with the sequence TTCGAT--N_{17}--TATTGT.

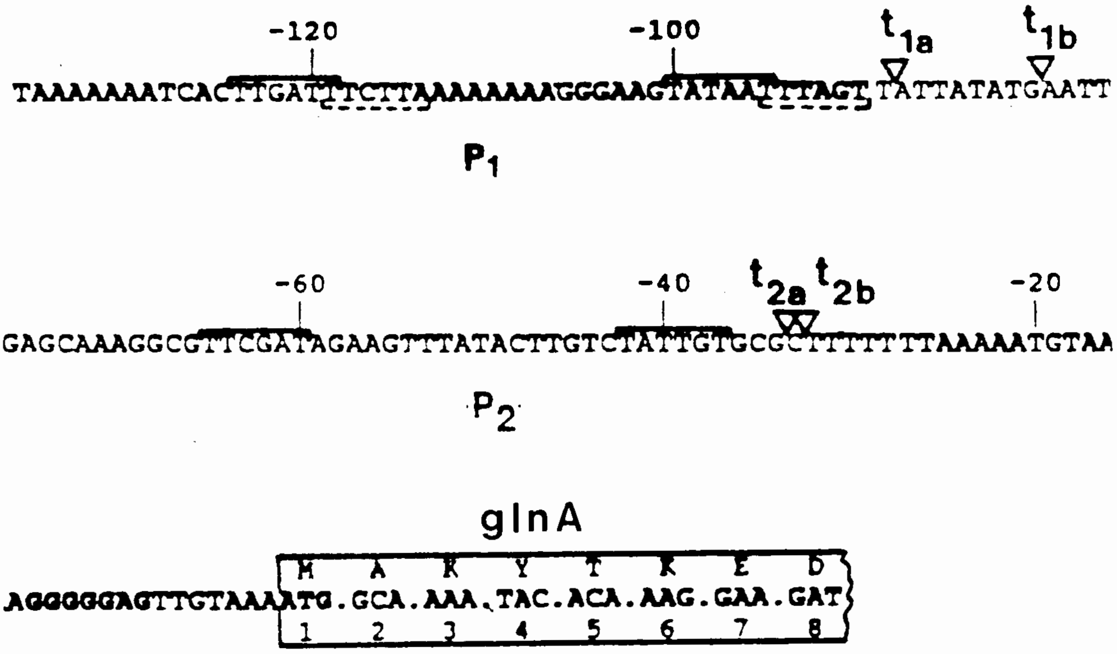


Fig. 3.3.: Transcription initiation sites of *glnA* mRNA transcripts under control of *C. acetobutylicm glnA* upstream promoter sequences. Transcription startpoints t_{1a}, t_{1b}, t_{2a}, and t_{2b}, are indicated by an open triangle above the sequence. Putative promoter $p_{1.1}$ is indicated with a dashed line under the sequence.

Start points t_{1a} and t_{1b} are located at positions -88 and -80, repectively. Because of the distance separating the start points t_{1a} and t_{1b}, it is possible that the transcript t_{1a} originated from the consensus promoter p_1, previously suggested to be between positions -124 and -95 (Janssen *et al.*, 1988), and that the transcript t_{1b} was controlled by an adjacent putative promoter ($p_{1.1}$) containing -10 and -35 RNA polymerase-binding consensus sequences, situated 5 bp downstream and overlapping p_1 (Fig3.3.). Initiation of transcription from t_{a1}, t_{1b}, and t_{2a} was repressed in exponentional phase *E. coli* YMC11(pHZ200) cells grown under nitrogen-excess conditions but transcript t_{2b} and the start points below t_{2b} were not markedly affected (Fig. 2.3.; Lane 5). It seems that regulation of the *glnA* gene at the level of transcription was mainly exerted via the repression/derepression of p_1 promoter activity. This promoter, which has the sequence TTGAtt--N18--TATAAT, and resembles the consensus sequence with 10 matches out of 12, is overlapped by a complex palindromic sequence (CPS). The involvement of this CPS in the regulation of the *C. acetobutylicum glnA* gene has been reported (Janssen *et al.* 1989) and will be discussed in Chapter 4.

Primer-extension analysis was also performed on RNA isolated from exponentional *C. acetobutylicum* cells grown in molasses medium. In *C. acetobutylicum*, *glnA* transcripts were initiated mainly at the start points t_{1a} and t_{2a}, with weak initiation from t_{2b} (Fig. 3.2.; Lane 7). Initiation from the start point t_{1b} was not observed in *C. acetobutylicum*, whereas it was a major start point regulated by nitrogen in *E. coli* YMC11(pHZ200). Another difference between the initiation of *glnA* transcripts in *C. acetobutylicum* and *E. coli* involved the transcription initiation products below t_{2b} which were only observed with mRNA isolated from *E. coli* YMC11(pHZ200) cells but not with mRNA isolated from *C. acetobutylicum* cells. In addition, these transcripts did not appear to be regulated by nitrogen levels in *E. coli*. It is suggested that they were spurious and perhaps due to the extensive A-T rich sequence in this region.

Although growth conditions (e.g. nitrogen availability) for the *E. coli* cells, grown in synthetic medium, differ from the growth conditions for the *C. acetobutylicum* cells, grown in FM medium, it is apparent that the same *glnA* transcription initiation signals are being used in both *E. coli* and *C. acetobutylicum*. The only significant difference in transcription initiation of the *C. acetobutylicum glnA* gene between the two organisms is the use of the promoter $p_{1.1}$ in *E. coli*. It may be argued that transcription of $p_{1.1}$ possibly interferes with regulatory events in *E. coli* and that its presence may distort experimental results on the expression of the *C. acetobutylicum glnA* gene when studied in *E. coli*. However, the fact that transcript t_{1b} is regulated by nitrogen in *E. coli* in the same way as transcript t_{1a}, indicates that this is unlikely. The various *glnA* upstream promoter sequences, investigated in this study, are summarized in Fig. 3.4.

Ec	------------	TTGACA	------------------	TATAAT	-----
G^+	T--AAAAA----	TTGACA	------A-------T-TG-	TATAAT	-AA-A
glnAp$_1$	TAAAAAAATCAC	TTGATT	TCTTAAAAAAAAGGGAAG	TATAAT	TTAGT
glnAp$_2$	TGAGCAAAGGCG	TTCGAT	AGAAGTTTATACTTGTC	TATTGT	GCGCT
glnAp$_{1.1}$	AAATCACTTGAT	TTCTTA	AAAAAAAGGGAAGTATAA	TTTAGT	TATTA

(Bases that appear in any given position in more than 50% of the examples are indicated in the consensus sequences; corresponding positions in the *glnA* promoters are underlined)

Fig. 3.4.: Comparison of *glnA* upstream promoters with the *E. coli* (Ec) and Gram-positive (G^+) consensus promoter sequences (McClure, 1985; Graves and Rabinowitz, 1986).

3.4.2.2. Stringent response controlled promoter.

Bacteria which experience starvation caused by amino acid deprivation or exhaustion of their primary carbon source make a swift and wide-ranging set of adjustments to their metabolism (reviewed by Cashel and Rudd, 1987). The major response of the cell to starvation is a decrease in protein synthesis, accomplished by the specific inhibition of the production of stable RNA's. At the same time expression of amino acid biosynthetic operons is allowed, improving cell economy. There is extensive evidence that this selective regulation is, at least partly, achieved by the control of transcription initiation (reviewed by Lamond and Travers, 1985b; Cashel and Rudd, 1987).

Amino acid starvation triggers the synthesis of the 'alarmone' guanosine 5'-diphosphate 3'-diphosphate (ppGpp) by the RelA protein. It has been proposed that ppGpp interacts with RNA polymerase, either directly or in conjunction with an acessory protein, and this would result in an altered promoter specificity of $E\sigma^{70}$ (Kajitani and Ishihama, 1984; Ishihama *et al.*, 1988). Riggs *et al.* (1986) have performed *in vitro* experiments on DNA templates containing the promoter proximal portion of the *S. typhimurium his* operon fused in-frame to *lacZ*. Addition of ppGpp stimulated *his::lacZ* expression 22-fold. Site-directed mutagenesis in the -10 region of the his promoter sequence resulted in a 10-fold increase of *his::lacZ* expression in the absence of ppGpp, and reduced the stimulation (after addition of ppGpp) to less than a factor of 2. By comparison of 49 ppGpp sensitive promoter sequences, Riggs *et al.* (1986) elucidated the importance of the fourth and fifth positions of the -10 hexamer sequence in the activation/repression by ppGpp.

The downstream region, or 'discriminator sequence', adjacent to the -10 region of ppGpp sensitive promoters, was also shown to be important in regulating transcription. Studies of ppGpp inhibition of promoter mutants produced by oligonucleotide mutagenesis and *in vitro* deletion have implicated the -7 to -4 positions, normally GCGC, followed by an A-T rich region (Mitzushima-Sugano and Kaziro, 1985). Changing each of the GC pairs to AT pairs individually at the -7 to -4 positions lessened the degree of ppGpp inhibition in vitro. Similar results were obtained by Lamond and Travers (1985a). Briggs *et al.* (1986) have suggested that ppGpp regulates a step after closed complex formation. Apparently, a GC rich discriminator sequence and a poor entry region would be rate-limiting for open complex formation. A detailed study on the effect of such a GCGC motif on growth-rate control and stringent sensitivity has recently been presented by Zacharias *et al.* (1989).

Promoter p_2 displays several features common to stringent response controlled promoters. It has an entry region (the -35 box), which deviates from the canonical TTGA sequence, a Pribnow box with a T and a G in the 4th and 5th positions, respectively, and a 5′-GCGCTTTTTTAAAAA-3′ discriminator sequence immediately downstream of the -10 region (Fig. 3.3). A positive stringent control system has been reported for glutamine/glutamate synthesis (*gltB*, *glnA*, *gdh*) in *S. typhimurium* (Sales and Brenchley, 1982).

In order to investigate the possible role of the relA gene product, ppGpp synthetase, in the regulation of promoter p_2 by nitrogen levels, a *glnA::lacZ* fusion plasmid, pLac106, was constructed. This plasmid, which contained only p_2 and the first eight codons of the *C. acetobutylicum* *glnA* gene fused to the *lacZ* gene (Materials and Methods), was transformed into *E. coli* strains JM105 (*lac*$^-$, *rel*$^+$) and JM107(*lac*$^-$, *relA1*). The production of the fusion protein by these strains was investigated in nutrient shift-down experiments designed to induce the stringent response. β-galactosidase activity of *E. coli* JM105 (pLac106) cells increased 2.4- and 2.0-fold in the first and second nutritional shift down media respectively. β-galactosidase activity of *E. coli* JM107 *relA1* (pLac106) cells increased 3.4- and 3.3- fold in the first and second shift down media respectively (Table 3.1.).

Table 3.1.: Nutrient shift down experiments using the *glnA::lacZ* fusion plasmid pLac106[c].

Strain[a]	LB[b]	Ngln	gln
rel$^+$	280	960	3,210
relA1	340	820	1,650

(a) *E. coli* host strains used were JM105 (*rel*$^+$) and JM107 (*relA1*).
(b) Both strains JM105 (pLac106) and JM107 (pLac106) were pre-grown overnight in Luria-Bertani (LB) medium untill stationary phase was reached. These cells were used to inocculate fresh LB medium (1:25) and strains were grown for 75' min till mid-exponentional phase. Samples were taken at this stage for enzyme activity assays. At the same time CSH minimal medium supplemented with 8 mM (NH4)2SO4, 4 mM glutamine (Ngln) was inocculated (1:10) and strains were grown for 6.5 h till mid-exponentional phase. Again samples were taken to measure enzyme activity and simultaneously, cells were inocculated (1:10) into CSH minimal medium containing only 4 mM glutamine (gln) and grown for 16 h untill mid-exponentional phase was reached. At this point, samples were taken and enzyme activities were measured.
(c) Enzyme activities are expressed as the amount of ONP released in μmoles/min/mg protein.

Since β-galactosidase acivities were not markedly affected by nitrogen starvation in the *E. coli rel*$^+$ strain in comparison with the *E. coli relA1* strain, it is concluded that although the p_2 promoter displayed structural features common to stringent response promoters, it is not regulated by a *relA* mediated factor.

Chapter 4: Regulatory studies on the expression of the *C. acetobutylicum glnA* gene in *E. coli*.

Chapter 4

Regulatory studies on the expression of the *C. acetobutylicum glnA* gene in *E. coli*

4.1. Summary.

Expression from the *C. acetobutylicum glnA* promoters p_1 and p_2 was affected by up- and downstream sequences of the *glnA* coding region. Regulatory studies on *glnA* upstream deletion mutants showed that the sites involved in the regulation of *glnA* gene expression were contained within a 550 bp region of the *C. acetobutylicum glnA* gene. The deletion of a complex palindromic sequence, which overlaps promoter p^1, resulted in high levels of GS activity. Studies on *glnA* gene expression were also performed using *glnA::lacZ* fusion plasmids containing decreasing amounts of *glnA* upstream sequences. β-galactosidase levels coded for by these plasmids were found to be regulated by nitrogen in the absence of functional GS enzyme, although repression ratios were low. It was established that repression of the *C. acetobutylicum glnA* gene was not affected by GS *in trans*.

Deletion of the 3.5 kb *glnA* downstream region caused deregulation of transcription initiation: removal of this region resulted in derepression of *glnA* transcription, irrespective of nitrogen levels. Although downstream sequences (i.e. regulatory genes) may play a role in the regulation of transcription, this was not reflected in the levels of GS activity. It is suggested that the differences between the regulation of transcription and the levels of GS activity are due to regulatory mechanisms at the level of translation. The possible involvement of anti-sense RNA, transcribed from the putative *glnA* downstream promoter p_3 (Chapter 2) and which would bind to the RBS of the *C. acetobutylicum glnA* gene, thereby inhibiting translation efficiency, is discussed.

4.2. Introduction.

Bacteria respond to nitrogen deprivation by increasing the intracellular concentration of GS, together with elevated levels of expression from a number of genes essential for the uptake or degradation of nitrogen-containing compounds. In *E. coli* and other enteric bacteria such as *S. typhimurium* and *K. pneumoniae*, this involves a global *ntr* regulatory system, which includes the use of core polymerase complexed with σ^{54}, the bacterial activator protein NR_I, and a set of several *gln* genes unlinked with the *glnALG* operon (see Chapter 1). The *ntr*-network exists also in non-enterobacterial species such as *R. meliloti* (Szeto *et al.*, 1987), *R. leguminosarum* (Filser *et al.*, 1986), and *A. tumefaciens* (Rossbach *et al.*, 1987; Wardham *et al.*, 1989), but in contrast to the operon arrangement in enterobacteria, *ntrBC* (*glnLG*) genes are not linked with *glnA*. *Ntr* regulated genes, whose products have different physiological functions, have been detected in several other organisms such as *Caulobacter crescentus*, *Neissaria gonorrhoeae*, and *Rhodobacter capsulatus* (reviewed by Kustu *et al.*, 1989).

However, there is no evidence for the presence of global-regulatory *ntr*-genes in the Gram-positive endospore forming aerobe *B. subtilis*. In contrast to the enteric bacteria, it appears that in *B. subtilis* GS regulates its own expression. Recently, it has been suggested by Schreier *et al.* (1989) that the GS of *B. subtilis* interacts with a regulatory protein, GlnR, which negatively regulates expresion of the *glnRA* operon. This interaction between GS and GlnR may involve the formation of a large protein complex with the potential to repress *glnA* expression or it may involve the modification of the *glnR* gene product in response to a metabolic signal, turning GlnR into a repressor. Furthermore, this interaction of the *B. subtilis* GS with *glnR* must be highly specific as the GS coded for by *C. acetobutylicum* was not able to substitute the *B. subtilis*

GS activity in *glnRA* regulation (Schreier, unpublished; cited in Schreier *et al.*, 1989), despite the high amino acid homology (60%) and structural similarity (both lack the central loop) between the two enzymes. In addition, Zhang *et al.* (1989) have identified - by nucleotide sequencing - a number of mutations within the *B. subtilis glnA* gene. These mutations, which resulted in altered catalytic properties and overproduction of GS antigen, mapped within or near sequences which contribute to the central loop region and the active site of the *S. typhimurium* GS enzyme. This confirmed an earlier report of Schreier *et al.* (1985), who showed that this region is required for the regulation in *trans* of β-galactosidase levels coded for by a plasmid in which the *glnRA* promoter region was fused to *E. coli lacZ*. Similar *in trans* studies have also been performed in various *B. subtilis glnA* mutants (Schreier *et al.*, 1986) previously isolated and characterised by Dean *et al.* (1977), and Fisher and Sonenshein (1984). A recent report of Nakano *et al.* (1989) indicates that regulation of *glnA* gene expression in *B. cereus* might resemble the control of *glnA* expression in *B. subtilis*.

The *C. acetobutylicum glnA* upstream region does not contain any *ntr* regulatory sites nor does it contain an $E\sigma^{54}$ consensus promoter sequence (Janssen *et al.*, 1988; this study). This is in agreement with earlier results of Usdin *et al.* (1986), who showed that, in an *E. coli glnALG* mutant, the cloned *glnA* gene was expressed and regulated by nitrogen levels from its own regulatory region. From this point of view, control of *glnA* gene expression in *C. acetobutylicum* could possibly be similar to *glnA* gene regulation in *Bacillus*, in which the GS enzyme itself has an important regulatory role.

C. acetobutylicum, *B. subtilis* and *B. cereus* are all three Gram-positive endospore forming organisms. In addition, the *C. acetobutylicum* GS enzyme, like the *B. subtilis* and *B. cereus* GS enzymes, not only lacks the central loop region, thought to be an integral part of the GS active site (Chapter 1, Chapter 2), but it also showed a high aa similarity (approximately 60%) with the *Bacillus* GS enzymes (Chapter 2). However, no upstream regulatory gene (*glnR*) could be detected (Chapter 2, Appendix E) and transcription of the *C. acetobutylicum glnA* gene was controlled by two promoters p_1 and p_2, which were shown to be functional in both *E. coli* and *C. acetobutylicum* cells (Chapter 3). In an attempt to elucidate the mechanisms whereby the *C. acetobutylicum glnA* gene is regulated by nitrogen levels in *E. coli*, *glnA* upstream deletion mutants and *glnA::lacZ* fusion plasmids were constructed. Furthermore, to establish whether the *C. acetobutylicum* GS itself might be involved in the regulation of *glnA* gene expression, *in trans* studies were performed in *E. coli glnALG* mutants.

4.3. Materials and methods.

4.3.1. Strains and plasmids. Plasmids pLK34, pHZ200, and pGln1300, have been described in Chapter 2. Plasmid pAcYc177 (Chang and Cohen, 1978) was used as a general cloning vector for *in trans* studies (Appendix D). Plasmid pMC1871 is a gene fusion vector (Casadaban *et al.*, 1980) (Appendix D) used to fuse 3′-end truncated genes with the nineth codon of the *lacZ* coding region. *E. coli* strains K514, LK111, YMC10, YMC11, and ET8051 were used as described in Chapter 2. All plasmids and strains used in this study are listed in Appendix A.

4.2. Media and buffers. *E. coli* strains containing plasmids carrying the *lacZ* gene or *glnA::lacZ* fusions were grown on either McConkey agar or Eosin Methylene Blue (EMB; Difco) agar (Appendix B). In general, *E. coli* strains were grown on Luria-Bertani (LB) medium or Cold Spring Harbor minimal medium (CSH; Miller *et al.*, 1972), supplemented with the appropriate antibiotic(s). Tetracycline (Tc), kanamycin (Kn) and ampicillin (Ap) were added to media at concentrations of 15, 25 and 100 μg/mL, respectively. Nitrogen-limiting conditions were achieved by growth of the *E. coli* cells in CSH minimal medium supplemented with 15 mM glutamate, 0.15 mM glutamine or 4 mM glutamine, and nitrogen-excess conditions by growth in CSH minimal medium supplemented with 15 mM glutamate, 15 mM glutamine or 4mM glutamine, 8 mM $(NH_4)_2SO_4$. All *E. coli* and strains were maintained as described in Appendix A.

4.3.3. Growth conditions. *E. coli* strains were grown as described in Chapter 2, unless otherwise stated in the text.

4.3.4. Plasmid constructions. Construction of plasmid pHZ202 has been described previously (Usdin *et al.*, 1986). Plasmid pGln1300, previously named pJP3 (Janssen *et al.*, 1988), containing the *C. acetobutylicum glnA* structural gene, approximately 1300 bp of upstream sequence and 614 bp of downstream sequence which included putative promoter p_3 and inverted repeat sequences, was the source of DNA for the plasmid constructs. Plasmids pLK34 and pMC1871 have been described in Chapter 2 and Chapter 3, respectively. Plasmid pgln6 (Backman *et al.*, 1981), which carries the *E. coli glnA* gene, was a generous gift from Dr. B. Magasanik, M.I.T., Boston, USA.

Upstream deletion pGln plasmids were constructed as follows: pGln1300 was digested with restriction endonuclease *SmaI*, quick purified by phenol extraction (Appendix C) and further digested under controlled conditions (reaction times 0 to 14 min in intervals of 2 min) using calibrated *Bal31* exonuclease (Appendix C). These truncated DNA fragments were subsequently digested by restriction endonucleases *NcoI* and *PstI* (the latter for biochemical selection against parental molecules) and *Bal31* truncated/ *NcoI* restricted fragments were ligated to pGln1300, opened by *SmaI* and *NcoI* restriction endonucleases. The ligation mixture was transformed into YMC11 *glnALG* (Appendix C) and Ap^r $glnA^+$ transformants were identified on CSH minimal agar plates containing 0.1 g/L $(NH_4)_2SO_4$ (Appendix A). The extent of each deletion was monitored by an *EcoRI* digest on purified

DNA. In this way, an ordered set (approximately every 100 bp) of upstream deletion mutants was obtained.

Upstream deletion mutants pGlnUR10, pGln272 and pGln211 contained 550 (approximately), 272 and 211 bp of upstream sequences, respectively, as was determined by detailed restriction endonuclease mapping and DNA sequence analysis (Fig. 4.4). Plasmid pGln146, which contained 146 bp of *glnA* upstream sequence, was constructed by subcloning a 449 bp *SspI-NcoI* fragment of pGln211, containing 146 *glnA* upstream sequence and 303 bp of *glnA* coding region, into pGln1300, previously opened with *SmaI* and *NcoI* restriction endonucleases. Deletion plasmids pGln211Δp_2 and pGln146Δp_2 were derivatives of pGln211 and pGln146, respectively, in which promoter p_2 was deleted by removal of a 28 bp *TaqI-HinPI* fragment (Fig. 4.4).

Plasmids carrying *glnA::lacZ* fusions, including pLac1300, pLac524, pLac300, and pLac106 (Fig. 4.4), were constructed as described in Chapter 3. Fusion plasmid pLac146 was constructed by the subcloning of a 170 bp *SspI-EcoVI* fragment of pGln1300 into the *SmaI* site of fusion vector pMC1871. Both *EcoVI* and *SmaI* half sites are in reading frame with the *lacZ* and *glnA* genes, respectively. *PstI* fragments, containing the complete *glnA::lacZ* fusion from plasmids pLac1300 and pLac524 (Chapter 2), were used to construct pTsLac1300 and pTsLac524, respectively, by cloning these fragments into the unique *PstI* site of pAcYc177 (Appendix D), which contains the p15A replicon (Chang and

Cohen, 1978). The p15A replicon is compatible with the pBR322 replicons in pLK34-, pGln- and pLac-derived plasmids.

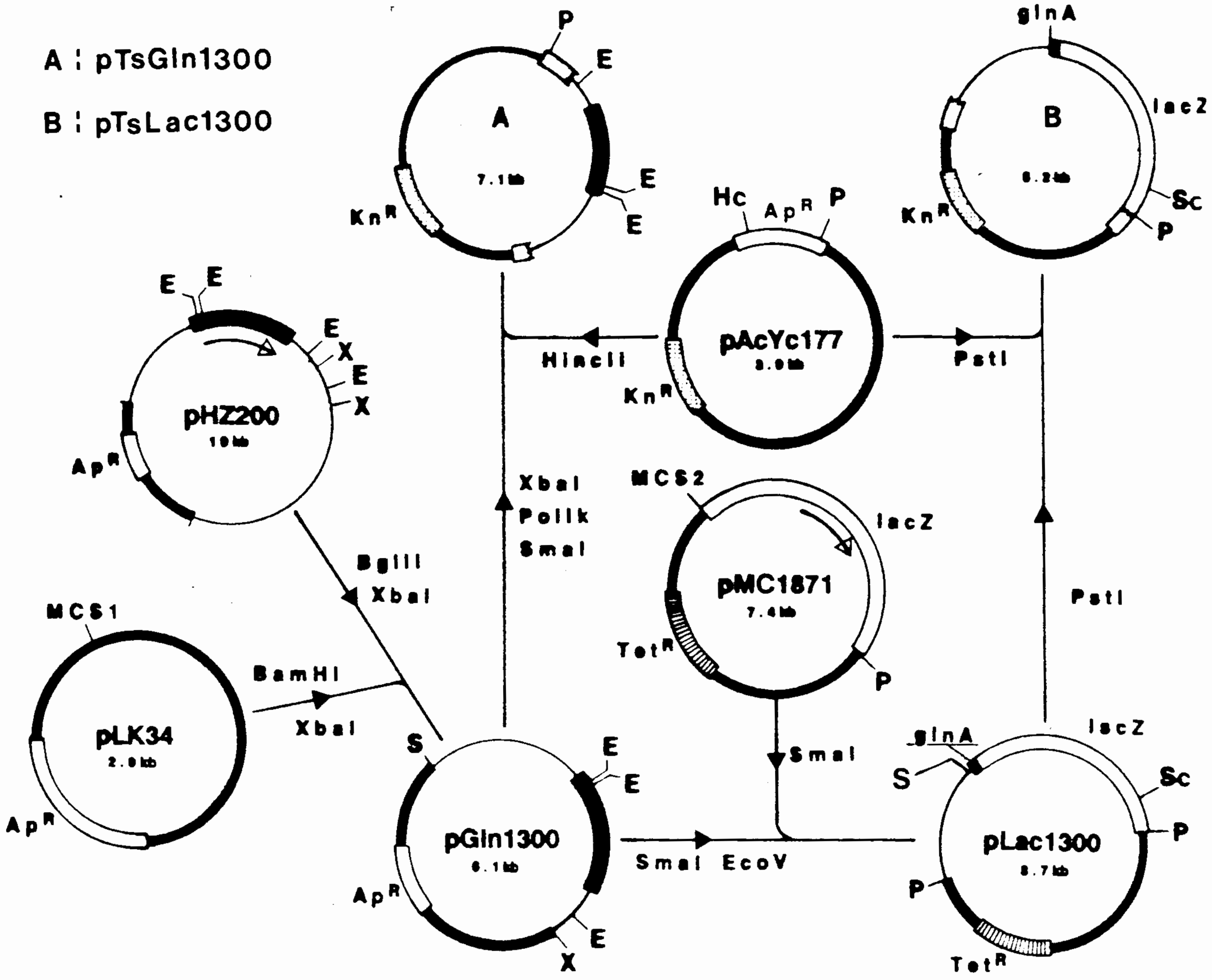


Fig. 4.1.: General outline of cloning procedures. Plasmids pGln1300 an pLac1300 were used as the source for construction of deletion mutants (see further; materials and methods) Restriction sites indicated are: E, *EcoRV*; X, *XbaI*; S, *SmaI*; P, *PstI*; Sc, *ScaI*. "PolkI" stands for DNA polymerase I - Klenow fragment. Multiple cloning site MCS1 contains (clockwise) *EcoRI*, *SmaI*, *BamHI*, *SalI*, *PstI*, *HindIII*, *XbaI*. Multiple cloning site MCS2 contains (clockwise) *PstI*, *SalI*, *BamHI*, *EcoRI*, *SmaI*, and *BamHI*. Solid lines (▬) indicate pBR322 derived sequences in plasmids pLK34, pHZ200, and pMC1871. Thin lines (———) indicate C. acetobutylicum DNA sequences. Antibiotic indicator genes are abbreviated as Ap (ampicillin) (▭), Tet (tetracycline) ▥, and Kn (kanamycin) (░)

Regulatory studies in which the *C. acetobutylicum glnA* gene was placed *in trans* with a *glnA::lacZ* fusion required the subcloning of a *SmaI-XbaI* fragment from pGln1300 into the HindII site of plasmid pAcYc177 (Appendix D), after the *XbaI* half-site was filled in by Klenow fragment of DNA polymerase I (Boehringer Mannheim, FRG) (Appendix C). This resulted in the construction of plasmid pTsGln1300. The constructs used in this study are summarized in Fig. 4.1.

Plasmid pEcGln was constructed by the subcloning of a 3.5 kb PstI-PvuII fragment from pgln6 (Backman *et al.*, 1981) into pLK34, linearized with *SmaI* and *PstI*. In both pgln6 as well as in pLK34, the *PstI* site is located within the β-lactamase gene at the same position - both are pBR322 derivatives - (Appendix D), which facilitated selection for the correct construction of pEcGln (Ap^R). Plasmid pEcGln was preferred as a control in *in trans* studies because all *glnA* upstream deletion mutants were derivatives of pLK34 which is a high copy number plasmid lacking the repressor of primer (*rop*) gene originally situated on a 309 nt *HpaII* fragment in pBR322 (Botterman and Zabeau, 1987). It is at this stage worthwhile mentioning that pEcoR251, and derivatives pHZ200 and pHZ202, are similar high copy number plasmids, lacking the same region of pBR322 containing the *rop* gene.

Sub-cloning strategies utilized were essentially according to the standard recombinant DNA techniques used by Maniatis *et al.* (1984). Otherwise, general procedures of

DNA recombinant technology used throughout this study have been summarized in Appendix C. Nucleotide sequences of plasmid constructs were confirmed by sequence analysis as described in Chapter 2, except for plasmid constructs pGln1300, pGlnUR10, and pGlnUR8, which were confirmed by detailed restriction endonuclease mapping.

4.3.5. Transformation procedures. Transformation of *E. coli* cells was carried out as described in Appendix C.

4.3.6. Enzyme activity assays. GS activity was determined by the γ-glutamyl transferase (γ-GT) assay according to Shapiro and Stadtman (1968). The amount of γ-glutamyl hydroxymate formed was calculated from a standard curve of synthetic γ-glutamyl hydroxymate (Sigma) (Appendix G). The γ-GT assay for *C. acetobutylicum* GS had a relatively sharp pH optimum between pH6.75 and pH7.00 as was determined by measuring the effect of pH on γ-GT activity of GS enzyme extracted from *E. coli* strain YMC11(pHZ200) or *C. acetobutylicum* strain P262 (K. Usdin, unpublished results; see also Appendix C). Throughout this study, a pH value of 6.75 was maintained for γ-GT assays. Specific activity of the enzyme was defined as the number of nmoles of γ-glytamyl hydroxymate formed per min per mg protein. β-galactosidase and total protein levels were measured as described in Chapter 3. All samples were taken in mid-exponentional growth, unless otherwise stated in the text. Cells were harvested and washed in either 1% KCl or CSH minimal medium (Appendix B). Cell pellets were quick frozen using liquid nitrogen and stored at -70°C until used. Assays were performed on cell-free lysates, obtained by sonication, using the conditions recommended for each particular assay (Appendix C).

4.3.7. Primer-extension reactions. Transcription initiation sites were determined by procedures described in Chapter 3.

4.4. Results and discussion.

4.4.1. The involvement of *glnA* up- and downstream sequences in the regulation of *glnA* gene expression. Two *glnA* upstream promoters, p_1 and p_2, were previously identified by primer-extension reactions, carried out on mRNA extracted from an *E. coli glnALG* mutant strain containing the cloned *C. acetobutylicum* GS and from *C. acetobutylicum* P262 (Chapter 3). Both *glnA* promoters were used in *E. coli* as well as in *C. acetobutylicum* cells, although the use of a third tentative upstream promoter sequence, $p_{1.1}$, apparently resulted in the abberant transcription of the *C. acetobutylicum glnA* gene in *E. coli* (Chapter 3; Janssen *et al.*, 1989). Previous results have shown that expression of the cloned *C. acetobutylicum glnA* gene was regulated by nitrogen levels in *E. coli* from its own regulatory region (Usdin *et al.*, 1986; Janssen *et al.*, 1988). Because transcription of the *C. acetobutylicum glnA* gene essentially originated from the same transcription initiation signals in *E. coli* and *C. acetobutylicum* cells, it was decided to study the regulation of the *glnA* gene entirely in *E. coli*. This facilitated additional work concerning *glnA* upstream deletion mutants and the study of their expression in *E. coli*, as well as the construction of *glnA::lacZ* fusions and *in trans* studies, as there was no suitable recombinant DNA system available for *C. acetobutylicum* until recently (Bertram and Dürre, 1989).

4.4.1.1. The involvement of *glnA* downstream sequences in the regulation of the *glnA* gene. In an earlier rapport (Janssen *et al.*, 1988), it was shown that deletion of *glnA* downstream sequences affected overall *C. acetobutylicum* GS activities in *E. coli* ET8051 (*glnALG*) cells.

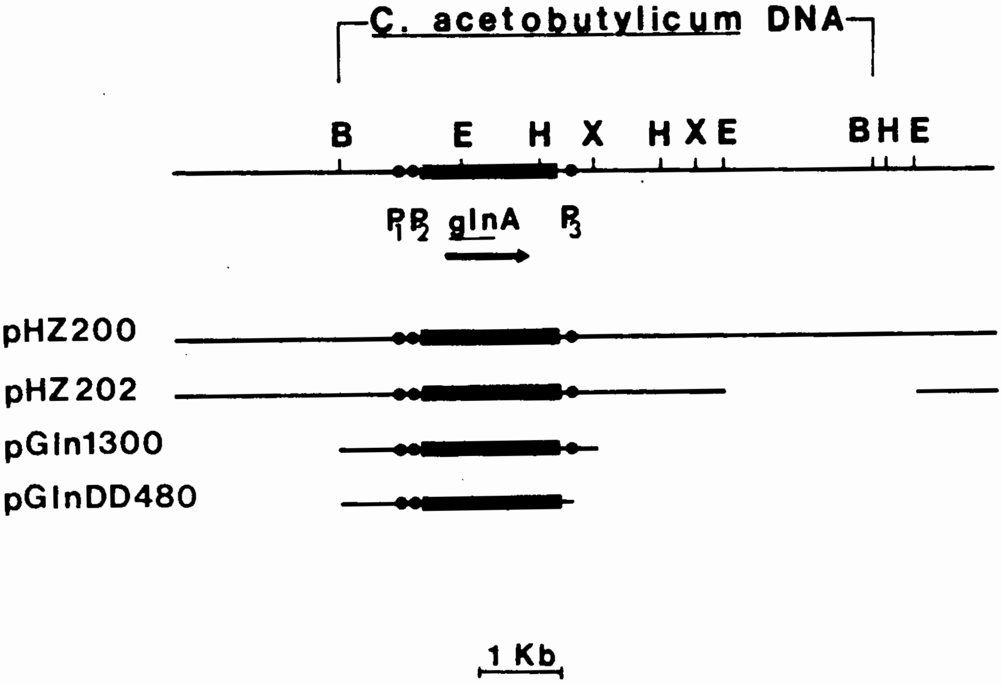


Fig. 4.2.: Deletion plasmids of pHZ200 (Usdin *et al.*, 1986) containing decreasing amounts of the *C. acetobutylicum glnA* downstream sequences. Solid circles represent the positions of the two *glnA* upstream promoters (Chapter 3) and the putative *glnA* downstream promoter p_3. The downstream region with an extensive stretch of inverted repeat sequences is situated between the end of the *glnA* structural gene (Heavy line) and the putative promoter p_3. Restriction endonuclease site abbrevations: B, *BglII*; E, *EcoRI*, H, *HindIII*; X, *XbaI*.

The following plasmids were used in this study (Fig. 4.1; Fig. 4.2): pHZ200, which contained a 6.3 kb DNA insert from *C. acetobutylicum* chromosomal DNA containing the *glnA* gene flanked by 1.3 kb upstream and 3.5 kb downstream sequences; pHZ202, an *EcoRI* deletion derivative of pHZ200 with approximately 1.7 kb *glnA* downstream sequences still present; pGln1300, previously named pJP3, which is a pLK34 derivative and contained 1.3 kb upstream and 614 bp downstream sequences; and pGlnDD480, previously named pJP3Δ 7.18, in which a further 480 bp was deleted and contained 134 bp the *glnA* downstream region. In constructing pGlnDD480, the downstream putative promoter p_3 was deleted and the inverted repeat sequences were disrupted (Fig. 2.3)

The effect of nitrogen levels on the expression of the cloned *C. acetobutylicum glnA* gene, contained on the deletion plasmids, in *E. coli*, was determined by comparing GS activity in cells grown in CSH medium containing 15 mM glutamate + 15 mM glutamine (nitrogen-excess) and CSH medium containing 15 mM glutamate + 0.15 mM glutamine (nitrogen-limiting) (Fig. 4.3).

Growth of *E. coli* ET8051 cells harbouring pHZ200, pHZ202, pGln1300, and pGlnDD480 was determined under nitrogen-excess and nitrogen-limiting conditions. Strains were first grown overnight in LB medium. The next day, cultures were harvested and washed with 1 x A minimal salts and resuspended (1:25 inocculum) in CSH minimal medium containing either 15 mM glutamate/15 mM glutamine or 15 mM glutamate/0.15 glutamine. Because cells were not aclimatised to minimal medium, strains grew very slowly. However, pGlnDD480 enabled the *E. coli* ET8051 *glnA* deletion mutant to grow extremely well in the nitrogen-limited minimal medium, and the cultures reached a maximal optical density at 600 nm (OD_{600}) of 2.8 units. *E. coli* K514 wild-type control cultures attained a maximum OD_{600} of 1.4 in the same medium, and the maximum OD_{600} of *E. coli* ET8051 cultures containing pHZ200, pHZ202, and pGln1300 was slightly less than that of the wild-type control.

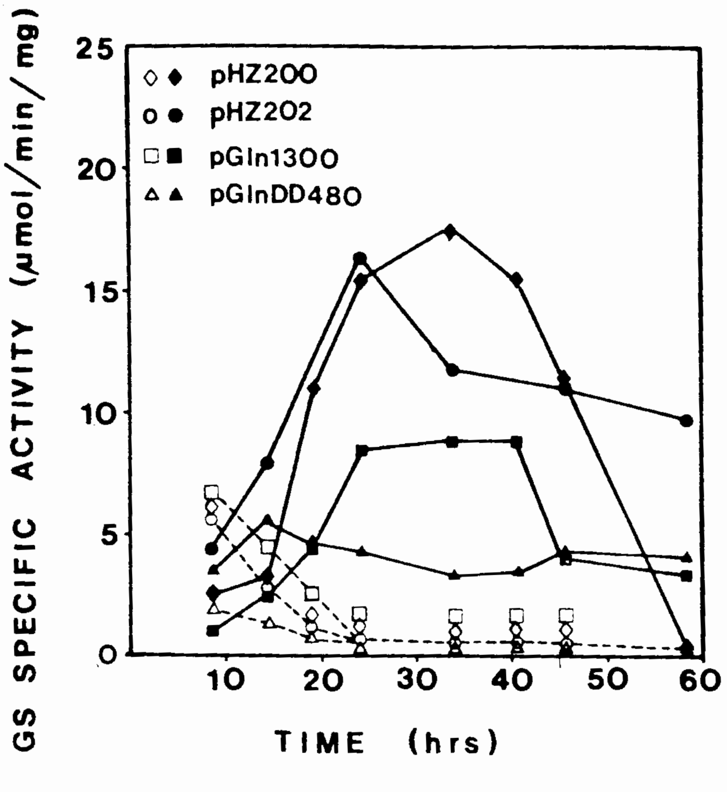


Fig. 4.3.: Production of GS in *E. coli glnALG* ET8051 cells by deletion derivatives of pHZ200 containing decreasing amounts of the *C. acetobutylicum glnA* downstream sequences. GS activity was measured by the γ-GT assay (Materials and methods) in cells grown under nitrogen-excess (open symbols) and nitrogen-limiting (closed symbols) conditions. GS activity was measured in ET8051 cells containing pHZ200 (◆), pHZ202 (●), pGln1300 (▲), and pGlnDD480 (■).

Under nitrogen- excess conditions, *E. coli* ET8051 (pGlnDD480) grew slightly better than the other cultures and reached a maximum OD_{600} of 4.85 units. In comparison, the other cultures reached an OD_{600} of between 3.5 and 4.5 units. Typical doubling times for cultures grown in nitrogen-excess and nitrogen-limiting conditions were approximately 6 h and 12 h, respectively.

The results given above indicate that DNA sequences downstream of the cloned *C. acetobutylicum glnA* structural gene have a regulatory role, since there was a correlation between the level of GS activity and the extent of the deletion. It was previously shown that the regulation of GS acivity by nitrogen levels in *E. coli* strains ET8051(pHZ200) and ET8051(pHZ202) was not due to changes in plasmid copy number since there was < 5% variation in b-lactamase levels after growth in nitrogen-excess and nitrogen-limiting media (Usdin *et al.*, 1986; K. Usdin, unpublished results). Likewise, it is not expected that pLK34 derived plasmids containing the *C. acetobutylicum glnA* gene undergo changes in copy number because the *bla ori* region of these plasmids is identical to that of the pHZ plasmids. However, it may be necessary in the near future to determine the copynumber of these plasmids, especially when *glnA* expression levels are compared with production levels of *glnA::lacZ* fusion protein (see further), because *glnA* and *glnA::lacZ* constructs are based on plasmids with different *ori* regions. Table 4.1 summarizes the expression of *C. acetobutylicum* GS in *E. coli* plasmid-carrying strains after 2.5 doubling times, when cells are in mid-exponentional growth.

Table 4.1.: Expression of the cloned *C. acetobutylicum glnA* gene contained on various *glnA* downstream deletion plasmids in *E. coli* ET8051.

Strain	GS activity^a		R^c
	15 mM gln^b	0.15 mM gln	
K514 (wild-type)	280	680	2.4
ET8051	0	0	0
ET8051(pHZ200)	2,600	15,500	6.0
ET8051(pHZ202)	2,000	12,000	6.0
ET8051(pGln1300)	4,000	7,400	1.8
ET8051(pGlnDD480)	2,000	3,800	1.9

(a) GS activities are given in γ-glutamyl hydroxymate per min per mg protein. (b) CSH media contained 15 mM glutamate and the indicated amount of glutamine (gln). (c) Repression ratios (R) were calculated by dividing the enzyme activity found in nitrogen-limited cultures by that in nitrogen-excess cultures. Represented values are the estimates from Fig. 4.2 at time points 15 h and 30 h for nitrogen-excess and nitrogen-limiting conditions, respectively. These time points were choosen because they represent approximately 2.5 doubling times for growing cells in both nitrogen-excess and nitrogen-limiting conditions (see text).

4.4.1.2. GS activity of glnA upstream deletion mutants. The upstream region of the *C. acetobutylicum glnA* gene is characterised by p_1 and p_2, a complex palindrome sequence (CPS) immediately upstream of p_1 and p_2 between nucleotides -130 to -195 (Fig. 4.4), and a 380 bp region between nucleotides -175 to -554 which had a very high DNA curvature score of 8.2 (Plaskon and Wartell, 1987; Chapter 2). This region also contained six copies of a direct repeat, with the consensus sequence 5'-ATATTGTAA-3' (Chapter 2). Similar 9 bp sequences (consensus 5'-ATATTGTTT-3') occur nine times in the *B. subtilis gltA/gltC* regulatory region

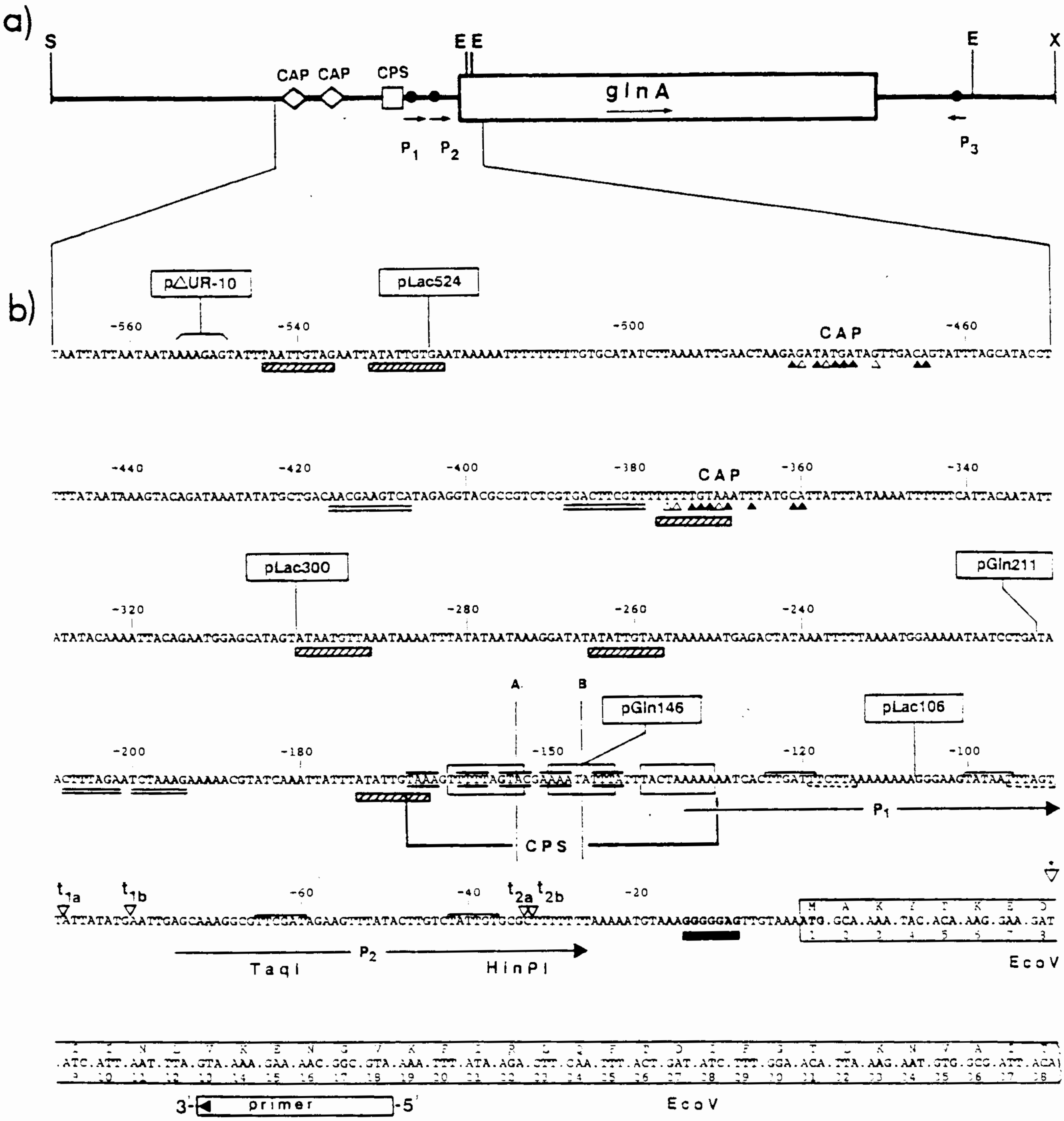
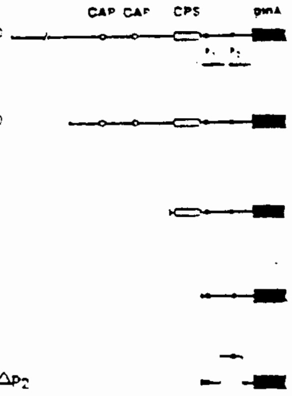


Fig. 4.4. Diagrammatic structure of pGln1300 (a) and nucleotide sequence of the upstream region of the *C. acetobutylicum glnA* gene (b). Deletion mutants and *glnA:lacZ* fusions are indicated. Promoters p_1 and p_2 are underlined with a full line. The putative promter $p_{1.1}$ is indicated with a dashed line. The fusion point between glnA and lacZ is demonstrated by an open triangle and a star above the first EcoRV site within the *glnA* coding region. A primer, used for sequence analysis and primer extension rections, is shown under the *glnA* coding sequence containing the primer binding site (codons 13 to 18). Relevant restriction sites are as in indicated in Fig. 4.1.

(Bohannon and Sonenshein, 1989), five times in the *B. cereus glnA/glnR* regulatory region (Nakano et al., 1989), three times in the *E. coli glnA* (Miranda-Rios et al., 1987) upstream region, and in the target DNAs of several LysR family proteins in *E. coli* and *R. trifolii* (Kölling and Lother, 1985; Schofield and Watson, 1986).

To investigate the effects of these regions on the regulation of the *C. acetobutylicum glnA* gene in *E. coli*, the production of GS by various *glnA* upstream deletion plasmids was determined in exponentional *E. coli* YMC11 *glnALG* cells in four different minimal media in which nitrogen was available in various forms (Fig. 4.5).

E. coli YMC11 cells containing pGln1300 and pGlnUR10, which contained approximately 1300 and 550 bp of the *glnA* upstream region, respectively, produced similar levels of GS in the two nitrogen-excess and the two nitrogen-limiting media (Fig. 4.5). GS production by YMC11(pGln1300) and YMC11(pGlnUR10) were regulated by nitrogen and in excess nitrogen showed similar repression ratios of 1.8 which were independent of the nature of the nitrogen source. Since the production of GS by pGln1300 and pGlnUR10 in different nitrogen-limiting and nitrogen-excess media was similar, it is concluded that the sites involved in the regulation by nitrogen are contained within a 550 bp region of the *C. acetobutylicum glnA* gene.



	15 mM glt 15 mM gln	15mM glt 0.15 mM gln	Repression ratio	4 mM gln 8 mM NH4	4 mM gln	Repression ratio
pGln1300	4,420	8,060	1.8	4,050	7,100	1.8
pΔUR-10	4,840	8,590	1.8	3,850	7,100	1.8
pGln211	2,910	4,920	1.7	2,180	2,630	1.2
pGln146	13,560	22,150	1.6	6,960	14,400	2.1
pGln146Δp_2	17,140	19,400	1.1	7,250	12,850	1.8

Fig. 4.5.: Production of GS by *C. acetobutylicum glnA* upstream deletion mutants in *E. coli* YMC11 *glnALG* cells in four different minimal media containing the following nitrogen sources: 15 mM glutamate + 15 mM glutamine; 15 mM glutamate + 0.15 mM glutamine; 4 mM glutamine + 8 mM $(NH_4)_2SO_4$; 4 mM glutamine. Putative CAP binding sites and the complex palindromic sequence (CPS) are indicated with open boxes. Promoters p_1 and p_2 are presented with closed circles. GS specific activity is expressed as nmole γ-glutamyl hydroxymate/mg/protein. Values reported are the average of at least two independent determinations. The repression ratios were calculated by dividing the enzyme activity found in nitrogen-limited cultures by that in nitroegn-excess conditions.

pGln211 was a deletion plasmid which contained 211 bp of the *glnA* upstream region (Fig. 4.4). Deletion of the 340 bp region, which had a high curvature score of 8.2 and contained the six direct repeats 5'–ATATTGTAA-3', resulted in a two-fold lower level of GS production in *E. coli* YMC11(pGln211) cells under nitrogen-limiting and nitrogen-excess conditions, but did not markedly affect the nitrogen repression ratio (Fig. 4.5). It appeared that this 380 bp region affected the overall level of GS activity but in the absence of this re/gion, the GS level was still regulated by nitrogen.

Construction of pGln146 which contained 146 bp of the *glnA* upstream region, resulted in the deletion of the majority of the CPS of the *glnA* upstream region but promoters p_1 and p_2 were not affected. *E. coli* YMC11(pGln146) cells showed an overproduction of GS in both nitrogen-limiting and nitrogen-excess media. However, GS levels produced by pGln146 in *E. coli* YMC11 were still repressed in nitrogen-excess media and nitrogen repression ratios varied between 1.6 and 2.1. The high GS levels produced by pGln146 indicated that the palindrome sequences in the *glnA* upstream region (between nt -130 and -195) were involved in the overall expression level of *C. acetobutylicum glnA* gene in *E. coli*. Similar results were obtained when plasmid YMC11(pGln146Δp2) cells were used (Figs. 4.4, 4.5), supporting the importance of the CPS in *glnA* gene expression. It is possible that these palindrome sequences represent a genetic signal in the form of DNA cruciforms. Such cruciforms are the result of the reorganisation of complementary base-pairing which goes through a transient unfolding from the original interstrand-bound duplex into two intrastrand-bound hairpin loop structures (Wada and Suyama, 1986). Because the CPS is in the vicinity of the -35 region of promoter p_1, it is tempting to speculate that differences in stacking energy between the interstrand-bond state and the intrastrand-bond state have an effect on p_1 promoter activity by modulating

RNA polymerase binding or open complex formation. This may require the action of a regulatory protein or additional features in the upstream sequence.

Although experiments involving pGln1300 derived *glnA* upstream deletion mutants are valid in respect of studies on gene expression, future experiments on *glnA* upstream deletion mutants should involve the construction of *glnA* upstream deletions derived from pHZ200. Unfortunately, no suitable restriction endonuclease sites are present within the 6 kb insert DNA of pHZ200. The entire insert of pHZ200, located on a 6 kb *BglII* fragment (Usdin *et al.*, 1986), could be subcloned into an *E. coli* vector with an appropriate multiple cloning site. Unidirectional deletion would then produce the required *glnA* upstream deletions. However, attempts to subclone the entire 6 kb *BglII* fragment of pHZ200 into pUC18 (Norrander *et al.*, 1983) failed (data not shown), possibly due to instability of the recombinant plasmids. An alternative approach to study the *glnA* upstream region in plasmid pHZ200 would be to perform saturated mutagenesis within the *glnA* regulatory region.

4.4.2 . The use of *glnA::lacZ* gene fusions to study the regulation of *glnA* gene expression.

4.4.2.1 Regulatory studies on the expression of *glnA::lacZ* usptream deletion mutants. A fusion plasmid, pLac1300, containing approximately 1300 bp of upstream region and the first eight codons of the *C. acetobutylicum* *glnA* gene fused to the nineth codon of *lacZ* (Materials and methods; Chapter 3) was constructed. This basic construct was used to produce a number of upstream deletions of which pLac524 contained 524 bp *glnA* upstream region (Fig. 4.4). *E. coli* (pLac524) cells showed similar levels of β-galactosidase activity as in *E. coli* (pLac1300) cells grown in nitrogen-limiting and -excess media (Table 4.2). This is in close agreement with the above results in which *E. coli* cells containing pGln1300 or pGlnUR10 (which contained approximately 550 bp of *the glnA* upstream region) were shown to produce similar levels of GS.

Table 4.2.: *E. coli* strain YMC11 transformed with various fusion plasmids and the expression of β-galactosidase measured in cells grown in the presence of the indicated nitrogen sources.

Fusion plasmid	Ngln(a)	gln(b)	R(c)
pLac1300	23.2	36.3	1.56
pLac524	21.3	35.5	1.67
pLac300	18.3	27.3	1.49
pLac146	7.0	12.6	1.80
pLac106	2.1	3.7	1.71

(a) nitrogen-excess CSH medium (Ngln; 8 mM ammonium sulphate, 4 mM glutamine).
(b) nitrogen-limiting CSH medium (gln; 4 mM glutamine).
(c) repression ratios (R) were calculated by dividing the respective enzyme activity found in gln cultures by the Ngln cultures. β-galactosidase levels are given in μmol.min^{-1} per mg protein.

Values reported are the averages of independent measurements of β-galactosidase activity on three samples during exponentional growth.

Although repression ratios were low, a relatively constant value of approximately 1.6 was observed (Table

4.2). Since only β-galactosidase activity was measured, it was unclear whether the higher β-galactosidase levels of the *glnA::lacZ* fusion plasmids in nitrogen-limiting conditions is due entirely to derepression of gene expression. Nevertheless, regulation of *glnA::lacZ* expression by nitrogen would be rather marginal. Overall expression levels of β-galactosidase, however, were affected when *glnA* upstream sequences were progressively removed (Table 4.2). Whereas β-galactosidase activities in *E. coli* (pLac300) cells were similar to b-galactosidase activities in *E. coli* (pLacl300) cells, removal of upstream sequences adjacent to promoter p_1 in pLacl46 resulted in a three-fold drop in β-galactosidase expression in *E. coli* (pGlnl46) cells. Deletion Of p_1 promoter sequences in pLacl06 (Fig. 4.4) resulted in very low expresion levels and a ten-fold drop in β-galactosidase activity in *E. coli* (pLacl06) cells. This is in agreement with primer-extension experiments performed in Chapter 3 in which it was concluded that regulation of the *glnA* gene at the level of transcription was mainly exerted via the repression/ derepression of p_1 promoter activity. Of interest is the decrease in β-galactosidase activity for *E. coli* YMCll (pLacl46) cells [compared to other pLac plasmid carrying cells] which contrasts with the elevated levels of GS production in *E. coli* YMCll (pGlnl46) cells (compared to other pGln plasmid carrying cells](*). It is difficult to explain this discrepancy in terms of gene regulation. It seems that *glnA* coding sequences or *glnA* downstream sequences, or both, are required for an increase in gene expression which was apparent in *E. coli* YMCll (pGlnl46) cells.

(*) Comparison of the gene expression levels between the pGln and pLac systems is not possible because the copynumber of the basic plasmids may differ.

E. coli YMC11 cells transformed with the various *glnA::lacZ* fusion plasmids were also grown in CSH minimal medium containing 15 mM glutamate supplemented with either 0.15 mM glutamine (nitrogen-limiting), or 15 mM glutamine (nitrogen-excess). In the absence of functional glutamine synthetase, *E. coli* YMC11(pLac) strains grew extremely slowly when grown in minimal medium containing 15 mM glutamate and 0.15 mM glutamine. Only residual growth was observed which came to a halt after a period of 12-15 h in which the cell density barely doubled. This is in contrast with the doubling times for CSH media containing 4

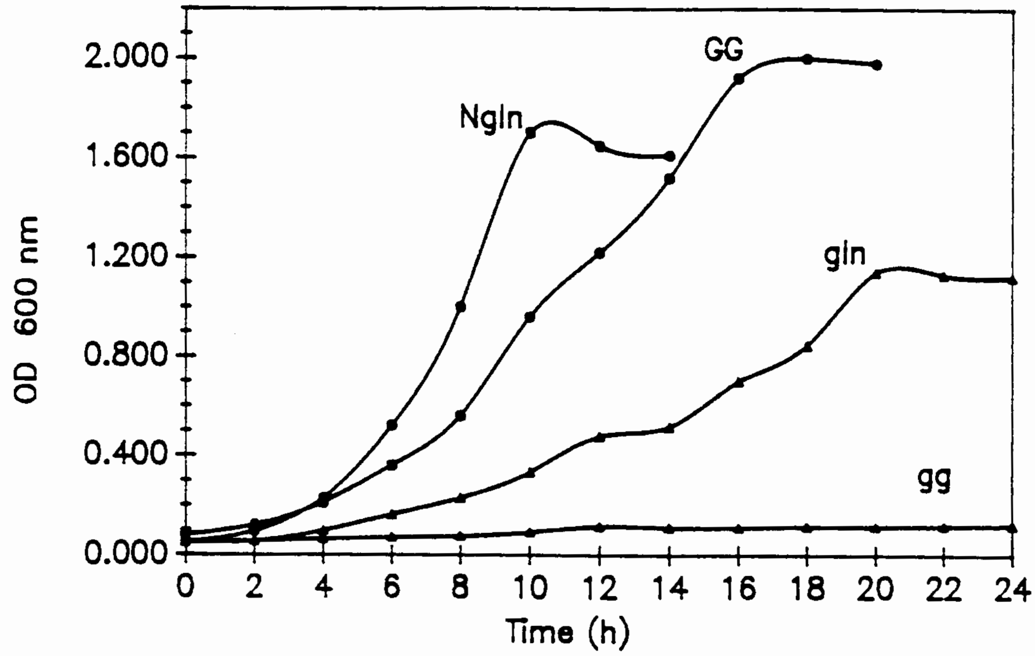


Fig. 4.6.: Characteristic growth curves for *E. coli* YMC11(pLac) strains grown under various nitrogen conditions. CSH media contained 15mM glutamate/15 mM glutamine (GG), 15 mM glutamate/0.15 mM glutamine (gg), 4 mM glutamine (gln), or 4m M glutamine/8mM $(NH_4)_2SO_4$ (Ngln).

mM glutamine (gln-medium) or 4 mM glutamine, 8mM $(NH_4)_2SO_4$ (Ngln-medium) with doubling times of 6.5 h and 2 h, respectively, and which sustained growth of *E. coli* YMC11(pLac) strains until the stationary stage was reached (Fig 4.6).

Although *E. coli* YMC11(pLac) cultures grown in CSH medium supplemented with 15 mM glutamate/15 mM glutamine grew nearly as well as *E. coli* YMC11(pLac) cultures grown in CSH-Ngln medium (Fig. 4.6), it was decided that regulatory experiments using CSH medium supplemented with glutamate (glt) and glutamine (gln) should be omitted because comparison of β-galactosidase levels between cultures grown under high (15 mM glt; 15 mM gln) and low (15 mM glt; 0.15 mM gln) nitrogen conditions was not possible.

4.4.2.2. *In trans* studies to investigate the regulatory role of GS in *glnA* gene expression. Regulatory studies *in trans* were performed to test wether *C. acetobutylicum* GS affected β-galactosidase expression levels from *glnA::lacZ* fusion plasmids (Table 4.3).

The *C. acetobutylicum glnA* gene on plasmid pTsGln1300 (Materials and methods) was placed *in trans* with pLac1300 by co-transformation of *E. coli* YMC11 with the two plasmids. The copy number of the pBR322-based *lacZ* fusion plasmid can be presumed to be ca. 18 copies/ *E. coli* genome (Covarrubias *et al.*, 1981) and that of the pAcYc177-based pTsGln1300 plasmid ca. 22 copies/ *E. coli* genome (Chang and Cohen, 1978). Since the *in trans* study was done in a multi-copy system, the *glnA* product is unlikely to be limiting in this experiment.

Table 4.3.: Expression of β-galactosidase and GS from plasmids pTsgln1300 and pLac1300 in *E. coli* YMC11 cells.

Plasmids	β-gal[a]			GS[b]		
	Ngln[c]	gln[d]	R[e]	Ngln	gln	R
pTsGln1300	-	-	-	2,000	5,150	2.6
pLac1300	23.4	39.9	1.7	-	-	-
pTsGln1300 pLac1300	24.0	40.1	1.7	1,480	5,080	3.4
pTsGln1300 pMC1871	<0.08	<0.08	-	2,100	5,120	2.4

(a) β-galactosidase levels were expressed in μmol.min^{-1} per mg protein.
(b) GS activity was measured as nmol.min^{-1} per mg protein.
(c) nitrogen-excess CSH medium (Ngln; 4 mM glutamine, 8 mM ammonium sulphate).
(d) nitrogen-limiting CSH medium (gln; 4 mM glutamine).
(e) Repression ratios (R) are calculated by dividing the respective enzyme activity found in gln cultures by the Ngln cultures.

Values reported are the average of two independent measurements. All samples were taken during mid-exponentional growth.

The presence of the *C. acetobutylicum* GS has no effect on the production of β-galactosidase in *E. coli* cells containing the *glnA::lacZ* fusion plasmid. It may be argued that due to the use of a multi-copy system, the expression of the *C. acetobutylicum glnA* gene is not subject to its normal regulation. From the above preliminary results, however, it would appear that the *C. acetobutylicum* GS has no direct role in regulation of its own expression.

A further experiment was designed to bring *glnA::lacZ* fusions *in trans* with various up- and downstream *glnA* deletion mutants. Restriction fragments carrying the *glnA::lacZ* gene fusions of pLac1300 and pLac524 were subcloned into pAcYc177, resulting in plasmids pTsLac1300 and pTsLac524, respectively. *E. coli* YMC11 cells containing plasmids pHZ200, pHZ202, pGln1300, pGln211, pGlnDD480 (a *glnA* downstream deletion derivative of pGln1300), and the control plasmid pEcGln (a pLK34 derivative containing the *E. coli glnA* gene), were co-transformed with either pTsLac1300 or pTsLac524 into *E. coli* strain YMC11. Growth of these strains was scored on LB medium and CSH medium, supplemented with 8 mM $(NH_4)_2SO_4$ (100 µg/mL Ap, 25 µg/mL Kn) (Table 4.4). Surprisingly, when pHZ202 was co-transformed with either pTsLac1300 or pTsLac524, no growth on LB or on CSH medium was observed, despite numerous attempts to co-transform these plasmids. Co-transformation of pHZ200 with the control plasmid pAcYc177, however, resulted in normal growth, both on LB as well as on CSH medium, when compared with growth of YMC11 transformed with pHZ200 only (Table 4.4).

Table 4.4.: *E. coli* YMC11 co-transformants, used for *in trans* studies, grown on CSH minimal medium supplemented with 8 mM $(NH_4)_2SO_4$ (identical results were obtained using LB medium).

Plasmids	none*(a)*	pAcYc177	pTsLcac1300	pTsLac524
pHZ200	+(b)	+(c)	-(c)	-(c)
pHZ202	++	++	++	++
pGln1300	++	++	++	++
pGln211	++	++	++	++
pGlnDD480	++	++	++	++

(a) Transformation of *glnA* plasmids only.
(b) Transformants (Ap^r) scored per ng plasmid: (-), no transformants; (+), 500 to 1200 transformants; (++), more than 1200 transformants.
(c) Co-transformants (Ap^r, Kn^r) were scored per ng plasmid: (-), no co-transformants; (+), 500 to 1200 co-transformants; (++), more than 1200 co-transformants.
Identical results were obtained in both LB and CSH media.

Apparently, pHZ200 *in trans* with either pTsLac1300 or pTsLac524 inhibits cell growth. A similar phenomenon was observed when the wildtype *E. coli* strain YMC10 was used as the recipient strain (results not shown). The *E. coli* strain YMC11(pHZ200) grew slowly and in general formed very small colonies on LB or CSH agar plates. It has been shown that the 3.5 kb *glnA* downstream region in plasmid pHZ200 plays an important role in the overall expression of the *C. acetobutylicum* GS in *E. coli* (Janssen *et al.*, 1988). When pHZ200 is co-transformed with pTsLac plasmids, *glnA* downstream sequences may affect expression of the *glnA::lacZ* fusion and possibly cause overproduction of the fusion product. Hence the observed toxic effect. An incompatibility problem between pHZ200 and pAcYc177-derived plasmids is not expected because the *E. coli* YMC11(pHZ200)(pAcYc177) strain displayed normal growth on both LB and CSH media.

E. coli YMC11(pTsLac524) cells were transformed with *glnA* upstream deletion mutants and β-galactosidase and GS activities were measured in mid-exponentional growth in CSH medium supplemented with the indicated nitrogen sources (Table 4.5.).

Table 4.5.: *E. coli* YMC11(pTsLac524) strain transformed with various *glnA* plasmids and the expression of β-galactosidase and GS in cells grown under nitrogen-limiting (gln) and nitrogen-excess (Ngln) conditions.

Plasmids	β-gal[a]			GS[b]		
	Ngln	gln	R[c]	Ngln	gln	R
pHZ202	20.1	61.1	3.0	2,650	12,130	4.6
pGln1300	22.3	49.5	2.2	3,700	7,300	2.0
pGln211	19.1	34.7	1.8	1,550	2,420	1.6
pGln146	12.8	41.8	3.3	6,120	14,300	2.3
pGlnDD480	22.3	30.7	1.4	5,280	7,700	1.5
pEcGln	18.9	35.0	1.9	2,100	1,950	0.9
none	18.7	33.4	1.8	-	-	-

(a) β-galactosidase acivities are given in μmol.min^{-1} per mg protein.
(b) GS levels are presented in nmol.min^{-1} per mg protein, as was determined by the γ-GT assay.
Values reported are the result of a single measurement during mid-exponentional growth.
(c) Repression ratios were calculated by dividing the respective enzyme activity found in glutamine cultures (gln) by the value for glutamine + $(NH_4)_2SO_4$ cultures (Ngln).

The data given in Table 4.5 are in agreement with the results presented in Table 4.3 in that no significant changes in β-galactosidase activity of *E. coli* cells containing the *glnA::lacZ* fusion plasmid pTsLac524 was observed when the *C. acetobutylicum glnA* gene was present *in trans*. The above data also agree with the GS levels presented in Table 4.2. It is concluded from the above results that the *C. acetobutylicum* GS *in trans* does not contribute significantly to the regulation of β-galactosidase levels in *E. coli* YMC11(pTsLac524) cells containing various *glnA* up- and downstream deletion plasmids and hence, most likely only has a minor role in the regulation of *glnA* gene expression in *E. coli*.

4.4.3. Regulation of *C. acetobutylicum glnA* transcription initiation sites in *E. coli*. The initiation sites of transcription for the *C. acetobutylicum glnA* gene have been determined in both *E. coli* and in *C. acetobutylicum* (Janssen *et al.*, 1989; Chapter 3). In total four major start points, t_{1a}, t_{1b}, t_{2a}, and t_{2b}, were identified for *E. coli* YMC11(pHZ200) cells grown in nitrogen-limiting medium. Transcripts t_{2a} and t_{2b}, located at positions -33 and -32, respectively, were controlled by the promoter p_2 with the sequence TTCGAT--N17--TATTGT and were also present in primer-extension experiments performed on mRNA extracted from *C. acetobutylicum* cells grown in Molasses Fermentation medium. Transcripts t_{1a} and t_{1b} originated from the p_1 promoter region at positions -88 and

-80, respectively. This promoter, with the consensus sequence TTGATT--N18--TATAAT, is overlapped by the CPS, thought to be a regulatory element important in the regulation of *glnA* gene expression (see above). Because transcript t_{1b} was not observed in *C. acetobutylicum,* it was suggested that this transcript in *E. coli* might be the product of the abberant transcription from promoter $p_{1.1}$ which overlapped promoter p_1 in a 5 bp-shift (Chapter 3).

As has been shown from the data presented in Fig. 4.1 and Table 4.1., *glnA* downstream sequences are important for the expression of the *C. acetobutylicum glnA* gene in *E. coli*. To test whether differences between GS activities of *E. coli* (pHZ200) and *E. coli* (pGln1300) cells were due to changes in transcriptional control, it was decided to perform primer-extension experiments on *E. coli* (pGln1300) cells (Fig. 4.7). Primer extension experiments were carried out on RNA extracted from exponentional *E. coli* YMC11 *glnALG* cells containing pHZ200 or pGln1300, which were first pre-grown in minimal (nitrogen-excess) medium and then grown under nitrogen-limiting (15 mM glutamate/0.15 mM glutamine) and nitrogen-excess (15 mM glutamate/15 mM glutamine) conditions for 16 and 6 h, respectively. These nitrogen conditions, which previously have been used by others (Usdin *et al.*, 1986; R. Maharaj, Ph. D. Thesis 1988 , University of Cape Town, South Africa), were the same conditions used in experiments described in Chapter 3 and in the experiments which resulted in data represented in Table 4.1.

The results of the primer extension experiments with exponentional phase *E coli* YMC11 *glnALG* (pGln1300) cells grown under nitrogen-limiting conditions were similar to those with *E. coli* YMC11 (pHZ200) cells. However, primer extension experiments showed an important difference between pHZ200 and pGln1300. Transcription initiation from t_{1a}, t_{1b}, t_{2a}, and t_{2b} in exponentional *E. coli* cells containing pHZ200 was markedly repressed by excess nitrogen, whereas

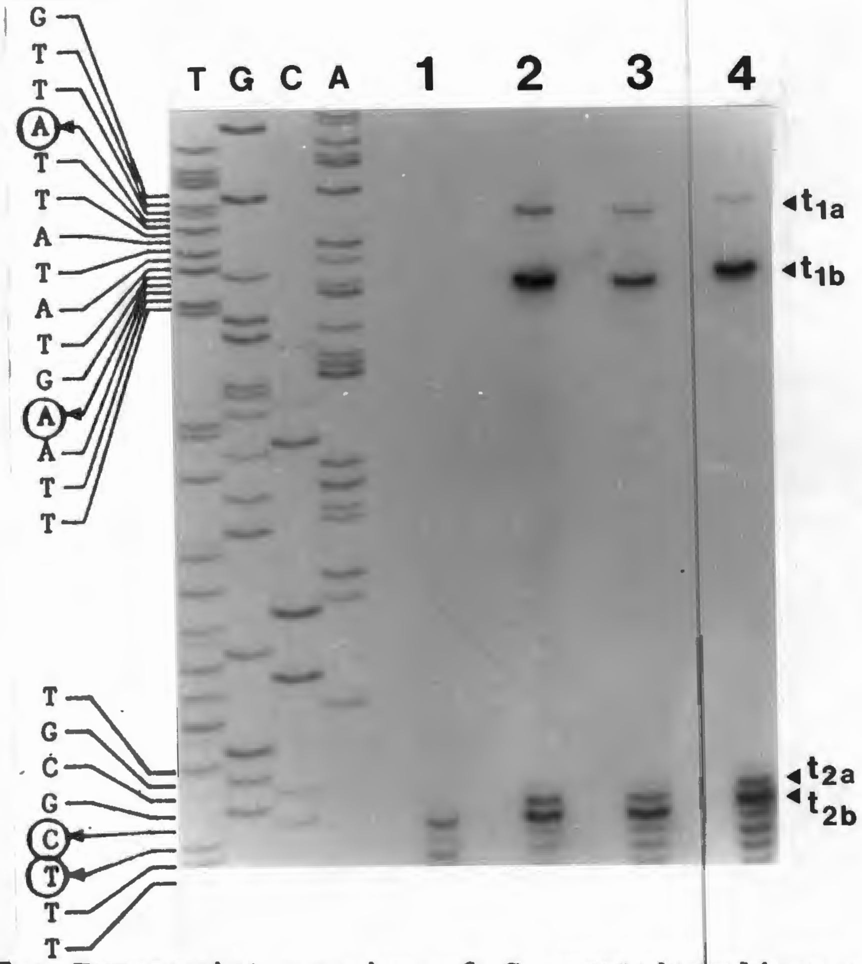


Fig. 4.7.: Transcript mapping of *C. acetobutylicum glnA* mRNA in *E. coli*. The 5′ ends of the *glnA* transcripts were determined by primer extension (Materials and methods; Chapter 3). Lanes 1 and 2 contained mRNA from *E. coli* YMC11(pHZ200) cells. Lanes 3 and 4 contained mRNA from E. coli YMC11(pGln1300) cells. *E. coli* cells were grown in CSH glucose minimal medium containing 15 mM glutamate and either 15 mM glutamine (Lanes 1 and 3) or 0.15 mM glutamine (Lanes 2 and 4). Sequencing ladders (T, G, C, A, from the left) were obtained using the same primer as for transcript mapping and were used as control. Transcripts t_{1a}, t_{1b}, t_{2a}, and t_{2b}, are indicated. Part of the glnA promoter proximal region is shown with the transcription start points.

there was little repression of the initiation of transcription from t_{1a}, t_{1b}, t_{2a}, and t_{2b} in *E. coli* cells containing pGln1300 (Fig. 4.7). This suggests that there may be a region of the downstream sequence of the *C. acetobtylicum glnA* gene which is involved in the control of transcription in *E. coli* from t_{2a} and t_{2b} and in particular from t_{1a} and t_{1b}. In addition,transcript t_{1a}, which was shown to be the major transcript in *C. acetobutylicum* cells transcribed from the p_1 promoter region (Fig. 3.2), was more abundant in *E. coli* YMC11 (pGln1300) cells grown in nitrogen-excess medium than when grown on nitrogen-limiting medium.

At the same time as mRNA was isolated from exponentional *E. coli* (pHZ200) and *E. coli* (pGln1300) cells for the primer extension experiments, the GS activity of the cells grown in nitrogen-limiting and nitrogen-excess media was determined. *E. coli* (pHZ200) cells grown in nitrogen-limiting and nitrogen-excess media had 15.0 and 4.5 μmol of τ-glutamyl hydroxymate/mg of protein per min respectively with a nitrogen repression ratio of 3.3. *E. coli* (pGln1300) cells grown in nitrogen-limiting and nitrogen-excess media had 13.8 and 5.1 μmol of τ-glutamyl hydroxymate/mg of protein per min respectively with a nitrogen repression ratio of 2.7. The GS levels and nitrogen repression ratios obtained with *E. coli* (pHZ200) and *E. coli* (pGln1300) cells did not correlate with the results of the primer extension experiments on *E. coli* (pHZ200) and *E. coli* (pGln1300) cells grown under the same nitrogen-limiting and nitrogen-excess conditions. These differences between the regulation of transcription and the levels of GS activity could be due to regulatory mechanisms operating at the level of translation. The presence of an functional *glnA* downstream promoter sequence (p_3) which controlled the transcription of a putative antisense RNA, complementary to 43 bases overlapping the ribosome binding site of the *C. acetobutylicum glnA* mRNA, has been reported (Janssen *et al.*, 1988). The action of such an antisense RNA would affect GS activity, either by inhibition of translation efficiency or by changing mRNA stability (Chapter 1), without affecting the initiation of transcription of the *C. acetobutylicum glnA* gene.

Differences in GS activity were also observed in *E. coli* (pHZ200) and *E. coli* (pGln1300) cells assayed in exponential or stationary growth phase (Fig. 4.3; Janssen *et al.*, 1988). *E. coli* cells containing pGlnDD480, however, showed a relatively constant level of GS activity throughout the exponentional and stationary growth phases. Plasmid pGlnDD480 is derived from pGln1300 in which a 480 bp *glnA* downstream sequences, including the p_3 promoter and adjacent nucleotides, were deleted. Preliminary results identified antisense RNA transcripts from the *glnA* downstream promoter p_3 in *E. coli* (pGln1300) and *C. acetobutylicum* cells, and indicated that the expression of this downstream promoter varied with the growth phase (see Chapter 5).

Because deletion of p_3 and adjacent nucleotides in pGlnDD480 also destroyed part of the extended palindrome sequence immediately downstream of the *glnA* gene, it was necessary to investigate the role of antisense RNA production in the regulation of the *C. acetobutylicum glnA* gene by site-directed mutagenesis of the p_3 promoter proximal region. This was done by the use of a novel system based on the gapped duplex approach (Kramer *et al.*, 1984) and a set of newly constructed phasmids (Stanssens *et al.*, 1989).

Chapter 5: The involvement of antisense RNA in the regulation of the *C. acetobutylicum glnA* gene in *E. coli*.

Chapter 5

The involvement of antisense RNA in the regulation of the *C. acetobutylicum glnA* gene in *E. coli*.

5.1. Summary

Antisense RNA produced from the *glnA* downstream promoter p_3 was mapped by primer extension on RNA extracted from *E. coli* YMC11 *glnALG* (pGln1300) cells grown in nitrogen-excess and nitrogen-limiting media. In parallel, primer extension experiments were also performed on RNA extracted from *C. acetobutylicum* cells grown in Molasses Fermentation medium with or without the addition of $(NH_4)_2SO_4$. Two antisense RNA transcripts were detected in both *E. coli* and *C. acetobutylicum* cells of which the 5′-end points were mapped at two adjacent adenine nucleotides immediately following the presumptive -35 region TATAAT of promoter p_3. It was further shown that levels of transcription from p_3 were not affected by nitrogen availability in *E. coli* or *C. acetobutylicum* cells.

Promoter p_3 was cloned in front of the chloramphenicol acetyl transferase (*cat*) gene in the *B. subtilis* promoter probe vector pPL603b1, and in front of the tetracycline resistance (*tet*) gene of the *E. coli* vector pBR322. The latter construct was used to monitor expression from p_3 during growth in a wildtype *E. coli* strain by determination of the tetracycline resistance, which was calculated as the MIC. It was shown that expression levels from p_3 varied with the growth phase.

A double mutation within the -35 region of the p_3 promoter was introduced by site directed mutagenesis and resulted in the alteration of the sequence in this region from TTGCTA to AAGCTA. Although it was expected that the mutation TT to AA would abolish promoter activity, the opposite occurred. Expression of the mutant promoter cloned in front of the *tet* gene in pBR322 resulted in a higher MIC value than for pBR322 or pBR322 containing the wildtype p_3 promoter. When the wildtype p3 promoter in pGln1300 was replaced by the mutant p_3 promoter sequence, lower levels of GS were measured. This is in accordance with the proposed inhibitory role of antisense RNA in the regulation of *glnA* gene expression.

5.1. Introduction.

In the past 20 years, gene expression has been shown to be controlled in both prokaryotes and eukaryotes by regulatory gene products such as activator and repressor proteins. More recently, naturally occurring antisense RNA molecules have been found to directly regulate gene expression (reviewed by Simons, 1988; see Chapter 1). In most cases, antisense and target RNAs are transcribed from opposite strands of the same template by opposing promoters, as is the case for the *E. coli crp* operon (Okamoto and Freundlich, 1986). However, a gene coding for antisense RNA may be unlinked with the target gene. A well known example is the *micF* gene, wich is separated from its target gene, *ompF*, by approximately 1000 kb (Aiba *et al.*, 1987). More recently, Carter-Münchau and Wolf (1989) reported growth-rate-dependent regulation of the *E. coli gnd* gene by an internal complementary sequence (ICS). This element was located within the *gnd* coding region and, because of its complementarity to the *gnd* ribosome-binding site, was able to sequester the ribosome-binding site in an mRNA secondary structure thereby inhibiting translation. Long-range transcriptional coupling has also been reported for the *rplJ* (Petersen, 1989) and the *EcoRV* (J. Botterman, Plant Genetic Systems, Ghent, Belgium; personal communication) genes in *E. coli*.

Since a 43 nt region downstream of the *C. acetobutylicum* gene showed extensive homology with a 43 nt region situated at the start of the *glnA* gene, including the Shine-Dalgarno sequence and the AUG start codon, it was of interest to research the possibilities concerning the presence of antisense RNA. As mentioned before (Chapters 2 and 4; Janssen *et al.*, 1988), a putative promoter (p_3), located downstream of the *C. acetobutylicum glnA* gene and orientated towards the caboxy-terminal part of *glnA* on the opposite strand, was detected by analysis of sequence data. It was also shown that this promoter was functional in *B. subtilis* by cloning a DNA fragment containing p_3 into a promoter probe vector (Janssen *et al.*, 1988). In view to this, it was suggested that the proposed promoter p_3 possibly initiated transcription of an antisense RNA which could form a strong RNA:RNA hybrid structure with the 5′ region of the *glnA* mRNA (ΔG value of -41.8 kcal/mol) (Salser *et al.*, 1977). To test this hypothesis, it was important to identify antisense RNA by primer-extension reactions performed on *E. coli* (pGln1300) and *C. acetobutylicum* cells under nitrogen-excess and nitrogen-limiting conditions.

Removal of p_3 in the downstream deletion mutant pGlnDD480 (Chapter 4) also destroyed part of the 158-bp stretch of inverted repeat sequences immediately following the *glnA* coding region. mRNA transcribed from this region has the potential to form a number of stem-loop structures (Chapter 2) which may be involved in either mRNA termination or stabilisation (for review, see Chapter 1). Because this could affect *glnA* gene expression, it was necessary to apply site-directed mutagenesis on the p_3 promoter proximal region. In this way an attempt was made to elucidate the involvement of the putative antisense RNA in the regulation of the *C. acetobutylicum glnA* gene in *E. coli* without disturbing other putative regulatory sequences.

5.3. Materials and methods.

5.3.1. Strains, plasmids, phasmids, and phages. *E. coli* strain K514 (Zabeau *et al.*, 1982), routinely used as the recepient strain to perform general DNA recombinant work, and *E. coli* strain YMC10, used in tetracycline (Tc) resistance determinations, have been described in Chapter 2. The *E. coli* strain YMC11 *glnALG* was used as a recipient for pGln constructs. *B. subtilis* strain 1A46 (*pro-1, trpC2, thr-5, recE4*) was used as a recepient strain for the promoter probe vector pPL603b1 and its derivatives and was obtained from the Bacillus Genetic Stock Center (USA). The *C. acetobutylicum* strain P262 has been described previously (Jones *et al.*, 1982) and was obtained from National Chemical Products (South Africa). Plasmids pGln1300 and pGlnDD480 have been described elsewhere (Chapters 2 and 4) and were formerly called pJP3 and pJP3Δ7.18, respectively (Janssen *et al.*, 1988). Vectors pBR322 (Bolivar *et al.*, 1977; Sutcliffe, 1978) and pUC13 (Messing, 1983) were used for general cloning procedures and their endonuclease restriction maps are presented in Appendix D. Plasmid pJP4 (Appendix D) is a pUC13 derivative containing the *C. acetobutylicum glnA* downstream region derived from a *HindIII-XbaI* fragment of pGln1300, and was used as an intermediate for the construction of plasmid pWTpal. Plasmid pGln1300/m carries a double mutation in the downstream promoter p_3. Phasmids (plasmid/phage hybrids) pMa5-8 and pMc5-8 were used in site-directed mutagenesis experiments and were a kind gift of Dr. P. Stanssens, Plant Genetic Systems, Ghent, Belgium. These vectors contain, in addition to a ColE1-type origin of replication, the origin of replication of filamentous phage f1. This makes it possible to package one specific strand of the phasmids into phage rods. Phasmid pMa5-8 confers resistance to ampicilin (Ap) and contains an amber codon in the *cat* gene. The complementary vector pMc5-8 confers resistance to chloramphenicol (Cm) and contains an amber mutation in the *bla* gene. With this 'twin vector system' (pMa/c), a single mutagenesis round not only yields the desired mutation, but simultaneously leads to the aquisition of Ap or Cm resistance and introduction of an amber-marker in the alternative resistance gene. The latter feature permits selection against the template strand in the following round of mutagenesis. Thus, the vectors allow sequential introduction of a series of mutations during multiple rounds of mutagenesis involving alternate selection for resistance to Ap or Cm (Stanssens *et al.*, 1989). The amber mutation in the *cat* gene (pMa5-8; Ap^R, Cm^S) destroys the unique *PvuII* site, and the amber mutation in the *bla* gene (pMc5-8; Ap^S, Cm^R) destroys one of the two *ScaI* sites. Therefore, a restriction digestion may be carried out to test the configuration of a specific mutant. Phasmid pMaWT was a pMa5-8 derivative containing the wildtype p_3 promoter region and was used to construct gapped duplex DNA (gpDNA). Phasmid pMcMut, the resultant of one round of mutagenesis, was a pMc5-8 derivative containing the mutant p_3 promoter and was used as a source for plasmid constructions. Plasmids pWTpal and pMUTpal were pBR322 derivatives in which the *tet* promoter was replaced by the wildtype and mutant p_3 promoter regions, respectively. Deletion of the direct repeats situated between p_3 and the 3′ end of the *glnA* coding region in plasmids pWTpal and pMUTpal resulted in the recombinant plasmids pWT and pMUT, respectively. Strains YMC10 and YMC11 were used as described previously. Strain WK6 (Δ*[lac-proAB*], *galE*, *strA*/ F′ *lacI*^{q}, *Z*Δ*M15*, *proA*$^+$*B*$^+$) was used as a recipient strain for pMa/c phasmid vectors and their derivatives. Elimination of progeny deriving from the template strand required the mismatch repair deficient strain WK6mutS (Δ[*lac-proAB*], *galE*, *strA*, *mutS*::Tn10/ F′ *lacI*^{q}, *Z*Δ*M15*, proA$^+$B$^+$), which is unable to suppress amber mutations. Strains QD5003 (Hfr, *supF*) and K514 (*thr1*, *leu6*, *thi1*, *supE44*, *lacY1*, *tonA21*, r^-m^+) were used to check the resistance levels of the respective amber mutations. All plasmids and strains are listed in Appendix A. Filamentous phage M13KO7 (Pharmacia, Sweden) was used as a helper phage to infect phasmid containing *E. coli* strains in order to obtain single stranded (ss) phasmid DNA. The use of M13KO7 has the advantage that routinely high yields of ss phasmid DNA was observed (an ss phasmid to phage DNA ratio of approximately 50:1), which improves subsequent purification of ss phasmid DNA.

5.3.2. Media and buffers. All *E. coli* strains were grown in LB medium or on LB agar (1.5 % w/v) plates as described in Chapter 2. Strains WK6 and WK6mutS were regularly streaked on M9 minimal medium plates (Miller,

1972) to select for the *pro* marker present on the F′-episome. The antibiotics Ap, Cm, and Tc were used at concentrations of 100 μg/mL and 25 μg/mL, respectively. *C. acetobutylicum* strain P262 was grown on MF medium and B. *subtilis* strain 1A46 was grown on Brain Heart Infusion Broth (BHIB; Merck), Tryptose Blood Agar Base (TBAB; Oxoid), or on DM3 protoplast regeneration agar base (Chang and Cohen, 1979). CSH mimimal medium has been described in Chapter 3 and was used to grow *E. coli* strain YMC11 containing pGln1300 in nitrogen-excess (15 mM glutamate, 15 mM glutamine) or nitrogen-limiting (15 mM glutate, 0.15 glutamine) conditions. All buffers and media are listed in Appendix B.

5.3.3. Growth conditions. As recommended by the supplier, infection of phasmid containing *E. coli* strains with M13KO7 was carried out with very good aeration to ensure a high yield of ss phasmid DNA. Growth of *E. coli* strains containing various plasmids or phasmids was carried out as described in Chapters 2 an 3. *E. coli* and *B. subtilis* strains were grown at 37°C unless otherwise stated in the text. *C. acetobutylicum* strain P262 was grown in MF medium (MFM; Appendix B) as described in Chapter 3.

5.3.4. Plasmid and phasmid constructions. Plasmid pGln1300 (section 2.3.4.; Janssen *et al.*, 1988) was used as the source for new constructs. For the construction of pCSGAp24, a 570 bp *Sau3A* fragment of pGln1300 containing the *glnA* downstream promoter p_3 was cloned in the *BamHI* site of the promoter probe vector pPL603b1, which is a derivative of pPL603 (Williams *et al.*, 1981), which lacks a promoter functional during late exponentional growth and contains a short *EcoRI* linker containing a *BamHI* site (Corfield *et al.*, 1984; V. Corfield, personal communication). After ligation and transformation in *B. subtilis* strain 1A46, recombinants were selected on DM3 agar plates supplemented with Cm (10 μg/mL). Plasmid pJP4 was constructed by subcloning the *C. acetobutylicum glnA* downstream region, contained on the 728 bp *HindIII-XbaI* fragment of pGln1300, into the vector pUC13. Phasmid pMaWT was constructed by subcloning the same *HindIII-XbaI* fragment into phasmid pMa5-8 (Fig. 5.1). Phasmid pMcMut was the result of one mutagenesis round on plasmid pMaWT (see section 5.3.9.). *E. coli* strain WK6 transformed with pMcMut conferred Cm resistance (25 μg/mL) but, as was expected, failed to grow on LB agar plates supplemented with 100 μg/mL Ap (section 5.3.1.). Phasmid construction pMcMut was also tested by a *ScaI* as well as a *PvuII* endonuclease digest. Plasmid pGln1300\m was constructed by the exchange of the *HindIII-XbaI* fragment from pGln1300 (wild type p_3 promoter) with the *HindIII-XbaI* fragment of pMcMut. Using the intermediate pJP4, which has a *EcoRI* site (situated in the pUC13 polylinker region) adjacent to the *XbaI* site (Appendix D), a recombinant plasmid pWTpal, in which the *E. coli tet* gene of pBR322 was placed under control of the *C. acetobutylicum glnA* downstream promoter p_3, was constructed. This was done by subcloning the *EcoRI-HindIII* fragment of pJP4 into pBR322, which was digested with endonucleases *EcoRI* and *HindIII*, thereby deleting the entry and core regions of the pBR322 *tet* promoter sequence. Consequently, the resulting plasmid, pWTpal, contained promoter p_3 on the *XbaI-HindIII* fragment originally derived from pGln1300. Plasmid pMUTpal was constructed by the exchange of the *XbaI-HindIII* fragment of pWTpal with the *XbaI-HindIII* fragment of pMcMut (Fig. 5.1). All recombinants were selected on LB agar plates supplemented with 100 μg/mL Ap. Unexpectedly, *E. coli* strains K514 (pWTpa)l and K514 (pMUTpal) failed to grow on LB agar plates supplemented with 100 μg/mL Ap + 15 μg/mL Tc, indicating that the *tet* gene was not expressed from promoter p_3. This is in contrast with the results obtained with plasmid pCAGSp24 in *B. subtilis*, in which the *cat* gene was found to be expressed under control of promoter p_3 (see above). An extensive stretch of inverted repeat sequences was located between the 3′-end of the *C. acetobutylicum glnA* gene and the downstream promoter p_3. It has been suggested that this region has the potential to form secondary structures when transcribed into RNA and could play a role in either *glnA* mRNA termination or stabilization (Chapter 2; Janssen *et al.*, 1988). Because these inverted repeat sequences are still present between the p3 promoter and the *tet* gene in both plasmids pWTpal and PMUTpal, it was decided to test wether putative secondary structures in the RNA transcribed from p_3 might hinder expression of the *tet* gene. This was done by deletion of a unique 182 bp *Sau96I* fragment, located between p3 and the *tet* gene and containing most of the inverted repeat sequences (Fig. 2.3; Appendix E). The *XbaI-HindIII* fragment of pWTpal was purified over

LMA, followed by digestion with endonuclease *Sau96I*. This mixture of restriction fragments was used in a ligation reaction with pWTpal opened with endonucleases *XbaI* and *HindIII* (of which the vector fragment was purified over LMA). Transformation of the *E. coli* strain K514 with this ligation mixture resulted in Ap Tc resistant (Ap^R, Tc^R) transformants when scored on LB agar plates supplemented with 100 μg/mL Ap + 15 μg/mL Tc. Plasmid DNA was isolated from a number of Ap^R Tc^R transformants and tested for deletion of the *Sau96I* fragment by digestion with endonucleases *XbaI* and *HindIII*. One of these isolates, which lacked the internal 182 bp *Sau96I* fragment, was called pWT and choosen for further study. In a similar way plasmid pMUT was constructed. The 728 bp *XbaI-HindIII* fragment of pGln1300 was substituted by the 546 bp *XbaI-HindIII* fragment of either pWT or pMUT which resulted in the construction of pGlnWT and pGlnMUT, respectively. These plasmids were used to test wether deletion of the *glnA* downstream region containing the inverted repeat sequences effected *glnA* gene expression (see text). The genealogy of the above plasmid constructions is presented in Fig.5.1. All constructs were verified by detailed restriction endonuclease digests and/or DNA sequence analysis. Procedures for restriction endonuclease digestion, ligation, fill-in reactions, and gel-electrophoresis, have been described previously (Maniatis *et al.*, 1982). All other basic techniques have been listed in Appendix C.

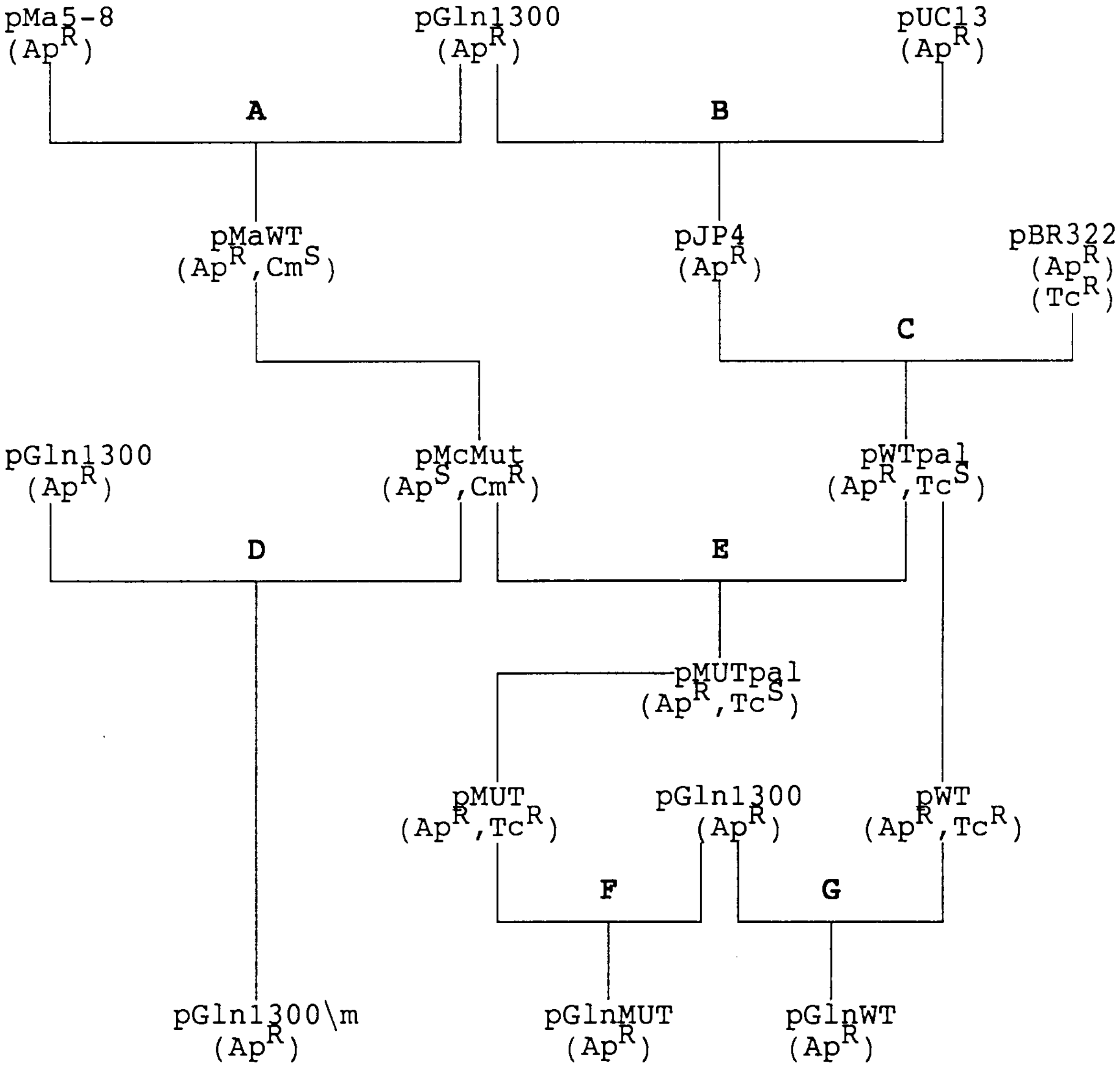


Fig. 5.1.: Genealogy of some constructs used in this study. A 728 bp *XbaI-HindIII* fragment of pGln1300, which contained the *glnA* downstream promoter p_3, was subcloned in the phasmid vector pMa5-8 (**A**) and into plasmid vector pUC13 (**B**). This resulted in phasmid pMaWT and the intermediate plasmid pJP4. Construct pMcMut was the result of one mutagenesis round on phasmid pMaWT whereas plasmid pWTpal was constructed by the exchange of the *tet* promoter in pBR322 by the *glnA* downstream promoter p_3 (see text) (**C**). Plasmids pGln1300\m and pMUTpal contained the mutant p_3 promoter which originated from pMcMut (**D** and **E** respectively). Constructs pMUT and pWT are derivatives of pMUTpal and pWTpal in which an internal *Sau96I* fragment, containing inverted repeat sequences, was deleted. Finally, pGlnMUT and pGlnWT were pGln1300 derivatives in which the *XbaI-HindIII* fragment was exchanged with the *XbaI-HindIII* fragments of pMUT (**F**) and pWT (**G**), respectively. Antibiotic resistance (R) and sensitivity (S) has been indicated (Ap, 100 μg/mL; Tc, 15 μg/mL; Cm, 25 μg/mL).

5.3.5. Transformation procedures. Transformation of *E. coli* and *B. subtilis* cells was performed as described in Appendix C. Transformation of *E. coli* strains WK6 and WK6mutS in site-directed mutagenesis procedures was performed as described below (section 5.3.9.).

5.3.6. Preparation of nucleic acids. RNA was prepared from *E. coli* and *C. acetobutylicum* strains as described in Chapter 3 and Appendix C. Procedures to extract plasmid DNA from either *E. coli* or *B. subtilis* cells have been outlined in Appendix C. Double stranded (ds) phasmid DNA was prepared according to the standard techniques used to prepare plasmid DNA (Appendix C). Single stranded (ss) phasmid DNA was produced as follows (Stanssens *et al.*, 1989). *E. coli* strain WK6 harbouring the vector or recombinant phasmid was grown overnight in LB medium (supplemented with 100 μg/mL Ap or 25 μg/mL), diluted 1:50 with LB (without antibiotic), and further grown at 37°C to a density of approximately 10^8 cells/mL. Cells (10 mL) were infected at a m.o.i of 20 with the helper phage M13KO7 (titer of 5.2 x 10^{10} pfu/mL), which supplies *in trans* all the products required for phasmid-directed ss DNA synthesis and packaging. The infected bacteria were incubated with agitation at 37°C for another 4-5 h period whereafter the culture was clarified by centrifugation (Sorvall HB4/ 10 000 rpm/ 10 min). PEG solution (20% polyethyleneglycol 6000; 2.5 M NaCl) was added (0.25 mL/mL supernatant) and phages were precipitated by cooling on ice for 15 min and subsequent centrifugation (Sorvall/ 8 000 rpm/ 10 min). The phage pellet was resuspended in 0.3 mL TE-Buffer (Appendix B) and purified with phenol and chloroform:isoamylalcohol extractions after which the nucleic acids were precipitated with $NaClO_4$/ isopropanol (Appendix C). DNA pellets were dissolved in 50 μL TE-buffer and stored at -20°C. The yield of ss DNA was determined photometrically ($\epsilon_{260} = 2.86 \times 10^{-2}$ $cm^2/\mu g$; 1 OD = 35 μg/mL). About 1.5-5 μg/mL culture was routinely obtained. The ratio of ss phasmid over ss helper phage DNA, as well as the total yield, was estimated by analysis of the ss DNA on a 1% agarose gel.

5.3.7. Specifications of the primer and mutagenic oligonucleotide. Primer-extension reactions on RNA extracted from *E. coli* or *C. acetobutylicum* cells required primer#2 (Chapter 2), which had a sequence complementary with a region 178 to 195 bp downstream of the *glnA* gene (Fig. 2.3). This region is situated 31 nt from the TATAAT Pribnow box of the putative promoter p_3. Primer#2 was also used to perform DNA sequence analysis on constructs containing the mutant p3 promoter. The mutagenic oligunucleotide used in site directed mutagenesis, which had the sequence 5'-GTT.TTT.AAT.TTT.T$\underline{AA}$.GCT.ATC.TC-3'(underlined residues differ from the wild type sequence), and which was partially complementary with a region containing the entry region TTGCTA of promoter p_3 (Fig. 2.3), was a gift of D. Botes, Department of Biochemistry, University of Cape Town, South Africa. This sequence contained an *AluI* site (AGCT) which was used for preliminary testing of the introduced mutation by an *AluI* endonuclease digest of constructs containing the mutant p_3 promoter.

5.3.8. Primer-extension reactions. The 5′-ends of putative antisense RNA transcribed from promoter p_3 were mapped using primer-extension procedures as described in Chapter 3.

5.3.9. Procedures for site-directed mutagenesis. To introduce a mutation in the *C. acetobutylicum glnA* downstream promoter p_3, oligonucleotide directed mutagenesis was performed essentially as described by Stanssens *et al.* (1989). This involved the use of newly developed phasmids based on the gapped duplex DNA (gd DNA) approach (Kramer *et al.*, 1984).

- Construction of gp DNA : Phasmid vector pMc5-8 was digested with *HindIII* and *XbaI* endonucleases and the largest fragment was purified over LMA (Appendix C). ssDNA of pMaWT (0.5 pmole) was mixed with the *XbaI-HindIII* fragment of pMc5-8 (0.1 pmol), 5µL of 1.5 M KCl/ 100 mM Tris-HCl (pH 7.5) was added and the volume adjusted to 40 µL with H_2O. This mixture was heated at 100°C for 4 min and placed at 65°C for 5-10 min. The formation of gp DNA was checked by loading a 20 µL sample on a 1% agarose gel. The rest of the hybridisation mixture can be stored at -20°C until the next step.

- Annealing of the mutagenic primer and DNA polymerase/ ligase reaction: Enzymatically phosphorylated mutagenic oligonucleotide (6 pmole; 45 ng) (Chapter C) was mixed with 8 μL hybridisation mixture in an Eppendorf tube. The mixture was heated to 65°C for 5 min and then allowed to cool to room temperature. After addition of H_2O (18 μL), fill-in buffer (4 μL) (Appendix B), T4-DNA ligase (1μL) (5 U/μL; Boehringer Mannheim, FRG), and DNA polymerase I (Klenow fragment; 1μL) (1 U/ μL; Boehringer Mannheim, FRG), the mixture was placed at 45°C for 45 min and then stored on ice until it was used for transformation.

- Transformation and segregation. A 5 μL sample of the filled-in gp DNA was used to transform the *E. coli* strain WK6mutS. The transformation efficiency was determined by spreading a 100 μL of the transformation mixture (Appendix C) on LB agar plates supplemented with 25 μg/mL Cm. The remainder of the suspension (approximately 1 mL) was used to inoculate 10 mL of LB medium supplemented with 25 μg/mL Cm. After overnight growth, phasmid DNA was isolated using a rapid small-scale purification method (Chapter C) and this mixed phasmid DNA population was allowed to segregate by transforming *E. coli* strain WK6. To minimize the number of double-transformants, a small amount (<20 ng) of the phasmid preparation was used. Aliquots (100 μL) of the transformation mixture were spread on LB agar plates containing 25 μg/mL Cm, and incubated overnight at 37°C. Transformants were tested for sensitivity for Ap (100 μg/mL) to avoid double transformants or parental pMac (Ap^R,Cm^R)

configurations (the result of marker decoupling). Phasmid DNA was prepared from each candidate mutant and further biochemically tested with endonucleases *PvuII* and *ScaI*. Finally, the introduced mutation was confirmed by DNA sequence analysis using primer#2.

5.3.10. Sequencing reactions. DNA sequence analysis was performed as described in Chapter 2 except for the sequencing reactions on pGln1300 DNA used to monitor the transcription initiation sites of antisense RNA transcripts, transcribed from the *C. acetobutylicum glnA* downstream promoter p_3. It was necessary to optimize reaction conditions to obtain sequence data from the antisense RNA transcription initiation sites, which were only 26 nt away from the priming site for primer#2. It was established that the use of a 1/15 (instead of normally 1/8) labeling mix and a twofold increase of Sequenase^{TM} in the sequencing reaction, together with a higher molar ratio of template to primer (2 instead of normally 1), gave the best results.

5.3.11. Tetracycline resistance determinations. The levels of Tc resistance for *E. coli* strains carrying different plasmids were calculated as the Minimum Inhibitory Concentration (MIC). Assay tubes containing 3 mL of LB medium were prepared using varying concentrations of Tc (from 0 to 50 μg/mL, in increments of 2 μg/mL), which was freshly made as a stock solution of 2 mg/mL. Tubes were inoculated with 50 μL of a full-grown cell culture of the strain to be tested and were incubated (with agitation) for

7 h at 37°C. Cell density was measured by absorbance of the samples at 420 nm. MIC values were obtained calculating the minimal Tc concentration needed to inhibit 50% of culture growth compared to the Tc-free culture.

5.4. Results and discussion.

5.4.1. Identification of antisense RNA by primer-extension experiments.

The start sites of antisense transcripts were determined by primer-extension using a synthetic oligonucleotide primer that hybridises to a region 178 to 195 bp downstream of the *C. acetobutylicum glnA* gene (Fig. 2.3), using the same procedures as has been described in Chapter 3.

Primer-extension reactions were carried out on RNA extracted from exponentional phase cells of *E. coli* YMC11 *glnALG* (pGln1300), grown in CSH minimal medium under nitrogen excess (15 mM glutamate/ 15 mM glutamine) and nitrogen-limiting (15 mM glutamate/ 0.15 mM glutamine) conditions (Fig. 5.2). Primer-extension reactions were also performed on RNA extracted from *C. acetobutylicum* cells, grown until the end of the exponentional phase (clostridial stage) in Molasses Fermentation (MF) medium with or without the addition of $(NH_4)_2SO_4$.

Start points t_{Ra} and t_{Rb} were mapped at adjacent nucleotides 4 and 5 bp away from the presumptive -10 region TATAAT of promoter p_3, respectively. This correlates with the third and fourth nucleotides of the region previously proposed to be transcribed into antisense RNA (Fig. 2.3; Janssen *et al.*, 1988). Antisense transcripts did not appear to be regulated by nitrogen availability in either *E. coli* or in *C. acetobutylicum* cells (Fig. 5.2).

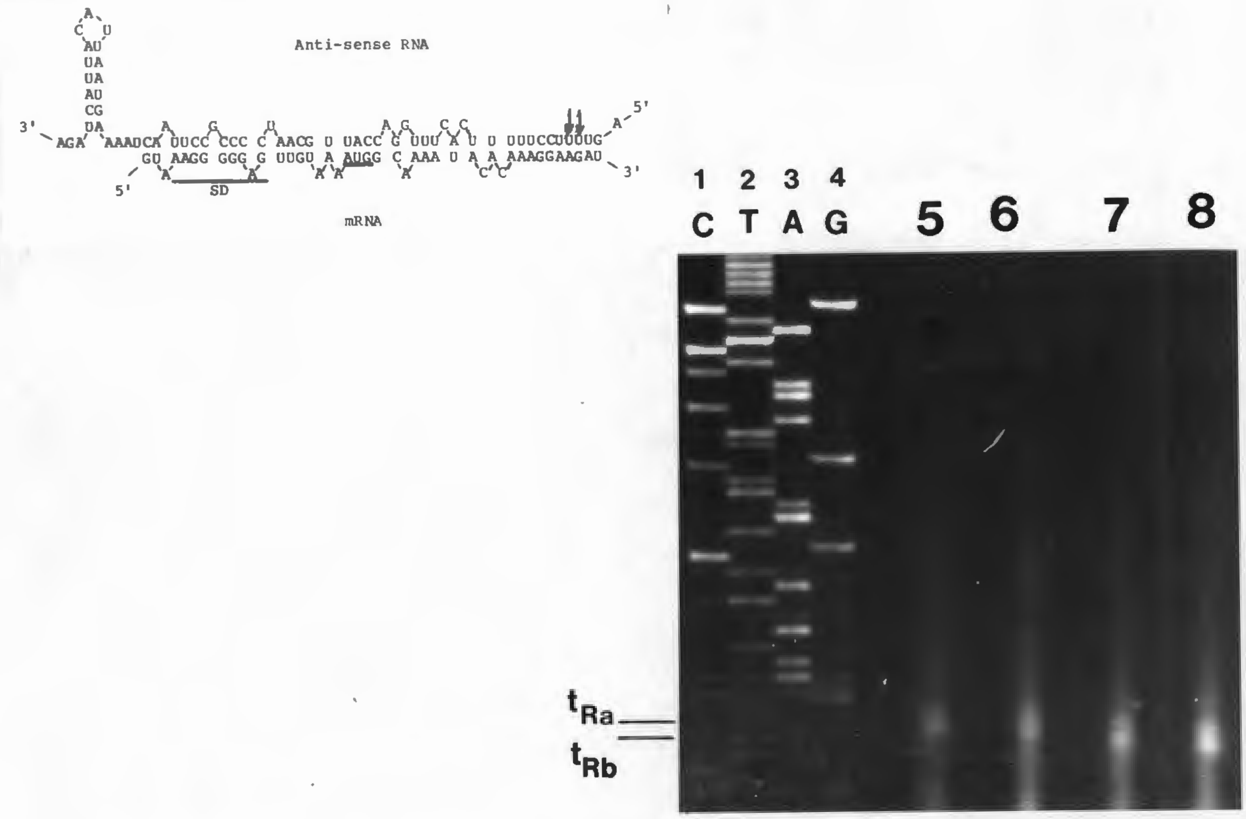


Fig. 5.2.: Identification of antisense RNA start points. Primer-extension analysis was performed using an 18-mer synthetic oligonucleotide primer as described in materials and methods. Lanes 1-4: sequencing ladders, derived from sequence analysis on pGln1300 with the same primer, are shown with the C, T, A, and G reactions from left to right. Lanes 5-6: *E. coli* YMC11 *glnALG* (pGln1300) cells, grown in nitrogen-excess (Lane 5) and nitrogen-limiting (Lane 6) medium. Lane 7-8: *C. acetbutylicum* P262 cells grown in MF medium with (Lane 7) or without (Lane 8) the addition of $(NH_4)_2SO_4$. Two transcripts, t_{Ra} and t_{Rb}, are indicated. The proposed RNA:RNA hybrid formation between divergent RNA transcribed from the antisense promoter p3 and the *glnA* mrRNA ($\Delta G = -41.8$ kcal/mol; Salser *et al.*, 1977) is presented. Hybrid formation would block translation of the *glnA* mRNA because the Shine-Dalgarno (SD) sequence and the initiation codon (AUG) are within the base-paired region.

5.4.2. Site-directed mutagenesis of the *glnA* downstream promoter p_3.

Oligonucleotide directed mutagenesis has become a widely used and invaluable means of probing the structure and function of DNA and the macromolecules it encodes. The usual strategy is to anneal a mutagenic oligonucleotide primer to a single-stranded DNA template, extend the primer by DNA polymerisation to form a complementary strand, ligate the ends and transfect competent bacterial cells with the resulting heteroduplex to obtain progeny containing the desired mutation. The ideal experiment produces a high frequency of mutants in a short time with minimal effort. In actual practice the efficiency of mutagenesis with this multistep procedure depends on the quality of each of the individual steps, each step having limitations that can reduce efficiency.

Table 5.1.: Strand asymmetries for site directed mutagenesis

Strand asymmetry	Selection principle	Reference
methylated/ unmethylated	mismatch repair	Kramer *et al.*, 1982
wild type/ amber codons	nonsense mutations	Kramer *et al.*, 1984
EcoK/EcoB or EcoB/EcoK	restriction modification	Carter *et al.*, 1985
α-S-dNTP/dNTP	restriction endonuclease	Taylor *et al.*, 1985
thymine/uracil	uracil glycosylase	Kunkel *et al.*, 1987

In the last few years numerous approaches for site directed mutagenesis have been developed to increase efficiency by overcoming one or more of these limitations (reviewed by Botstein and Shortle, 1985; Carter, 1986). A common feature of these approaches is that an asymmetry is created between the two strands of the heteroduplex to reduce the frequency of the non-mutant phenotype. The asymmetries that heve been employed are shown in Table 5.1.

A double mutation was introduced into the -35 region of the *C. acetobutylicum glnA* dowstream promoter p_3 which eventually resulted in a new construct, pGln1300/m (Fig. 5.1). This was done by the use of a set of newly developed phasmid vectors (Stanssens *et al.*, 1989) based on the gapped duplex DNA (*gp*DNA) approach of Kramer *et al.* (1984) (Table 5.1; Materials and methods). The mutation in the -35 region of p_3 consisted of the alteration of the sequence TTGCTA into AAGCTA (Fig.5.3). This alteration was choosen for the following reasons. It is well known that *E. coli* promoters, recognized by the RNA polymerase complexed with sigma-70 ($E\sigma^{70}$), contain a higly conserved entry region with consensus sequence TTGACA (reviewed in Chapter 1) of which the thymidine nucleotides at the first two positions of the entry region are extremely well conserved. In fact, to our knowledge, no naturally occuring $E\sigma^{70}$ promoter sequence lacking both thymidine nucleotides at the first two positions of its entry region has been reported, and data on mutant $E\sigma^{70}$ promoters in which the two first nucleotides of the -35 region is changed in two adenine nucleotides, is not available. It appears that the entry region of *E. coli* $E\sigma^{70}$ and $E\sigma^{32}$ promoters, in particular the TTGA sequence, which is a common feature in these promoters, makes contact with the helix-turn-helix motif found in σ factors (Gardella *et al.*, 1989; Siegele *et al.*, 1989). It was therefore reasonable to expect a changed, if not diminished, promoter activity for the mutant p_3 promoter.

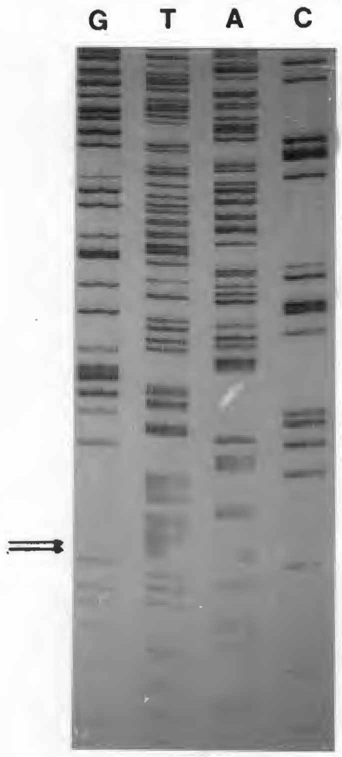


Fig. 5.3.: Site-directed mutagenesis of the *C. acetobutylicum glnA* downstream promoter p_3 was done by using a double mismatch mutagenic oligunucleotide (Materials and Methods). Sequence analysis was performed as described in Fig. 5.1. Sequencing ladders are shown with the **G**, **T**, **A** , and **C** reactions from left or right for both pGln1300 and pGln1300/m.

5.4.3. Comparison of expression levels from mutant and wildtype p_3 promoters.

Exchange of the *tet* promoter region in pBR322 with the wildtype or mutant *glnA* downstream p_3 promoter regions resulted in the recombinant plasmids pWTpal and pMUTpal, respectively (Fig. 5.1). It was noticed that the presence of the extensive stretch of repeat sequences, located immediately downstream of the *glnA* coding region, hampered expression of the *tet* gene, presumably due to the formation of secondary structures, and it was decided to remove these sequences by an internal deletion in order to allow optimal transcription elongation (Table 5.2). This resulted in the construction of plasmids pWT and pMUT, respectively (Fig. 5.1). Wildtype *E. coli* YMC10 cells were transformed with pWTpal, pMUTpal, pWT, and pMUT, and MIC values for tetracycline resistance (Tc^R) were measured during growth in LB medium (Table 5.2).

Table 5.2.: Tetracycline (Tc) resistance of *E. coli* strain YMC10 containing pBR322 and pBR322 derived plasmids carrying various *glnA* downstream promoter fragments.

Plasmid	Construct	MIC (μg/mL)[a]
none	-	2
pBR322	-	22
pWTpal	contains wildtype p_3 region	2
pWT	removal of direct repeats	40
pMUTpal	contains mutant p_3 region	2
pMUT	removal of direct repeats	> 50

(a) MIC values were obtained calculating the minimal Tc concentration needed to inhibit 50% of culture growth compared to the Tc-free culture (Material and methods).

It is interesting that the *cat* gene under control of the wildtype p_3 promoter in plasmid pCAGSp24 (Materials and Methods) was expressed in *B. subtilis* strain 1A46 (Janssen *et al.*, 1988), despite the presence of the inverted repeat sequences located immediately behind the p_3 promoter. It is possible that, in this particular case, the *B. subtilis* transcription elongation complex is less sensitive to the putative secondary structures in the nascent transcript than the transcription elongation complex of *E. coli*.

Because Tc resistance levels for *E. coli* strains used in the above experiment were measured in LB medium containing 0 to 50 µ/mL tetracycline in increments of 2 µg/mL, it was not possible to give an exact MIC value for *E. coli* YMC10 (pMUT) cells. However, no inhibition of growth was observed at a tetracycline concentration of 50 µg/mL. The MIC value for Tc resistance of *E. coli* YMC10 (pMUT) cells must thus be at least 50 µg/mL, which is much higher then for *E. coli* YMC10 (pWT) cells which had an MIC value of 40 µg/mL, i.e. showed a 50% inhibition of growth at a tetracycline concentration of 40 µg/mL. It is therefor concluded that expression levels from the mutant p_3 promoter were considerably higher than expression levels from the wildtype p_3 promoter. This was unexpected for reasons explained above. It is possible that the mutant promoter with the new entry region AAGCTA is stronger *in vivo* then the wildtype promoter with an entry region TTGCTA, but in our opinion, this is unlikely. Alternatevily, a new promoter sequence may have been created (Fig. 5.4), in which the newly formed -10 region TTTAAG is preceded by a suitable TTGCGT -35 region, located 17 bp away. It could be argued that additional mutations outside the target site have been introduced due to the fact that RNA polymerase has a high error frequency (Fersht and Knill-Jones, 1983). For this reason, the entire segment, exposed as single stranded DNA during the experiment, was sequenced after mutagenesis (Fig. 5.3; complete data not shown). The only alterations involved the substitution of the two TT nt by the AA nt.

Newly created promoter sequence ?

```
   -35                       -10
  TTGCGT.................TTTAAG
TCTTGCGTTTACGAAGTTTTTAATTTTTAAGCTATCTCAACATTGCTTAATGCTATAATA

TCTTGCGTTTACGAAGTTTTTAATTTTTTTGCTATCTCAACATTGCTTAATGCTATAATA
283                         TTGCTA................TG.TATAAT 223
                             -35                      -10
```

WT p_3 promoter

Fig. 5.4.: Mutation of the -35 region of the *glnA* downstream p_3 promoter resulted in the creation of a new promoter consensus sequence. A 60 bp region between 223 and 283 nt from the *glnA* stop codon is represented. The double mutation is indicated in bold letter type.

Ponnambalam *et al.* (1986) have constructed promoter mutations in the *E. coli gal* operon which resulted in promoter sequences lacking a convential -35 region. These deletion mutant promoters were shown to be functional *in vivo*. Apparently, in these cases studied, the Pribnow box (-35 region) was sufficient to provide initiation of transcription. However, a TG pair, situated 1 bp from the -10 region, was shown to be important for promoter activity in such mutant constructs. From their work, and from similar studies on the P_{re} Promoter of bacteriophage lambda (Keilty and Rosenberg, 1987), it was concluded that the TG motif 1 bp upstream of the Pribnow box not only was essential to promoter activity (for these promoters), but also was sufficient to compensate the lack of a conventional -35 region. It is suggested that this extraordinary process of promoter recognition and transcription elongation (open complex formation) occurring at the same region in the DNA, requires additional primary information within DNA sequences of the -10 region, for instance to allow a conformational change in the $E\sigma^{70}$ complex. It is interesting that the mutant *galP2* promoter and the P_{re} promoter both contained a perfect (TATAAT) or near-perfect (TATACT) Pribnow box, respectively.

The *C. acetobutylicum glnA* downstream promoter p_3 contains a TG nt pair 1 bp away from a perfect consensus TATAAT Pribnow box. It is therefore possible that the -35 region TTGCTA of promoter p_3 is not absolutely required for promoter activity, or in other words, that the mutation TT » AA within this region does not result in diminished promoter activity for the mutant p_3 promoter. In the contrary, the mutant p_3 promoter has even higher promoter activity *in vivo* than the wildtype p_3 promoter.

It was regulary observed that late stationary phase cells of *E. coli* strains YMC10(pWT) and YMC10(pMUT) grown overnight on LB medium failed to grow on fresh agar plates supplemented with 15 µg/mL tetracycline. To monitor expression from the *glnA* downstream wildtype p_3 and mutant p_3 promoters during growth of *E. coli* YMC10(pWT) and YMC10(pMUT) cells, respectively, the following experiment was designed. *E. coli* YMC10(pWT) and YMC10(pMUT) cells were grown in Tc-free LB medium and samples were taken during growth. From each sample, a 10^{-5} dilution was made and 20 µL of diluted cells were plated on LB agar plates containing either 100 µg/mL Ap or 100 µg/mL Ap + 15 µg/mL Tc. The numbers of Ap^r and Ap^r Tc^r colonies were scored after growth at 37°C for 14 h (Table 5.3).

Table 5.3.: Expression of wildtype and mutant p_3 promoters during growth of *E. coli* strains YMC10(pWT) and YMC10(pMUT), respectively.

Wildtype p_3 promoter			
OD_{600}	Ap^R	Ap^R Tc^R	R^a
0.120	24	7	0.29
0.340	50	8	0.16
0.760	120	18	0.15
1.800	270	130	0.48
2.100	325	150	0.46
2.000	348	160	0.46
2.000	308	66	0.21
2.000	460	140	0.30

Mutant p_3 promoter			
OD_{600}	Ap^R	Ap^R Tc^R	R
0.118	33	8	0.24
0.336	50	18	0.36
0.840	166	38	0.23
1.800	340	220	0.65
2.100	360	209	0.58
2.000	295	189	0.64
2.000	372	124	0.33
2.000	460	90	0.20

(a) The ratio of Ap^R/Tc^R transformants to Ap^R transformants is represented as R.
Samples were taken every 45 min.

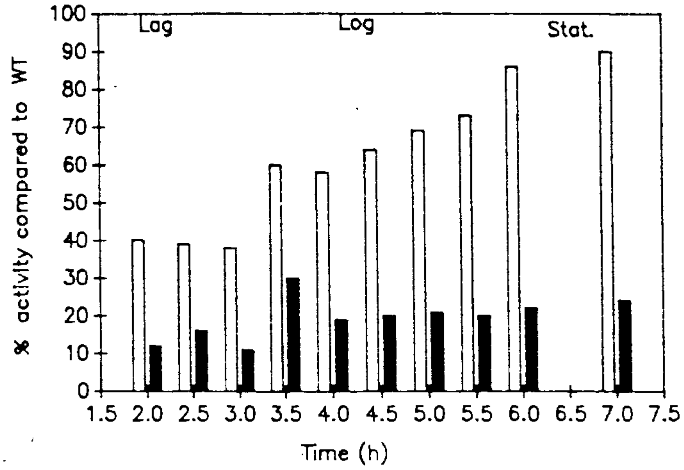
Expression levels of the wildtype and mutant p_3 promoters appear to be very similar although a slightly higher ratio of Ap^RTc^R to Ap^R transformants was observed for the *E. coli* strain YMC10(pMUT), especially during early stationary phase (30-40 % increase) (Table 5.3). In addition, the ratios of Ap^RTc^R to Ap^R transformants for both *E.coli* strains YMC10(pWT) and pYMC10(pMUT) fluctuate and are maximal during early stationary phase. It is known that transcription from a strong promoter towards the origin of replication produces a negative effect on plasmid stability (Stueber and Bujard, 1982). In the case of pBR322, transcription from either the wildtype or the mutant p_3 promoter may enhance the synthesis of RNA I, the specific inhibitor of replication initiation for plasmids of the ColE1 type, whose direction of transcription coincides with that of the *tet* gene. However, it is not anticipated that this would affect the interpretation of the results presented in Table 5.3 because the numbers of Ap^RTc^R and Ap^R transformants were compared (reflected in the R-value) for each particular sample. It may be also argued that the *tet* gene product, which is localized in the cytoplasmic membrane of *E. coli* cells (Chopra, 1986), may affect membrane functioning, thereby decelerating the growth rate of plasmid-containing *E. coli* cells (Lee and Edlin, 1985). Because the *E. coli* strains YMC10(pWT) and YMC10(pMUT) were pre-grown in Tc-free LB medium, plasmid-free cells may arise and their proportion would gradually increase because the difference in the growth-rates between plasmid-containing and plasmid-free cells. It is not apparent that this was the case in the above experiment, because the number of Ap^R transformants correlated well with the optical density (OD_{600}), measured on samples from the Tc-free grown cultures (Table 5.3).

It is therefore proposed that changes in Tc resistance of the various *E. coli* strains used in this study are a direct result of the variable expression of the *tet* gene and that the function of the wildtype and mutant p_3 promoters varied with the growth phase in *E. coli* cells.

5.4.4. Effect of the mutant p_3 promoter on the expression of the *C. acetobutylicum glnA* gene in *E. coli*. The mutant p_3 promoter was introduced in pGln1300 by exchanging the *XbaI-HindIII* fragment of pGln1300 by the *XbaI-HindIII* fragment of pMcMut (Materials and methods). This resulted in the plasmid pGln1300\m (Fig. 5.1). DNA sequence analysis was performed on pGln1300\m to confirm the introduction of the TT to AA mutation in the entry region TTGCTA of promoter p_3 (Fig. 5.3). No other mutations were detected within the entire *XbaI-HindIII* fragment of plasmid pGln1300\m. *E.coli* YMC11 *glnALG* cells containing either pGln1300, pGln1300\m, pGlnWT, or pGlnMUT (Fig. 5.1), were grown in CSH minimal medium under nitrogen-limiting (4 mM glutamine) or nitrogen-excess (4 mM glutamine, 8 mM ammonium sulphate) conditions and GS activity was measured during growth.

GS levels in *E. coli* YMC11(pGln1300\m) cells were generally lower than in *E. coli* YMC11(pGln1300) cells, irrespective of the nitrogen conditions (Fig. 5.5). This is in agreement with the hypothetical inhibitory effect of antisense RNA on GS activity. It was previously shown that expression of the *tet* gene in *E. coli* YMC10(pMUT) cells was considerably higher (MIC of >50 μg/mL Tc) than in *E. coli* YMC10(pWT) cells (MIC of 40 μg/mL Tc) (Table 5.2). It is

High nitrogen (Ngln medium)



Low nitrogen (gln medium)

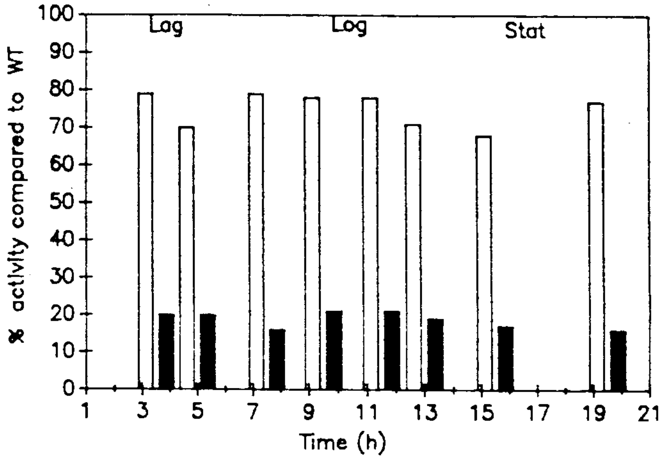


Fig. 5.5: Comparison of GS levels for *E. coli* YMC11 cells containing pGln1300, pGln1300\m (open bars), and pGlnWT (solid bars). Cells were grown under nitrogen-excess (Ngln; 4 mM glutamine, 8 mM ammonium sulphate) and nitrogen-limiting (gln; 4 mM glutamine) conditions. GS activities were measured throughout all growth stages (Lag, Log and Stationary phase) and are expressed as percentages of the GS activity of *E. coli* YMC11(pGln1300) cells.

therefore likely that levels of antisense RNA are accordingly higher in YMC11(pGln1300/m) cells than in *E. coli* YMC11(pGln1300) cells. Consequently, this would result in a more profound effect of antisense RNA on GS activity in *E. coli* YMC11(pGln1300\m) cells.

In order to obtain a better understanding of the role of antisense RNA production in the regulation of the *C. acetobutylicum glnA* gene in *E. coli*, experiments involving saturated mutagenesis of the p_3 promoter proximal region should be designed. Such mutations should be brought back into the original recombinant plasmid pHZ200 (Chapter 2) as this plasmid contains, in addition to approximately 1.3 kb *glnA* upstream region, a 3.5 kb *glnA* downstream region, which has been shown to be essential for overall expression of the *glnA* gene (Table 4.1; Janssen *et al.*, 1988). This was not done for the TT to AA mutation in the p_3 promoter used in the above experiments due to the lack of suitable restriction sites in pHZ200 (which contains multiple *XbaI* and *HindIII* sites).

The exchange of the *XbaI-HindIII* fragment of pGln1300 by the *XbaI-HindIII* fragments of either pWT or pMUT resulted in the construction of plasmids pGlnWT and pGlnMUT, respectively (Fig. 5.1; Materials and methods). These plasmids lack part of the region transcribed into antisense RNA, together with the inverted repeat sequences located between the p_3 promoter and the 3′-end of the *glnA* gene but the p_3 promoter was not deleted. *E. coli* strains

YMC11 (pGlnWT) and YMC11 (pGlnMUT) were grown under nitrogen-excess and nitrogen-limiting conditions and the GS activity was measured during growth (Fig. 5.5; data obtained for the *E. coli* YMC11 (pGlnMut) strain are identical to those obtained for the *E. coli* YMC11 (pGlnWT) strain and are omitted for clarity sake). A five-fold decrease in GS activity was observed when compared to GS activities for *E. coli* YMC11 (pGln1300). Similar results were obtained with *E. coli* YMC11 cells containing plasmid pGlnDD480 (Fig 4.1, Table 4.1), which lacked the p3 promoter and part of the inverted repeat sequences. A two-fold decrease in GS activity was observed when compared to GS levels in *E. coli* YMC11 (pGln1300) cells (Table 4.1; Janssen *et al.*, 1988). It is therefor concluded that the inverted repeat sequences, which have the potential to form secondary structures in the transcribed *glnA* mRNA, are essential for optimal *glnA* gene expression.

Chapter 6

General Discussion.

Glutamine synthetase (GS), a key enzyme in bacterial nitrogen metabolism, is a well characterised enzyme and extensively studied by numerous research groups, and its regulation and structure is widely documented. Two model systems emerge from the many reviews and publications: the *E. coli glnALG* and the *B. subtilis glnRA* operons.

The *E. coli glnALG* operon consists of the GS structural gene, *glnA*, and two regulatory genes, *glnL* (*ntrB*) and *glnG* (*ntrC*), whose products are part of a large family of signal-transducing proteins and of which phosphorylated NtrC functions as an activator of *glnA* transcription. Activation of *glnA* gene expression in *E. coli* requires also a specific sigma factor (σ^{54}) which is the product of the unlinked *ntrA* (*glnF*; *rpoN*) gene. Other unlinked genes, which are scattered throughout the chromosome, regulate GS activity in *E. coli* on post-translational level via a cascade of regulatory processes, resulting in the adenylylation/ deadenylylation of GS subunits. The interplay between these diverse regulatory mechanisms allows *E. coli* cells to adapt readily and ensures a swift and adequate response to nitrogen deprivation.

In *B. subtilis*, however, such an intricate control of GS activity is not present. There is no evidence for an adenylylation system and *ntr*-genes have not been detected. Instead, GS activity is controlled mainly by the regulation of *glnA* transcription, and involves a novel regulatory protein, the *glnR* gene product. GS appears to modify this regulatory protein thereby enabling GlnR to bind to the *glnRA* promoter proximal region and repress transcription. Compared with the *E. coli* system, the *glnRA* operon of *B. subtilis* appears to be regulated by less complex mechanisms in which only one regulatory protein is involved. This may be due to the fact that *B. subtilis* cells, when subjected to drastic nutritional changes in the environment, are able to sporulate. Another spore-forming organism, *B. cereus*, has been shown to contain an analogous *glnRA* operon, with a *glnA* gene that is very similar (87.3% amino acid homology) to the *B. subtilis glnA* gene and which is preceded by a corresponding *glnR* gene (Nakano *et al.*, 1989).

It is of considerable interest to study the genes and enzymes of nitrogen metabolism in *C. acetobutylicum* because nitrogen levels in *C. acetobutylicum* are important for solventogenesis and sporulation (Long *et al.*, 1984). An interesting finding was that *C. acetobutylicum* cells only sporulate in excess of nitrogen. Sporulation appears to be triggered by factors such as oxygen exposure, butanol toxicity, and differences between intra- and extracellular pH. However, sporulation did not occur as a result of any of these stimuli when nitrogen levels where insufficient.

The *C. acetobutylicum glnA* gene was cloned previously in *E. coli* (Usdin *et al.*, 1986) and it was shown that the cloned *glnA* gene was expressed and regulated by levels of nitrogen from its own regulatory region (Usdin *et al.*, 1986). In addition, there was no evidence for a global Ntr system, and the GS enzyme was not adenylylated.

The aim of this study was to characterize the *C. actobutylicum glnA* gene and to gain fundamental knowledge about the regulatory mechanisms involved. The DNA sequence of the *C. acetobutylicum glnA* coding region and of approximately 2 kb of the flanking regions was consequently analysed. The primary structure of GS was compared with that of GS enzymes of 10 other organism and a phylogenetic study was performed to estimate the evolutionary relationship between the various enzymes. The *C. acetobutylicum* GS enzyme resembles the GS enzymes of *B. subtilis* and *B. cereus* in several ways. It is higly similar in amino acid sequence (60%), lacks the central loop region, and is not regulated by adenylylation nor by any ntr-genes.

In spite of these resemblances, no *glnR* gene could be detected upstream of the *C. acetobutylicum glnA* gene. Instead, several interesting features within this region, including a complex palindromic sequence, a set of direct repeats 5'-ATATTGTAA-3', and long tracks of dA_n and dT_n nt, were observed. Furthermore, two promoters p_1 and p_2 in front of the *glnA* gene were identified by transcipt mapping in both *E. coli* and *C. acetobutylicum* cells. Regulation of the *glnA* gene on the level of transcription was mainly exerted via the repression/derepression of p_1 promoter activity whereas the p_2 promoter seemed to play a minor role.

As has been elucidated by *glnA* up- and downstream deletion mutants and various *glnA::lacZ* fusions, sequences located far downstream of the *C. acetobutylicum glnA* gene play an important role in GS production. It is tempting to speculate that regulatory genes are involved whose products interact with the glnA promoter proximal region. For this reason it is essential to analyse the DNA sequence of the entire 3.5 kb *C. acetobutylicum glnA* downstream region

The most striking result of this study is the observed discrepancy between transcription levels and amount of GS produced. It is proposed that these differences are due to regulatory mechanisms operating at the level of translation. Although antisense RNA transcribed from the glnA downstream promoter p_3 has been detected in both *E. coli* and *C. acetobutylicum* cells, it is not entirely clear what the precise effect is of antisense RNA on glnA mRNA translation. Site-directed mutagenesis was used in an attempt to abolish p_3 promoter activity. However, the opposite occurred as it was shown that expression levels from the mutant p_3 promoter were much higher then those of the wildtype p_3 promoter. Nevertheless, some interesting results were obtained. GS production levels were shown to be affected by the mutation and were approximately 25% lower compared to GS levels produced by the wildtype. This inhibition was observed during all growth phases and irrespective to nitrogen levels present in the medium. Because no other mutation was introduced in the glnA downstream region then the TT to AA change in the -35 region of the p_3 downstream promoter, it is concluded that antisense RNA indeed can have an inhibitory affect on the production of *C. acetobutylicum* GS in *E. coli*. To study the effects of this antisense RNA on glnA gene expression, it will be necessary to isolate a p_3 promoter down mutant. Work in this direction is already in progress using recently obtained *galK* promoter probe vectors.

In order to understand this complex regulatory system, in which regulation of gene expression is exerted at transcriptional as well as at translational level, I suggest that the following experiments should be undertaken (in roughly descending order of importance):

1) Sequence analysis of the 3.5 kb *glnA* downstream region in order to identify putative regulatory genes.

2) The isolation and identification of p_3 promoter down mutants and the use of such mutants in regulatory studies. Without doubt, this would pave the way for further investigations on the regulation of the *C. acetobutylicum* *glnA* gene.

3) The use of p_3 promoter down mutants fused to *galK* to monitor expression during growth.

4) Transcript mapping of the mutant (TT to AA) p_3 promoter to elucidate the true nature of the mutation. From an academical point of view, it would be interesting to see wether a new promoter was constructed or wether the -35 region of this promoter is non-essential for promoter activity.

5) Saturated mutagenesis of the *glnA* upstream region, preferentially on plasmid pHZ200, to investigate the importance of the CPS. Mutations in the *glnA* promoter proximal region would also be usefull in regulatory studies.

6) The isolation and study of *C. acetobutylicum* GlnA$^-$ mutants. More physiological work has to be done on wildtype and *glnA* mutant strains of *C. acetobutylicum* and as soon techniques come available to transfer DNA back to *C. acetobutylicum* cells, regulatory studies on specific mutants could be performed.

Appendix: contents

Appendix A: Strains and plasmids

A
Appendix A

Bacterial strains and plasmids.

A1. Bacterial strains and plasmids used in this study.

Strain or plasmid	Characteristics	Reference/ source
Bacterial strains:		
C. acetobutylicum P262	wildtype	Jones *et al.* (1982)
E. coli K514	*thr-1 leu-6 thi-1 supE44 lacY1 tonA21* r^- r^+	Zabeau and Stanley (1982)
E. coli LK111	*lacI*^{q} *lacZ*ΔM15 (LacY^{+} derivative of K514	Zabeau and Stanley (1982)
E. coli JM105	Δ(*lac-proAB*) *thi strA sbc15 hsdR4* /F′ *traD36 proAB lacI*^{q} *lacZ*ΔM15	Yannisch-Perron *et al.* (1985)
E. coli JM107	Δ(*lac-proAB*) *endA1 gyrA thi-1 hsdR17 supE44 relA1* /F′ *traD36 proAB lac I*^{q} *lacZ*ΔM15	Yannisch-Perron *et al.* (1982)
E. coli YMC10	*thi endA hsr hutC*$_{K}$ Δ*lac U169*	Backman *et al.* (1981)
E. coli YMC11	Δ(*glnG-glnA*) 2000 derivative of YMC10	Backman *et al.* (1981)
E. coli ET8051	*rbs lacZ*::IS1 *gyrA* Δ(*glnA-rha*)VIII	Pahel and Tyler (1979)
E. coli WK6	Δ(*lac-proAB*) *galE strA*/ *F′ lacI*^{q} *lacZ*ΔM15 *proAB*	Zell and Fritz (1987)
E. coli WK6mutS	*mutS*::*Tn*10 derivative of WK6	Zell and Fritz (1987)
E. coli QD5003	Hfr *supF*	E. Remaut, RUG Ghent, Belgium
B. subtilis 1A46	*pro-1 trpC2 thr-5 recE4*	Bacillus Genetic Stock Center, USA
Plasmids$^{(*)}$**:**		
E. coli		
pLK34	Ap^{R}	Botterman and Zabeau (1987)

pBR322	Ap^R Tc^R	Bolivar *et al.* (1977)
pUC8, pUC9	Ap^R	Vieira and Messing (1982)
pUC13	Ap^R	Messing (1983)
pUC18, pUC19	Ap^R	Norrander *et al.* (1983)
pMC1871	Tc^R *lac*	Casadaban *et al.* (1980)
pMa5-8	Ap^R Cm^S	Stanssens *et al.* (1989)
pMc5-8	Ap^S Cm^R	Stanssens *et al.* (1989)
pAcYc177	Ap^R Kn^R	Chang and Cohen (1978)
M13mp18, M13mp19		Yanisch-Perron *et al.* (1985)
pHZ200,	Ap^R $GlnA^+$	Usdin *et al.* (1986)
pGln1300 (=pJP3)	pLK34[a] pHZ200[b] / *BglII-XbaI*	Janssen *et al.* (1988)
pLac1300	pMC1871[a] pGln1300[b] / *SmaI-EcoV*	Janssen *et al.* (1989)
pMaWT	pMa5-8[a] pGln1300[b] / *HindIII-XbaI*	This study
pMcMut	pMaWT (one mutagensis round)	This study
pGln1300\m	pGln1300[a] pMcMut[b] / *HindIII-XbaI*	This study
B. subtilis		
pPL603b1	Kn^R	Corfield *et al.* (1984)
pPL608	Kn^R	Mongkolsuk *et al.* (1983)

(*) Only basic cloning vectors, or plasmids used as the primary source of DNA in cloning procedures, have been presented.

(a) vector fragment (derivative).

(b) insert fragment / restriction endonucleases used.

A.2 Maintenance of bacterial strains.

A.2.1 *C. acetobutylicum* P262. *C. acetobutylicum* P262, which was used to extract RNA for primer-extension reactions, was maintained as a spore stock in distilled water at 4°C. Spore stocks were prepared as described by Long *et al.* (1983).

A.2.2 *B. subtilis* stock cultures. *B. subtilis* strain 1A46 was maintained as a spore stock in distilled water at 4°C and was prepared using Schaeffer sporulation agar (Appendix B).

A.2.3 *E. coli* stock cultures. Stock cultures of *E. coli* strains were maintained at -70°C. Cultures (Typically 5 mL LB medium, supplemented with the required antibiotic(s); Appendix B) were inoculated from overnight (16 h) cultures (dilution factor 200 to 1000) or from isolated colonies on LB plates, and incubated with aeration at 37°C to an cell density of approximately 6×10^8. Sterile glycerol was added (15%, v/v final concentration), and aliquots (100 μL) were stored at -70°C in fourfold. In parallel, soft agar (0.7% w/v) LB medium was used to store *E. coli* strains on room temperature.

E. coli strains JM105, JM107, WK6, and WK6mutS, were regularly streaked on M9 minimal medium agar plates (Miller, 1972) to select for the *pro* marker on the F′ episome.

A.2.4. Competent *E. coli* cells. *E. coli* competent cells, prepared as described in Appendix C, were stored at -70°C in 0.1 M $CaCl_2$ containing 15% (v/v) glycerol.

Appendix B

Media, buffers, and solutions.

B.3 Buffers and solutions

All media, buffers and, solutions were sterilized by autoclaving at 121 °C for 17 min unless indicated otherwise. Heat labile substances were sterilized by filtration through 0.22 μm nitrocellulose membrane filters (Millipore).

B.1 Media

B.1.1 Clostridial Basal Medium (CBM) (O'Brien and Morris, 1971).

Glucose	10 g
Casein hydrolysate	4 g
Yeast extract	4 g
Distilled water	972 mL

The following constituents were added from stock solutions as indicated.

$MgSO_4.7H_2O$	(20%, w/v)	1 mL
$MnSO_4.4H_2O$	(1%, w/v)	1 mL
$FeSO_4.7H_2O$	(1%, w/v)	1 mL
p-Aminobenzoic acid	(0.1%, w/v)	1 mL
Biotin	(0.0002%, w/v)	1 mL
Thiamine HCl	(0.1%. w/v)	1 nL
$NaHCO_3$	(10%, w/v)	10 mL
Cysteine HCl	(5%, w/v)	10 mL
Resazurin	(0.25 mg/mL)	2 mL

The stock solutions were filter sterilized and were stored at 4°C, except FeSO4, which was stored at -20°C.

Liquid CBM medium was autoclaved and allowed to stand for 15 h under stringent anaerobic conditions in an anaerobic glove box (Forma Scientific Inc, Marietta, Ohio; atmosphere 70% N_2 : 20% CO_2 : 10% H_2 (v/v/v)) before inoculation. Cysteine HCl was added to all liquid anaerobic cultures, in order to scavenge any residual O_2. Agar (1.5%, w/v) was added to CBM broth made as above without resazurin and with the exception that the

Cysteine HCl and $NaHCO_3$ were added after the medium had been autoclaved and prior to pouring.

Hungate culture tubes for small volume (10 mL) overnight cultures of *C. acetobutylicum*, were prepared by boiling the medium for 15 min to drive off most of the O_2, dispensing into glass Hungate tubes and perfused with H_2 and CO_2, before sterilization by autoclaving. Spore stocks of *C. acetobutylicum* were heat shocked at 70°C for 2 min and then plaved on ice for 1 min, before inoculation.

B.1.2 Molasses Fermentation Medium (MFM) (Barber *et al.*, 1979)

Molasses (6.5% total invert sugar required)	134.0 g
$(NH_4)_2SO_4$	2.0 g
$CaCO_3$	1.0 g
Steepwater concentrate	1.5 mL
Distilled water	to 1000 mL

The pH of the medium was adjusted to pH 7.0 with 1 N NaOH before autoclaving.

B.1.3 Brain Heart Infusion (BHI) broth (*B. subtilis*)

BHI (Merck)	18.5 g
Yeast extract	2.5 g
Distilled water	to 500 mL

B.1.4 Ttryptose Blood Agar Base (TBAB) (*B. subtilis*)

This medium was made according to the recommendations of the supplier (Oxoid, code CM233) (ready-to-use agar base).

B.1.5 DM3 protoplast regeneration agar base (*B. subtilis*) (Chang and Cohen, 1979)

The medium is made from the following sterile solutions. The 4% agar can be mixed with the Na-succinate solution when both are are hot to dilute the agar. This mixture and all the other solutions are brought to 50°C and then combined. All the solutions are autoclaved except BSA which is filtered through a nitrocellulose filter (0.22 μm).

(per L)		
Yeast extract	(10% w/v)	200 mL
Casamino acids	(5% w/v)	100 mL
Na-succinate	(1 M, pH 7.3)	500 mL
Agar	(4%, w/v solution)	200 mL
Phosphate buffer	(see below)	100 mL
$MgCl_2$	(1 M)	20 mL
Glucose	(20%, w/v)	25 mL
BSA fraction V	(2%, w/v)	5 mL

Phosphate buffer: Combine K2HPO4 (3.5%, w/v) and KH2PO4 (1.5%, w/v), and autoclave.

Antibiotics: for regeneration of *B. subtilis* protoplasts, less antibiotic is preferred for direct selection (5 μg/ mL Cm; 30 μg/ mL Kn).

B.1.6 PenAssay Broth (PAB) (*B. subtilis*) (4x)

Beef extract	6 g
Yeast extract	6 g
Peptone	20 g
Glucose	4 g
NaCl	14 g
K_2HPO_4	14.72 g
KH_2PO_4	5.28 g
Distilled water	to 1000 mL

B.1.7 Schaeffer's sporulation agar base (*B. subtilis*)

Nutrient broth	8 g
$MgSO_4.7H_2O$	0.25 g
KCl	1.0 g
Agar	15 g
Distilled water	to 1000 mL

This mixture is autoclaved, cooled to 50°C, and supplemented with sterile (separately autoclaved) solutions of:

$CaCl_2$	(0.1 M)	5 mL
$MnSO_4.4H_2O$	(5 mM)	1 mL

B.1.8 Minimal A medium (Cold Spring Harbor; CSH) (*E. coli*)
(Miller *et al.* 1972)

<u>Solution 1</u>:

Minimal agar (Oxoid No. 1)	15 g
Distilled water	600 mL

<u>Solution 2</u>:

Minimal salts solution (5x)	200 mL
Distilled water	200 mL

Minimal A salts solution (5x)

K_2HPO_4	52.5 g
KH_2PO_4	22.5 g
$C_6H_5Na_3O_7 \cdot 2H_2O)$ (Sodium citrate)	2.5 g
$(NH_4)_2SO_4$ (added when required)	5.0 g
Distilled water	to 1000 mL

Solutions 1 and 2 were autoclaved separately, cooled to 50°C, and then combined. The remainder of the constituents were added from stock solutions as indicated below.

Carbon source	(20%, w/v)	10 mL
Amino acids	(0.2%, w/v)[a]	1 mL
$MgSO_4$	(20%, w/v)	1 mL
Thiamine HCl	(0.5%, w/v)	1 mL

(a) For D,L-aminoacids, 0.4% (w/v) was used

B.1.9 Minimal M9 medium (*E. coli*) (Miller *et al.*, 1972)

<u>10 x M9 salts:</u>

Na_2HPO_4	60 g
KH_2PO_4	30 g
NH_4Cl	10 g
NaCL	5 g
Distilled water	to 1000 mL

M9 minimal salts were autoclaved and stored at 4°C

The following reagents were autoclaved separately and cooled before mixing:

a) 15 g minimal agar in 900 mL H2O
b) 100 mL 10x M9 salts
c) 1 mL $MgSO_4$ (1 M)
d) 1 mL $CaCl_2)$ (0.1 M)
e) 1 mL thiamine-HCl (1 M)
f) 10 mL glucose (20%, w/v)

B.1.10 YT medium (2x) (*E. coli*)

Bacto tryptone	16 g
Yeast extract	10 g
NaCl	5 g
Distilled water	to 1000 mL

For selection of pUC plasmids and their derivatives, IPTG (0.1 mL) and X-gal (0.8 mL) were added to 250 mL agar (50°C) before pouring the plates.

B.1.11 Luria-Bertani medium (*E. coli*). to 1000 mL

The pH of the medium was adjusted to pH 7.0 with 0.1 N NaOH before autoclaving. Solid media contained 1.5% (w/v) agar.

B.1.12 McConkey agar base (*E. coli*).

This medium was made according to the specifications of the supplier (Difco). Lactose was added to a final concentration of 1% after sterilization.

B.1.13 Eosin Methylene Blue agar base (EMB) (*E. coli*)

This medium was made according to the specifications of the supplier (Oxoid, code CM69).

B.1.14 H plates

Bacto tryptone	10 g
NaCl	8 g
Agar	12 g
Distilled water	to 1000 mL

B.1.15 H top agar

Bacto tryptone	10 g
NaCl	8 g
Agar	8 g
Distilled water	to 1000 mL

B.2 Media additives.

Media were cooled to 50°C before addition of antibiotics, X-gal or IPTG. Plates containing these additives were usually stored for no longer than one week at 4°C.

B.2.1 Antibiotics.

Antibiotic	quantity	volume	stock
Ampicillin	1.0 g	10 mL	100 μg/ mL
Chloramphinecol	0.1 g	10 mL	10 μg/ mL
Tetracycline	0.15 g	1 mL	15 μg/ mL
Kanamycine	0.25 g	10 mL	25 μg/ mL

Tetracycline (Tc) stock solution was freshly prepared and Tc-containing plates were stored in the dark. All antibiotics were dissolved in distilled water, except for chloramphinecol, which was dissolved in 100% ethanol.

B.2.2 **IPTG** (isopropyl-β-D-thio-galactopyranoside)

IPTG (100 mM)	23.8 mg
Distilled water	1.0 mL

The solution was stored in aliquots at -70°C.

B.2.3 **X-gal** (5-bromo-4-chloro-3-indolyl-β-galactoside)

X-gal (2%, w/v)	0.2 g
Dimethyl formamide	10 mL

The solution was stored in aliquots at -70°C.

B.3 Buffers and solutions

B.3.1 Universal restriction enzyme buffer

(M. Zabeau, Plant Genetic Systems, Ghent, Belgium)

Stock solution		final conc.
Tris-HCl pH7.9	(1 M)	0.1 M
$MgCl_2$	(1 M)	0.1 M
DTT	(0.5 M)	10 mM
BSA	(10 mg/ mL)	1 mg/ mL

Glycerol		44% (w/v)
NaCl	(5 M)	0, 50, 100, or 150 mM

The buffers (10 mL) were made using the following table and stored at -20°C.

Stock solution	Salt concentration (mM)			
	0	50	100	150
1 M Tris-HCl (pH 7.9)	1 mL	1 mL	1 mL	1 mL
$MgCl_2$	1 mL	1 mL	1 mL	1 mL
5 M NaCl	-	1 mL	2 mL	87.7 mg
DTT	0.2 mL	0.2 mL	0.2 mL	0.2 mL
BSA	1 mL	1 mL	1 mL	1 mL
Glycerol	4.4 mL	4.4 mL	4.4 mL	4.4 mL
Sterile water	2.4 mL	1.4 mL	0.4 mL	2.4 mL

B.3.2 Universal restriction enzyme dilution buffer

(M. Zabeau, Plant Genetic Systems, Ghent, Belgium)

The buffer was made according to the following table and stored at -20°C.

Stock solution	per 10 mL	Final conc
1 M Tris-HCl pH 7.5	0.1 mL	10 mM
5 M NaCl	0.1 mL	50 mM
14 M β-mercaptoethanol	7 μL	10 mM
Gelatin (10 mg/ mL)	0.1 mL	100 μg/ mL
Glycerol	4.4	44% (v/v)
Distilled water (sterile)	5.3 mL	-

B.3.3 *SmaI* restriction endonuclease buffer.

The buffer was made according to the following table and stored at 4°C.

Stock solution	per 10 mL	Final conc.
1 M Tris-HCl pH 8.0	1 mL	100 mM
1 M KCl	2 mL	200 mM
1 M $MgCl_2$	1 mL	100 mM
14 M β-mercaptoethanol	43 μL	60 mM
Glycerol	4.4 mL	44% (v/v)
Distilled water (sterile)	1.6 mL	-

B.3.4 DNA polymerase I reaction buffer (10x).

The buffer was made according to the following table and stored at -20°C. DNA polymerase I enzyme can be diluted in 50 mM potassium-phosphate buffer pH 7.2 / 44% glycerol.

Stock solution	per 10 mL	Final conc.
1 M Tris-HCl pH 7.6	1 mL	100 mM
1 M $MgCl_2$	1 mL	60 mM
5 M NaCl	1 mL	0.5 M
14 M β-mercaptoethanol	50 μL	70 μM
Distilled water (sterile)	6.95 mL	-

B.3.5 *Bal31* dilution/ storage buffer.

The buffer was made according to the following table and stored at -20°C.

Stock solution	per 10 mL	Final con.
1 M Tris-HCl pH 8.0	0.2 mL	20 mM
1 M $CaCl_2$	50 μL	5 mM
1 M $MgCl_2$	50 μL	5 mM
0.5 M EDTA pH 8.0	20 μL	1 mM
5 M NaCl	0.2 mL	100 mM
Glycerol	4.4 mL	44% (v/v)
Distilled water (sterile)	5.1 ml	-

B.3.6 *Bal31* reaction buffer (5x).

The buffer was made according to the following table and stored at 4°C.

Stock solution	per 10 mL	Final conc.
1 M Tris-HCl pH 8.0	1 mL	100 mM
1 M $CaCl_2$	0.6 mL	60 mM
1 M $MgCl_2$	0.6 mL	60 mM
0.5 M EDTA pH 8.0	0.1 mL	5 mM
5 M NaCl	6 mL	3 M
Distilled water (sterile)	1.7 mL	-

B.3.7 T4 ligase dilution buffer.

The buffer was made according to the following table and stored at -20°C.

Stock solution	per 10 mL	Final conc.
1 M Tris-HCl pH 7.6	0.2 mL	20 mM
1 M KCl	0.6 mL	60 mM
0.5 M EDTA pH 8.0	2 µL	1 mM
0.5 M DTT	10 µL	5 mM
Glycerol	4.4 mL	44% (v/v)
Distilled water (sterile)	5.2 mL	-

B.3.8 T4 Ligase reaction buffer (5x).

The buffer was made according to the following table and stored in aliquots at -70°C.

Stock solution	per mL	Final conc.
1 M Tris-HCl pH 7.6	250 µL	250 mM
1 M $MgCl_2$	50 µL	50 mM
0.1 M ATP	50 µL	5 mM
0.5 M DTT	50 µL	25 mM
Distilled water (sterile)	600 µL	-

B.3.9 Mung Bean nuclease dilution buffer.

The buffer was made according to the following table and stored in aliquots at -20°C.

Stock solution	per mL	Final conc.
1 M NaOAc pH 4.6	100 µL	10 mM
10 mM ZnOAc	10 µL	0.1 mM
50 mM Cysteine	20 µL	1 mM
Triton X-100 (0.1%, v/v)	100 µL	0.001 % (v/v)
Glycerol	440 µL	44.0 % (v/v)
Distilled water (sterile)	330 µL	-

B.3.10 Mung Bean nuclease reaction buffer (5x).

The buffer was made according to the following table and stored at -20°C.

Stock solution	per mL	Final conc.
1 M NaOAc pH 4.6	300 µL	300 mM
5 M NaCl	100 µL	500 mM
0.1 M $ZnCl_2$	100 µL	10 mM
Glycerol	440 µL	44% (v/v)
Distilled water (sterile)	60 µL	-

B.3.11 Calf Intestinal Alkaline Phosphatase (CIP) buffer. (10x)

The buffer was made according to the following table and stored in aliquots at -20°C.

Stock solution	per mL	Final conc.
1 M Tris-HCl pH 9.0	500 µL	500 mM
1 M $MgCl_2$	10 µL	10 mM
0.1 M $ZnCl_2$	10 µL	1 mM
0.1 M spermidine	100 µL	10 mM
distilled water (sterile)	380 µL	-

B.3.12 T4 kinase reaction buffer (5x).

The buffer was made according to the following table and stored at -20°C. When this buffer is used in P-exchange reactions, potassium-phosphate (pH 8.0) has to be included at a final concentration of 0.08 mM.

Stock solutions	per mL	Final conc.
1 M Tris-HCl pH 8.0	660 µL	660 mM
1 M $MgCl_2$	80 µL	80 mM
0.5 M DTT	160 µL	80 mM
10 mM EDTA pH 8.0	10 µL	0.1 mM
Distilled water (sterile)	90 µL	-

B.3.13 Reverse transcriptase (RT) reaction buffer (1x).

The buffer was made according to the following table and stored at -20°C in an RNAse free Eppendorf. The $MgCl_2$ solution was DEPC treated whereas Tris-HCl Buffer and DTT were prepared using DEPC treated distilled water.

Stock solution	per mL	Final conc.
1 M Tris-HCl pH 8.3	50 µL	50 mM
0.1 M $MgCl_2$	80 µL	8 mM
0.1 M DTT	20 µL	2 mM
Distilled water (sterile)	850 µL	-

B.3.14 Mutagenesis fill-in reaction buffer (10x).

(Stanssens *et al.*, 1989)

The buffer was made according to the following table and stored at -20°C.

Stock solution	per mL	Final conc.
1 M Tris-HCl pH 7.5	275 µL	275 mM
1.5 M KCl	417 µL	625 mM
1 M $MgCl_2$	150 µL	150 mM
0.5 M DTT	40 µL	20 mM
0.1 M ATP	5 µL	0.5 mM
25 mM dATP	10µL	0.25 mM
25 mM dCTP	10µL	0.25 mM
25 mM dGTP	10µL	0.25 mM
25 mM dTTP	10µL	0.25 mM

B.3.15 Nick-translation buffer (10x).

The buffer was made according to the following table and stored at -20 °C.

Stock solution	per 5 mL	Final conc.
1 M Tris-HCl pH 7.8	2.5 mL	500 mM
1 M $MgCl_2$	250 µL	50 mM
14 M β-mercaptoethanol	3.6 µL	10 mM
BSA (10 mg/ mL)	250 µL	500 µg/ mL
Distilled water (sterile)	2 mL	-

B.3.16 Buffers for the extraction of DNA and for general use in DNA recombinant methods

TE (Tris/EDTA) buffer (100x):

Stock solution/ reagent	per L	Final conc.
Tris-HCl	121 g	1 M
0.5 M EDTA pH 8.0	200 mL	100 mM
Distilled water	to 1 L	-

The pH was adjusted to pH 7.6 and the buffer was autoclaved after which it was diluted appropriately in distilled sterile water.

TES (Tris/EDTA/sucrose) buffer:

Stock solution	per 400 mL	Final conc.
1 M Tris-HCl pH 8.0	20 mL	50 mM
0.5 M EDTA pH 8.0	40 mL	50 mM
sucrose (60%, w/v)	100 mL	15% (w/v)
Distilled water (sterile)	240 mL	-

TET (Tris/EDTA/Triton) buffer:

Stock solution	per 400 mL	Final conc.
1 M Tris-HCl pH 8.0	20 mL	50 mM
0.5 M EDTA ph 8.0	40 mL	50 mM
Triton X-100	8 mL	2% (v/v)
Distilled water (sterile)	332 mL	-

Lysozyme buffer :

Stock solution	per 200 mL	Final conc.
1 M Tris-HCl pH 8.0	6 mL	30 mM
0.5 M EDTA pH 8.0	20 mL	50 mM
5 M NaCl	2 mL	50 mM
Sucrose (60% w/v)	83 mL	25% (w/v)
Distilled water (sterile)	89 mL	-

SDS (sodium dodecyl sulphate) buffer:

Stock solution	per 400 mL	Final conc.
0.1 M Tris-HCl pH 8.0	34.4 mL	8.6 mM
0.5 M EDTA pH 8.0	50 mL	625 mM
SDS (10 % w/v)	50 mL	(1.25% w/v)
Distilled water (sterile)	265.6 mL	-

SMM (sucrose/ maleic acid/ magnesium chloride) buffer (2x):

sucrose	342.3 g
maleic acid	4.64 g
$MgCl_2$	8.12 g
Distilled water	to 1000 mL

The pH of the buffer was adjusted to pH 6.5 with 10 N NaOH and the buffer was autoclaved.

PEG 6000 (poly ethylene glycol 6000) buffer:

PEG 6000	40 g
2x SMM buffer	50 mL
Distilled water	to 100 mL

The PEG was dissolved by heating in the 2x SMM buffer and some of the distilled water, then made up to volume and autoclaved.

SMMP medium/ solution:

Mix equal volumes of: 4x PenAssay Broth (PAB)
2x SMM buffer

B.3.17 DNA sample loading solutions.

For agarose gels:		
	Sucrose	(8%, w/v)
	Bromophenol blue	(0.04%, w/v)
	EDTA	(10 mM)

(RNAse at a final concentration of 20 μg/ mL was added when required, e.g. for the degradation of RNA in small-scale plasmid DNA preparations)

For acrylamide gels:

Formamide	(99%, v/v)
Bromophenol blue	(0.01%, w/v)
xylene cyanol	(0.01%, w/v)

(The sequencing tracking dye, supplied in the Sequenase kit, was also frequently used)

B.3.18 Ethidium bromide solution. An ethidium bromide (2,7-diamino-10-ethyl-9-phenyl-phenanthidinium bromide) solution of 0.1% (w/v) was prepared and stored in a dark bottle at 4°C.

B.3.19 Phenol solution (buffer saturated). Phenol (250 g, Merck) was melted at 45°C and the following was added:

8-hydroxyquinoline	0.3 g
1 M Tris pH 8.0	32 mL
Distilled water	62.5 mL

Aliquots (10 mL or 1mL) were frozen at -20°C.

B.3.20 Tris-acetate buffer (50x).

Tris base	242 g
Acetic acid	57.1 mL
EDTA (0.5 M, pH 8.0)	100 mL
Distilled water	to 1000 mL

B.3.21 Z-buffer. (Miller *et al.* 1972)

The buffer was made according to the table below and was stored at 4°C (adjusted to pH 7.0).

Stock solution	per 1000 mL	Final conc.
1 M $Na_2HPO_4.7H_2O$	60 mL	60 mM
1 M $NaH_2PO_4.H_2O$	40 mL	40 mM
1 M KCl	10 mL	10 mM
0.1 M MgSO4.7H_2O	10 mL	1 mM
14 M β-mercaptoethanol	3.6 mL	50 mM
Distilled water	876 mL	-

B.3.22 Buffers for the γ-glutamyl transferase (GGT) assay.

Stock solutions:

1 M Imidazole buffer	6.808 g/ 100 mL
0.8 M Hydroxylamine-HCl	5.560 g/ 100 mL
0.1 M $MnCl_2$	1.979 g/ 100 mL
0.28 M Potassium arsenate	5.040 g/ 100 mL
0.2 M glutamine	0.292 g/ 10 mL
0.1 M ADP (sodium-salt)	0.017 g/ mL

Resuspension buffer :

The buffer was made according to the following table and stored at 4°C.

Stock solution	per 200 mL	Final conc.
1 M Imidazole pH 7.15	4 mL	20 mM
1 M $MgCl_2.6H_2O$	2 mL	10 mM
14 M β-mercaptoethanol	28 μL	2 mM

Assay reaction buffer (5x):

A fresh assay reaction mixture was prepared using the following stock solutions:

Stock solution	per 20 mL	Final conc.
1 M Imidazole	3.71 mL	185 mM
0.8 M Hydroxylamine	0.61 mL	24.5 mM
0.1 M $MnCl_2$	74 μL	0.37 mM
0.28 M Potassium arsenate	2.47 mL	35 mM
0.28 M ADP (sodium salt)	247 μL	3.5 mM
Distilled water	12.9 mL	-

Reaction termination solution:

$FeCl_3.6H_2O$	55 g
Trichloroacetic acid (TCA)	20 g
HCl (concentrated)	21 mL
Distilled water	to 1000 mL

B.3.23 Folin-Lowry assay solutions.

A)	Na2CO3	(2%, w/v)	1 g
	Sodium citrate	(2%, w/v solution	0.5 mL
	CuSO4.5H20	(1%, w/v solution	0.5 mL
	Distilled water		to 50 mL

B)	Folin & Ciocalteu's phenol reagent	5 mL
	H_2O	5 mL

The solutions was prepared immediately before use. Folin & Ciocalteu's phenol reagent was obtained from BDH (No.19058)

B.2.24 Northern blotting buffers and reagents.

MOPS (3-[N-morpholino] propanesulfonic acid) buffer (10x):

MOPS	41.8 g
NaOAc (3M, pH 5.2)	16.6 mL
EDTA (0.5 M, pH 8.0)	20 mL
Distilled water	to 1000 mL

MOPS was pre-dissolved in 800 mL H2O and the intermediate pH was adjusted to pH 7.0 with 5 N NaOH before sodium acetate and EDTA were added.

SSPE buffer (20x):

NaCl	174 g
NaH2PO4.H2O	27.6 g
EDTA	7.4 g
Distilled water	to 1000 mL

The pH was adjusted to pH7.4 with 5 N NaOH

Denhardt's solution (50 x):

Ficoll	1 g
Polyvinylpyrolidine (PVP)	1 g
Bovin serum albumin (BSA)	1 g
Distilled water	to 100 mL

Prehybridisation solution:

Stocksolution	per 100 mL	Final conc.
20x SSPE	30 mL	6x
SDS (10% w/v)	5 mL	0.5% w/v
50x Denhardt's	10 mL	5x
Salmon sperm DNA (10 mg/mL)	1 mL	100 µg/mL
Distilled water (sterile)	54 mL	-

Salmon sperm solution was prepared by dissolving 0.1 g Salmon sperm DNA in 10 mL TE buffer. After the DNA solution was sonicated at full power (20 microns) for 10 min in an MSE Soniprep sonicator, the solution was aliquotted and stored at -20°C. Salmon sperm DNA was denatured by boiling for 10 min and cooled on ice before adding to the prehybridisation solution.

Appendix C

General Techniques

C. 1 Preparation of *E. coli* plasmid and phasmid DNA.

Plasmid and ds phasmid DNA were isolated from 1.5 mL overnight cultures (LB + appropriate antibiotic) according to modified procedures described by Gough *et al.* (1980). The cells were collected in a 1.5 mL microfuge tube by centrifugation (1 min, Eppendorf microfuge). The pellet was resuspended in 0.2 mL TES buffer plus 20 ml lysozyme solution (10 mg/mL) and held on ice for 15-20 min. After addition of ice cold H_2O (700 µL), the suspension was gently mixed and was allowed to stand on ice for not more than 5 min, followed by heating to 73°C for 2 min. Phenol (70 µL) was added and the two phases were carefully mixed. Precipitated cellular debris and denatured chromosomal DNA were pelleted by centrifugation for 30 min. Traces of phenol were removed by repeated extraction in chloroform/isoamyl alcohol (24:1). Each time the tube was briefly vortexed and centrifuged. The supernatant was retained. Plasmid (or ds phasmid) DNA was precipitated from the supernatant (0.5 mL) by addition of isopropanol (0.5 mL) and 5 M $NaClO_4$ (0.1 mL). The precipitate was washed with 70 % ethanol, air dried, and resuspended in 100 µL TE buffer. This method yielded approximately 2-3 µg DNA/culture which was sufficiently pure for restriction endonuclease digests.

Large scale preparations of *E. coli* plasmid DNA were performed according to the method of Clewell *et al.* (1972). Harvested cells from a 200 mL overnight culture (LB) were suspended in 7 mL TES buffer plus 0.8 mL lysozyme buffer (10 mg/mL). After incubation on ice for 15-20 min, 6.5 mL of TET buffer was added. The mixture was held on ice for a further 10-15 min and cellular debris was removed by centrifugation (19 500 rpm; Beckmann SS34) for 30 min at 4°C. Per mL supernatant, 0.1 mL ethidium bromide (Etbr; 10 mg/mL) and 1 g CsCl were added and the solution was centrifugated (13 000 rpm; Beckmann HB4). The liquid phase was transferred to a Beckmann Quickseal ultra-centrifuge tube and centrifugated (60 000 rpm; Beckmann Vti 65.2) for a minimum of 6 h. The plasmid band was visualized by long wave UV light (310 nm) and removed in as small a volume as possible. If the DNA was to be used for sequence analysis, a second CsCl centrifugation was preferred. The plasmid DNA was precipitated by adding 2 volumes of H_2O and 3 volumes of isopropanol, followed by centrifugation (10 000 rpm; Beckmann HB4) for 15 min. The DNA pellet was dissolved in

700 µL TE buffer and Etbr was removed by repeated phenol (700 µL) extraction. Traces of phenol were twice extracted in chloroform/isoamyl alcohol (24:1), and the DNA precipitated in a 1/2 volume of isopropanol plus 1/10 volume of 5 m $NaClO_4$. After centrifugation (10 000 rpm; Eppendorf micro centrifuge), the DNA pellet was dissolved in 300 µL TE buffer and the concentration determined spectrophotometrically as described below. Typical yields of 150 µg (for pBR-derivatives) to 500 µg (for pUC-derivatives) were obtained from 200 mL cultures.

C.2 Preparation of *B. subtilis* plasmid DNA. Cells were harvested from a 1.5 mL culture (PAB) and resuspended in 0.1 mL lysozyme buffer containing 1 mg/mL lysozyme. After incubation at 37°C for 15 min, SDS buffer (0.1 mL) and 5 M NaCl (40 µL) were added and the mixture was placed on ice for 45 min, followed by centrifugation (19 500 rpm; Beckmann SS34) for 20 min. The supernatant was tranferred to a new tube and DNA was precipitated by the addition of one volume isopropanol, followed by centrifugation (10 000 rpm; Eppendorf micro centrifuge) for 10 min. The DNA pellet was dissolved in 50 µL TE. When this DNA (yield approximately 1 µg per mL culture) was used for restriction endonuclease digestion, one phenol and one chloroform/isoamyl alcohol extraction was performed and the DNA was precipitated with 1/2 volume of isopropanol and 1/10 volume of 5 M $NaClO_4$.

The method of Williams *et al.* (1981) was used for large-scale extraction of plasmid DNA from *B. subtilis* cells. Harvested cells of a 250 mL culture (PAB) were resuspended in 4 mL lysozyme buffer (Appendix B) containing 1mg/mL lysozyme and incubated at 37°C for 30 min. SDS buffer (Appendix B) (16 mL) and 5 M NaCl (5 mL) were added and the mixture was placed on ice overnight. The next day, cell debris and chromosomal DNA are removed by centrifugation (19 500 rpm; Beckmann SS34) for 30 min. The supernatant was mixed with 5 mL of PEG6000 solution (50%) and held on ice for 4 h, which was followed by centrifugation (10 000 rpm; Beckmann HB4) for 10 min. The pellet was dissolved in 2 mL TE buffer, 0.2 mL RNase (2 mg/mL) was added, and the mixture was placed at 65°C for 30 min. Finally, after addition of 0.2 mL Etbr (10 mg/mL), the volume was adjusted to 4 mL with TE buffer and 3.65 g CsCl was added. This mixture was loaded in a Beckmann Quickseal ultra-centrifuge tube and centrifugated (60 000 rpm; Vti 65.2). Visualization of the plasmid band, extraction of Etbr, and subsequent precipitation of plasmid DNA was performed as explained in C.1.

C.3 Determination of DNA concentration. The concentration of DNA solutions was determined spectrophotometrically by monitoring the absorbance of the solutions between 220 and 310 nm. The concentration was determined by using the conversion where 1 absorbance unit at 260 nm is equivalent to 50 μg DNA/mL for ds DNA and 30 μg DNA/mL for ss DNA.

C.4 Restriction endonuclease digests. Restriction digests were carried out using one of the four restriction buffers (Appendix B) according to the salt and temperature requirements of the particular enzyme (New England Biolabs Catalog 1988/1989). A specific *SmaI*-buffer (Appendix B) was used for *SmaI* restriction endonuclease. Multiple restrictions, requiring different enzymes, could be combined in a single digest provided the salt requirements of the enzymes were compatible. If this was not possible, then the digestion was done sequentially using the enzyme with the lowest salt and highest temperature requirement first and the salt and incubation temperature were adjusted before the addition of the next enzyme. Typically 200-300 ng plasmid DNA was digested in a 25 μL volume using 1 unit of enzyme for 1 h. Concentrated enzyme stocks were diluted to 1 or 2 units/μL using a universal restriction enzyme dilution buffer (Appendix B). If necessary, DNA was purified in between enzymatic reactions using the Quick Clean-Up Method (see below). For electrophoretic analysis, digestions were stopped by the addition of DNA sample loading solution (Appendix B).

C.5 Quick Clean-up Method. This method is a micro-scale clean-up of DNA in order to perform many sequential operations on the same basic DNA solution. The volume of a typical enzyme reaction mixture (20- 50 μL) was adjusted to 100 μL with TE buffer, and 15 μL phenol was added. The phases were mixed well (vortexed) and immediately 110 μL chloroform/isoamyl alcohol (24:1) was added. Phases were mixed well again and the solution was centrifugated (10 000 rpm; Eppendorf micro centrifuge) for 5 min. To the supernatant, 10 μL NaClO4 (5 M) and 60 μL isopropanol were added and the DNA was precipitated by centrifugation (10 000 rpm; Eppendorf micro centrifuge). To assist precipitation of small quantities of DNA (< 2 μg), 5 μg tRNA was added. Generally, this did not affect further enzymatic reactions. The DNA pellet was finally dissolved in 20-50 μL.

C.6 Agarose gel electrophoresis. Agarose gel electrophoresis was carried out using a horizontal submerged gel system similar to that described by Maniatis *et al.* (1982). Tris-acetate buffer (Appendix B) was used routinely. Sigma type II agarose was used in varrying concentrations (0.8 - 2.0 % w/v) depending on the sizes of the fragments being examined. The amount of DNA loaded per lane also varied with the sizes and number of fragments but under normal circumstances 200-300 ng of plasmid DNA was used, which was electrophoresed at 2 V/cm for 16 h. For 'quick' gels a Hoefer Minnie Submarine agarose gel unit (model HE33; Hoefer Scientific Instruments, USA) was used at 5 V/cm for 2-3 h. Gels were stained in electrophoresis buffer containing 4 μg/mL Etbr for 15 min. DNA bands were visualized using a 254 nm transilluminator (Chromato-Vue, Calif., USA). A 310 nm transilluminator was used if the DNA had to be recovered from the gel and used for further manipulations.

Gels were photographed using a Polaroid CU-5 Land camara fitted with a red filter and a fixed focal length attachment. Polaroid type 667 film (ASA 3 000) was used with an exposure time of 1-2 s at f4.7. If a better quality photograph and a negative were required, then Polaroid type 665 film (ASA 64) with an exposure time of 120-140 s at f4.7 was used.

Fragment sizes were calculated by extrapolation from a standard curve of the mobility of lambda DNA fragments, digested with either *HindIII*, *PstI*, or *BstEII*, plotted against the $\log_{10}$ of their molecular weights.

C.7 Ligation reactions. The method of Maniatis *et al.* (1982) was routinely used. DNA concentrations in the reactions varied depending on the aim of the ligation experiment. Recircularization reactions for the isolation of deletion plasmids contained DNA concentrations of 1 pmole/mL or less. A DNA concentration of approximately 5 pmole/mL was used for recombination reactions where limited polymerization with circle closure was required. To improve cloning efficiency of DNA fragments into vector DNA, a molar insert to vector ratio of 3:1 was applied. Typically, a ligation reaction was performed in a 25 µL volume containing DNA, ligation buffer (Appendix B), and water. Sticky-end ligations were performed at room temperature (19-21°C) for 2 h using 0.1 U of T4 ligase (Boehringer Mannheim, FRG). Blunt-end ligations were performed at 15°C for 16 h (overnight) using 1 U of T4 ligase. In some cases, the ligation reaction was analysed *in vitro* on agarose (1%, w/v) gels in the presence of 20 µg/mL Etbr.

C.8 Rapid subcloning techniques. Crouse *et al.* (1983) have described a method in which specific restriction fragments were separated by gel electrophoresis in low-melting-point (LMP) agarose and ligated directly to appropriately endonuclease-digested vector DNA in the presence of agarose. The following method (Struhl, 1985) was used. DNA fragments were excised from a LMP agarose (SeaPlaqueR , Marine Colloids, Rockland, USA) gel in as small a volume as possible. The gel slices were melted at 70°C for 5 min, equilibrated at 37°C, and combined in appropriate proportions in a total volume of 10 µL. An equal volume of 2x concentrated ligation buffer containing T4 ligase was added, the ligation mixed quickly, and incubated at room temperature (19-21°) for 2 to 16 h. Prior to transformation of *E. coli* cells, the gel mixture was heated to 70°C for 5 min, and then diluted with 10 volumes of 0.1 M $CaCl_2$. When large quantities of linearized vector DNA (5-10 µg) was electrophoresed over LMP agarose, it was preferred to purify vector DNA with phenol and subsequent chloroform extractions (2x) and precipitation after which the DNA was stored ready-for-use in 100 µL TE buffer at 4°C.

C.9 Transformation of *E. coli* and *B. subtilis* cells.

C.9.1. Preparation and transformation of competent *E. coli* cells. Competent cells of *E. coli* were prepared essentially as described by Dagert and Ehrlich (1979). An overnight culture of *E. coli* was diluted 1/100 into 50 ml prewarmed (37°C) LB medium and grown with shaking to OD_{600} = 0.2. After the culture was cooled on ice for 5 min, cells were collected by centrifugation (5 000 rpm; Beckmann HB4) at 4°C for 5 min, washed with 50 ml ice cold 0.1 M $MgCl_2$, and resuspended in 25 mL of ice cold 0.1 M $CaCl_2$. After 20-60 min on ice the cells were collected by centrifugation and resuspended in 5 mL 0.1 M $CaCl_2$. The competent cells were kept on ice for at least 1 h before use or aged overnight at 4°C to improve their competency. For long term storage, competent cells were treated with glycerol to a final concentration of 15% (v/v) and the cells frozen at -70°C. Competent cells prepared in this way were suitable for transformation for more than 1 year.

Competent *E. coli* cells were transformed as follows. Plasmid DNA (1-5 ng) or a fraction of a ligation reaction (20-100 ng DNA) was added to 100 µL of competent cells on ice. After 10 min the cells were induced to take up the DNA by heat-shocking the transformation mix at 42°C for 2 min. One mL of LB medium was added and the transformation mix left at 37°C to allow expression of the plasmid borne antibiotic marker (Ap: 15 min; Cm:30 min; Kn:30 min; Tet:45 min). Controls included: competent cells with the addition of 1) no DNA, 2) pBR322 plasmid DNA, and 3) linearized and ligated vector DNA. Transformation frequencies of 3000-8000 cfu/ng pBR322 DNA were routinely obtained.

C.9.1 Preparation, transformation, and regeneration of *B. subtilis* protoplasts. Protoplasts were prepared from *B. subtilis* cells as follows (Chang and Cohen, 1979; Levi-Meyrueis *et al.*, 1980). Prewarmed (37°C) PAB (Appendix B) (50 mL) was inocculated with a 200 µL overnight culture and the *B. subtilis* cells were further grown on 37°C for approximately 3 h to OD_{600} 0.3-0.4. Cells were harvested by centrifugation (5 000 rpm; Beckmann SS34) for 5 min, and resuspended in 4 mL SMMP (Appendix B) containing 0.5 mg/mL lysozyme. The suspension was incubated for 45-120 min at 37°C with gentle shaking and development of protoplasts was monitored by microscopy. Protoplasts were harvested by centrifugation (4 000 rpm; Beckmann SS34) for 5 min and washed with 5 mL SMMP. The protoplasts were again harvested and resuspended in a final volume of 4 mL SMMP, in which protplasts were stable for at least 3 h at room temperature (19-21°C). Alternatively, protoplasts were frozen and stored at -70°C until used.

To transform *B. subtilis* protoplasts, 10-100 ng plasmid DNA in 50 µL TE buffer was mixed with 50 µL of 2X SMMP, and added to 0.5 mL protoplasts at room temperature. PEG6000 solution (Appendix B) was immediately added and the mixture was gently mixed. After 2 min, this preparation was plated on DM3 regeneration agar base (Appendix B), supplemented with the appropriate antibiotic(s).

C.10 *Bal31* shortening of DNA fragments. Bidirectional shortening of linearized DNA was accomplised essentially as described by Misra *et al.* (1985). The enzyme was first calibrated as follows. Linearized DNA (10 µg; 66 µL) was mixed with 5x *Bal31* buffer (50 µL) and water (133 µL), equilibrated at 37°C, and 10 U *Bal31* (Boehringer mannheim, FRG) was added (*Bal31* enzyme was diluted using *Bal31* enzyme dilution buffer - Appendix B). At 1 min inervals 25 µL aliquots (1 µg) were removed and tranferred into tubes containing 75 µL TE, 4 µL tRNA (1mg/mL) and 10 µL phenol. The mixture was well mixed and the DNA was purified according to the Quick Clean-up Method (see C.5). DNA precipitates were resuspended in 18 µL TE. This material was digested with an appropriate restriction endonuclease to analyse the extend of deletion by electrophoresis of the digest on a 1.2% (w/v) agarose gel. In this way, it was estimated which time interval produced fragments of the desired size.

Specific digestion of linearized DNA by *Bal31* endonuclease was repeated under exact the same reaction conditions as above, but only for the required times. Digested DNA fragments were purified as above and resuspended in a final volume of 20 µL TE. Truncated ends were repaired to blunt ends by application of Mung bean nuclease.

C.11 Mung bean nuclease generation of blunt-ended DNA fragments. Mung bean nuclease digests either 5′ or 3′ single-stranded extensions and can be used to blunt-end restriction DNA fragments (or *Bal31* digested DNA, see above). To 16 µL linearized DNA (typically 5 µg), 2 µL Mung bean nuclease reaction buffer (10x ;Appendix B) and 2 µL Mung bean nuclease enzyme (10 U; New England Biolabs, USA) were added (when required, Mung bean nuclease was diluted with Mung bean nuclease dilution buffer - Appendix B). The reaction mixture was incubated at 37°C for 30 min and the DNA was purified according to the Quick Clean-up Method (see C.5).

C.12 Fill-in reactions with Klenow fragment. DNA fragments with 5′-protruding termini were blunt-ended using the 'filling-in' polymerisation at the 3′-ends by the large fragment (Klenow fragment) of DNA polymerase I. Restriction endonuclease-digested DNA (typically 5 μg) was dissolved in 25 μL and 3 μL DNA polymerase I reaction buffer (10x ; Appendix B), 3 μL dNTP (10 mM; 0.25 mM final for each dATP, dCTP, dGTP, and dTTP), and 1 μL of Klenow fragment (1 U/μL) (Boehringer Mannheim, FRG) were added. DNA was purified according to the Quick Clean-up Method described in section C.5 after the reaction was incubated for 10 min at room temperature.

C.13 Dephosphorylation reactions. To prevent recircularization of vector DNA, and when biochemical or biological selection against parental constructs was not possible, 5′- phosphates of linearized vector DNA were removed by the use of Calf Intestinal Phosphatase (CIP). Ligation of such termini will only occur with DNA that has not been treated with CIP, thus considerably improving the efficiency of cloning insert DNA into vector DNA. After linearized vector DNA was dissolved in 10 μL water (for precipitation of DNA, tRNA was omitted because CIP also acts upon RNA), 2 μL CIP reaction buffer (10x; Appendix B) and CIP enzyme (0.05 U/ pmole ends) were added. The reaction was incubated at 37°C for 30 min, after which the DNA was purified by phenol and subsequent chloroform extractions, and precipitated with 1/2 volume of isopranol plus 1/10 volume of $NaClO_4$ (5 M).

C.14 Preparation of radio-active probes.

C.14.1 Nick-translation reaction. DNA probes were labelled with [α-^{32}P]dATP to high specific radioactivity by nick-translation (Rigby *et al.*, 1977). The required restriction DNA fragment (typically 1-1.5 μg) was purified over LMP agarose as described in section C.8 and dissolved in 30 μL TE. After addition of 5 μL nick-translation buffer (10x; Appendix B), 1 μL of each dGTP, dCTP, and dTTP (2.5 mM), 5

μL [α^{32}P]-dATP (3000 Ci/mmole; Amersham, UK), 1 μL DNase (0.1 μg/mL), and 1 μL DNA polymerase I (1 U/μL; Boehringer Manheim, FRG), the reaction mixture was incubated at 16°C for 2 h. Non-incorporated triphosphates were separated from the probe by the use of a Sephadex G50 spin column as described by Maniatis *et al.* (1982). Specific activities of approximately 10^7 cpm/μg of DNA were routinely obtained. Radioactively labelled probes were stored in lead containers at -20°C.

C.14.2 Kination reaction. T4 polynucleotide kinase (Boehringer Mannheim, FRG) was used to phosphorylate 5'-termini of oligonucleotides used in primer-extension reactions and site-directed mutagenesis. Typically, 15 pmoles of oligonucleotide were diluted in water to a final volume of 38 μL. After addition of 10 μL T4 kinase reaction buffer (Appendix B), 2μL [τ-32P]ATP (32 μCi/μL; D. Botes, Department of Biochemistry, University of CapeTown) and 1 μL T4 kinase (1 U/μL), the reaction mixture was incubated at 37°C for 30 min. The T4 kinase enzyme was inactivated by heating on 60°C for 10 min, non-incorporated triphosphate was removed as above (section C.14.1) and the ^{32}P-5' end-labelled oligonucleotide was stored in a lead container at -20°C. Extreme care was taken when radioactive labeled primers were used for the reverse transcription of RNA and all reagents (except Tris-HCl and DTT) were DEPC (0.1% v/v) treated. Phosphorylation of the mutagenic oligonucleotide, used in Chapter 5, was performed as above, except that non-radioactive ATP was used.

C.15 Northern blotting of RNA.

C.15.1 Formaldehyde agarose gel electrophoresis. To create RNase-free conditions, the glass plates, comb, and electrophoresis tank were rinsed with formamide and a formaldehyde agarose (1.2% w/v) gel was prepared as follows. Agarose (18 g) was dissolved in 15 mL MOPS buffer (10x; Appendix B) and 110 mL H_2O by heating to 90°C. This mixture was cooled to 50°C and 25 mL formaldehyde (37%) was added, after which the gel was poured immediately. The gel was covered by a glass plate to avoid diffusion of formaldehyde into the running buffer (1 x MOPS). The RNA (typically 10 µg), previously dissolved in 12 µL H2O, was diluted in 25 µL formamide, 5 µL MOPS buffer (10x) and 8 µL formaldehyde, heated at 65°C for 5 min, and chilled on ice. The RNA was mixed with 5 µL loading buffer (6x; Appendix B) and electrophoresed at 6 V/cm for 5 h.

C.15.2 Transfer of RNA from the agarose gel to nylon membrane. A glass plate, previously rinsed with formamide, was used as support for the gel. Four pieces Whatmans 3MM filter paper, of which the top one was soaked in 20x SSPE buffer (Appendix B), were layered on the glass plate and the gel was placed onto the filter paper. A nylon membrane (Hybond), soaked in 20x SSPE buffer, was applied to the gel being carefull not to entrap air bubbles. Three more Whatmans 3 MM filter papers (soaked in 20x SSPE buffer) were placed on top, followed by a 5 cm stack of dry paper towel and a second glass plate. A 1 kg weight on the glass plate ensured an even contact, and transfer was allowed to continue for 12 h. The membrane was air dried and the RNA was cross-linked by exposure to UV (254 nm) for 3 min.

C.15.3 Hybridisation and washing procedures. The membrane was soaked in 6x SSPE buffer for 2 min and placed in a plastic bag. Prewarmed (68°C) prehybridization solution (0.2 mL/cm^2 of membrane; Appendix B) was added, the bag sealed (ensuring the exclusion of air) and prehybridization was allowed to proceed at 68°C for 2 h with constant agitation. Thereafter, half the prehybridization fluid was removed, the denatured probe added, the bag resealed and hybridization was continued at 68°C for 10-16 h with constant agitation. Solutions of increasing stringencies were used to wash the membrane. The membrane was washed with 2x SSPE, 0.5% SDS twice at 40°C for 5 min, followed by two washes with 0.1x SSPE, 0.1% SDS at 68°C for 20 min. The wet membrane was sealed in a plastic bag and taped flat under Kodak XAR-5 autoradiography film in an X-ray cassette fitted with a Fuji X-ray intensifying screen. Exposure was allowed to proceed for 2 d at -70°C after which the film was processed using Kodak GBX X-ray developer and fixer according to the manufacturer's instructions.

C.16 The γ-glutamyl transferase assay. The γ-glutamyl transferase (GGT) assay, outlined by Bender *et al.* (1977), was used to measure GS activity. *E. coli* cells, grown in the appropriate media (as described in the text), were harvested by centrifugation (5 000 rpm; Beckmann SS34, or 10 000 rpm; Eppendorf micro centrifuge, depending on the volume of the sample). Typically, 2 x 10^8 to 3 x 10^8 cells were collected (one duplicate was directly froozen and stored at -70°C) and washed with CSH minimal medium (Appendix B) (without ammonium sulphate). After centrifugation, the cells were resuspended in 0.5 mL resuspension buffer (20 mM imidazole pH 7.15, 10 mM MgCl2.6H2O, 2 mM β-mercaptoethanol; Appendix B) and sonicated in an MSE Soniprep sonicator (3 bursts of 10 s on full power with intermediate cooling on ice). The lysate was clarified by centrifugation (10 000 rpm; Eppendorf micro centrifuge) at 4°C for 5 min and the cell-free supernatant was placed on ice until used. Generally, a dilution series was made using the above resuspension buffer. Samples (0.1 mL) were added to 0.4 mL GGT- assay reaction buffer (Appendix B) and equilibrated at 37°C for 5 min. The reaction was initiated by the addition of 50 μL glutamine (0.2 M). After 15 min at 37°C, the reaction was stopped by the addition of 1 mL 'stop mix' (Appendix B) and the absorbance was read spectrophotometrically at 540 nm. A standard curve for this assay, which was drawn up by using L-glutamine acid γ-monohydroxamate (GAMH; Sigma G-2253) as a substrate, gave an absorbance at 540 nm of 0.755 for 1 μmol of glutamyl hydroxamate (Appendix G). GS activity was expressed in μmoles of glutamyl hydroxamate produced per min per mg protein.

If *E. coli* cells were to be assayed for GS and β-galactosidase activities simultaneously, then the cells were sonicated in 0.5 mL water and the cell-free lysate was prepared as described above. For the GGT assay, 0.1 mL lysate was mixed with 0.1 mL resuspension buffer (2x), and for the β-galactosidase assay, 0.2 mL lysate was mixed with 0.2 mL phosphate buffer (50 mM; pH 7.0). Protein concentration determinations (in this particular case) were performed on the original cell-free lysate in water.

C.17 The β-galactosidase assay. β-galactosidase activity in *E. coli* cells was measured using the chromogenic substrate o-nitro-phenyl-β-D-galactoside (ONPG) essentially as described by Miller *et al.* (1972), except that cells were lysed by sonication as described in section C.16. Cell-free lysate (250 μL) was mixed with 250 μL Z-buffer (Appendix B) and equilibrated at 28°C for 5 min. After addition of 100 μL ONPG (4 mg/mL in 0.1 M phosphate buffer pH 7.0), the mixture was further incubated at 28°C and timed until a pale yellow colour developed. The reaction was stopped with 250 μL Na_2CO_3 (14% w/v), 3 mL of water was added, and hydrolysis of ONPG was measured by reading the absorbance spectrophotometrically at 420 nm. The β-galactosidase activity was expressed in μmoles of ONP produced per min per mg of protein. A standard curve was obtained using o-nitrophenyl (ONP) and, under exactly the same conditions as outlined above, an absorbance at 420 nm of 0.0570 corresponded with 50 nmoles of ONP (AppendixG).

C.18 Determination of protein concentration. Protein concentrations in solutions were determined by the method of Lowry *et al.* (1951). The reaction contained protein solution (0.1 mL), 0.1 M NaOH (0.1 mL) and 1 mL Folin-Lowry solution A (Appendix B). After 10 min at room temperature, 0.1 mL of Folin-Lowry solution B (Appendix B) was added followed by vigerous shaking. After 30 min at room temperature, the optical density of the reaction was monitored at 660 nm. Protein concentrations were calculated using a standard curve (BSA fraction V; 0.05-0.7 mg/mL) (Appendix G). Protein samples were diluted such that OD_{660} did not exceed 0.8.

Appendix D

Plasmid Restriction Maps

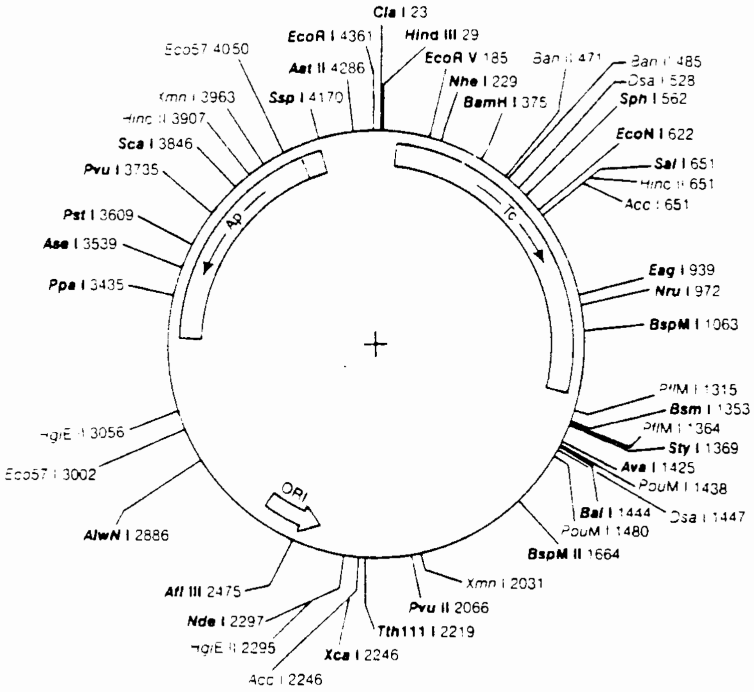


Fig. D.1.: Restriction map of plasmid pBR322 (Bolivar *et al.*, 1977). From Biolabs 1988-1989 Catalogue, New England Biolabs, Inc.

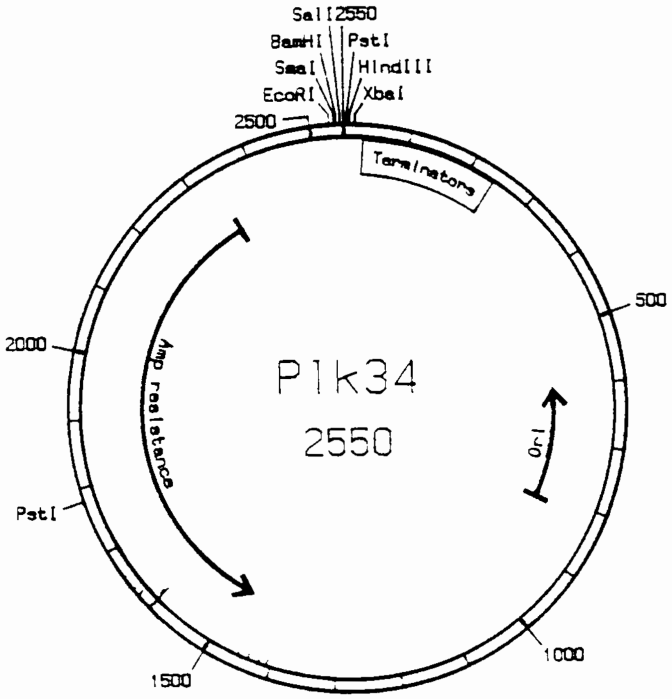


Fig. D.2.: Restriction map of plasmid pLK34 (Botterman *et al.*, 1987)

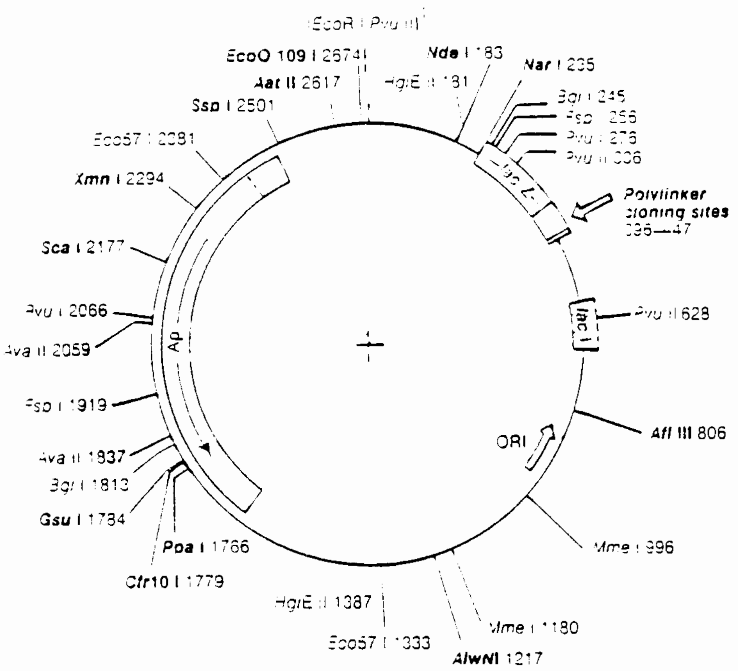


Fig. D.3.: Restriction map of pUC8 and derivatives (Yannisch-Perron *et al.*, 1985). From Biolabs 1988-1989 Catalogue, New England Biolabs, Inc.

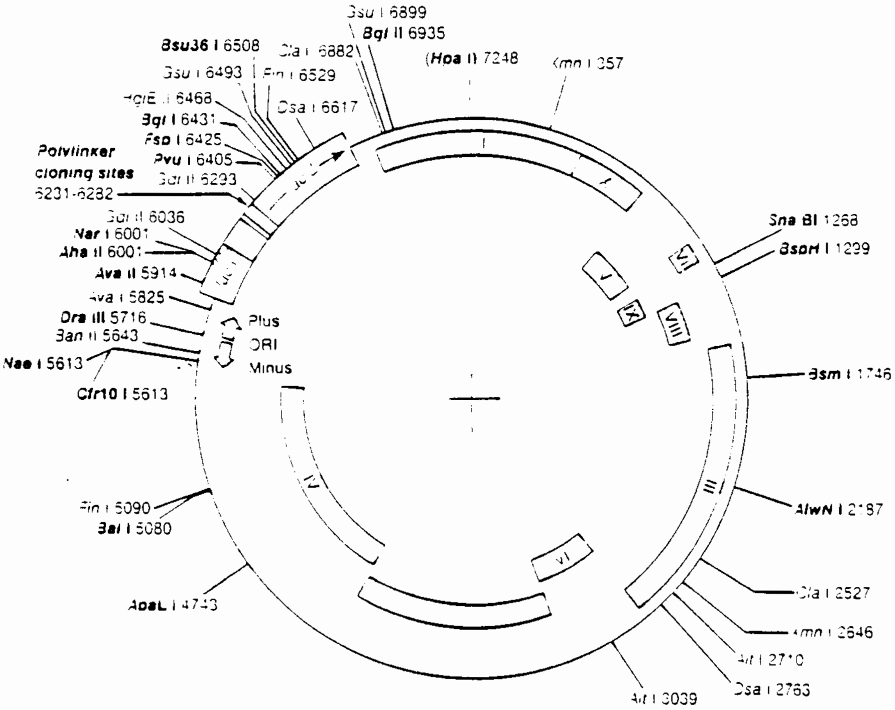


Fig. D.4.: Restriction map of M13mp18 DNA (Messing *et al.*, 1983). From Biolabs 1988-1989 Catalogue, New England Biolabs, Inc.

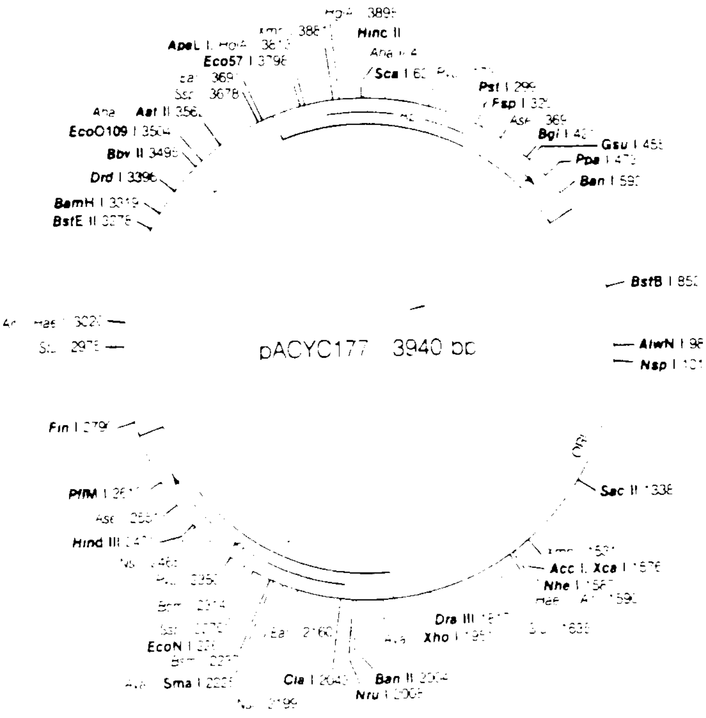


Fig. D.5.: Restriction map of plasmid pAcYc177 (Chang and Cohen, 1978).

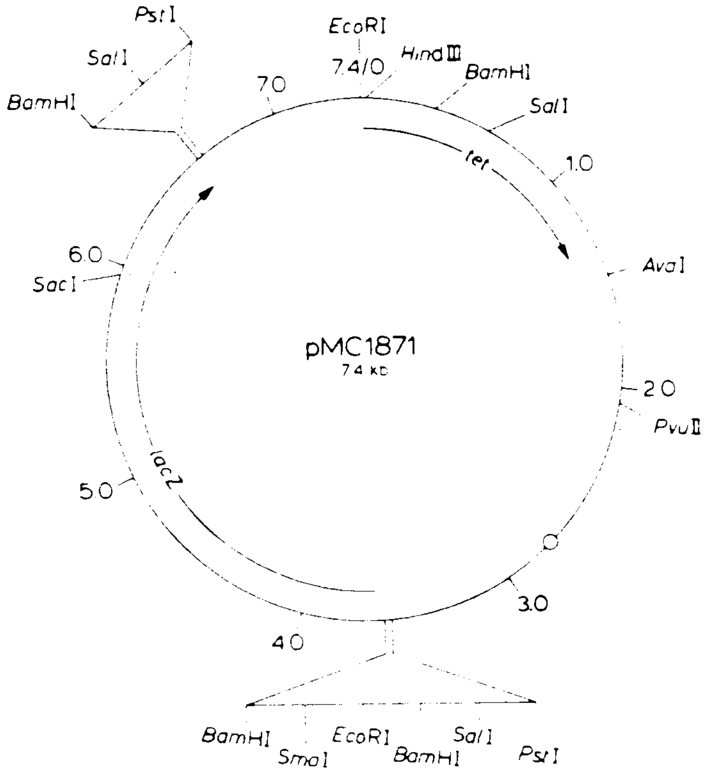


Fig. D.6.: Restriction map of plasmid pMC1871 (Casadaban *et al.*, 1980).

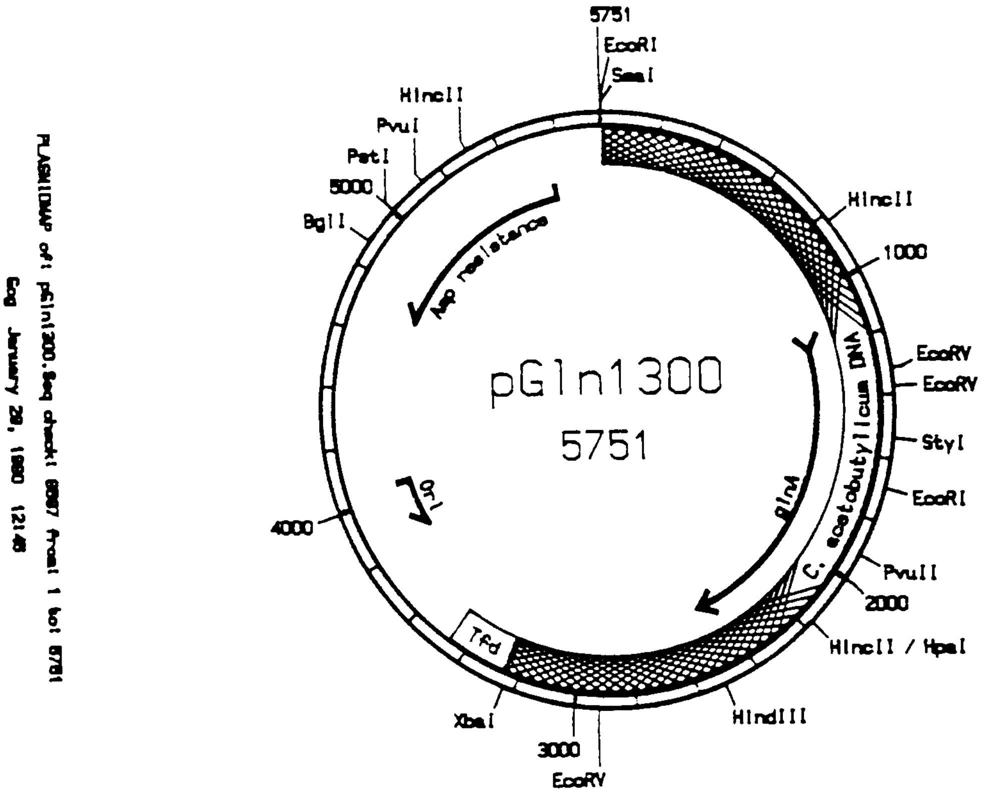


Fig. D.9.: Restriction map of plasmid pGln1300 (This study). This plasmid was constructed by subcloning the *XbaI-BglII* fragment containing the *C. acetobutylicum glnA* gene into the vector pLK34.

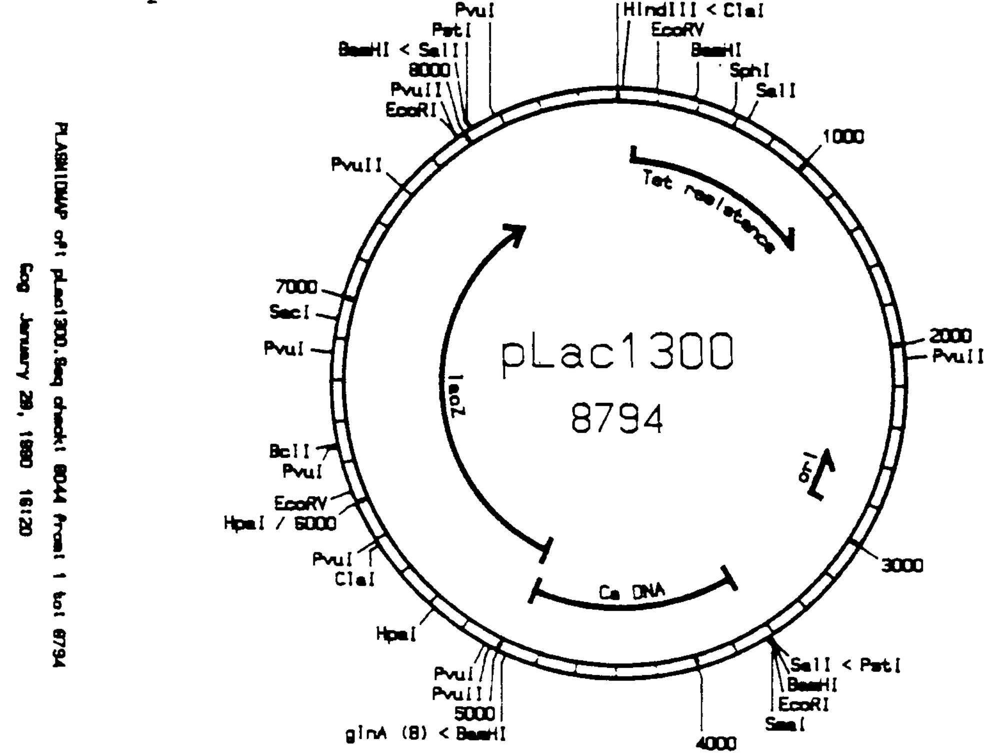


Fig. D.10.: Restriction map of plasmid pLac1300 (This study). The first eight codons of the *C. acetobutylicum glnA* gene is fused to the nineth codon of the *E. coli* lacZ gene.

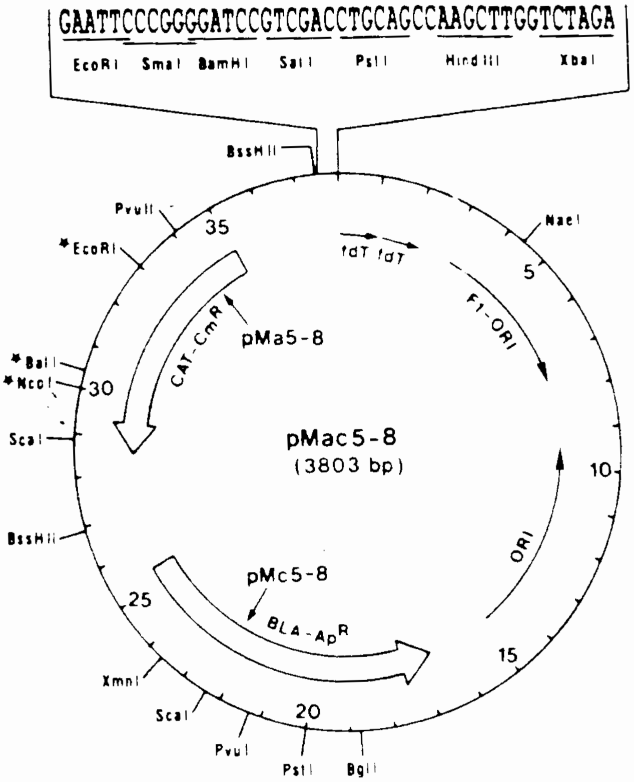


Fig. D.7.: Restriction map of phasmids pMa5-8 (Ap^R Cm^S) and pMc5-8 (Ap^S Cm^R) (Stanssens *et al.*, 1989).

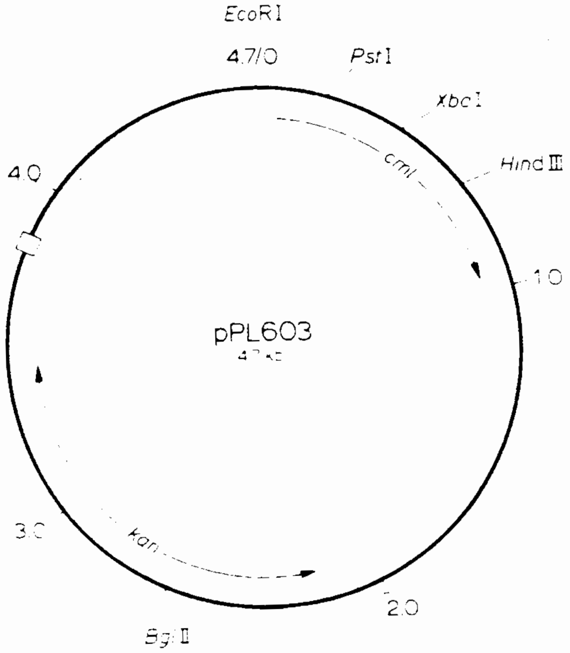


Fig. D.8.: Restriction map of plasmid pPL603b1 (Corfield *et al.*, 1984).

Appendix E

Complete DNA sequence of the *BglII-XbaI* fragment carrying the *C. acetobutylicum glnA* gene

```
AGATCTNTGCCATTAAGAAACCAGAAACATNGGTCCAATNTTCATAAAGATTTGCATTGATNNCTTGCAGCANCTAAT
GCTCCAGTAACATCAGANCCACCTCTTGAGAATGTAACAATGTTACCAGATTTATCAGCACCATAAAATNCAGGAATA
ACAGCCTTAGAAACATTAGATAATCTGTTGTGTAGAGCGCAGTTNNNNGCTTCATTATCTAATGAGCCATCCTTTTNN
AAAACTATNACATCTTNNNCATCAACAAATTCAAANGTCTAANTATTTCNCAAGAATTAACCCATTTAAATATTCNNT
CTCTGCGTTGCAGCGTAATCACTTGATNCACCAGCTTCAAGGTCTTTCTTTATTGTATTTAATTGATCTTCTATACTA
AAATCTAATTTTAAATCATTTATTATATCAGAATATCTTTGTCTTATGTGATTAAAGACATCATCTAAAGCGATTCCA
GCAGCAACATGAGCATGACAAAGGTATAATAAGTCAGTTACTTTTGAATCTTTAGAGTCTCTTTTTCCAGGCGCAGAA
GGTATTACATATTTTCTTGCATCATTAGAATAAATTATGTCTTTAACCTTTCTGAATTGATTGGCATCAGCTAATGAG
CTGCCACCAAATTTTGTTATAATTGTGTTCATAAGTATAGCCCCCAATATATTATAAATTTTATATATAAGATTGTAG
CAATTTATCTAAAAATGTTCAAGATAATTATTAATAATAAAAGAGTATTTAATTGTAGAATTATATTGTGAATAAAAA
TTTTTTTTTGTGCATATCTTAAAATTGAACTAAGAGATATGATAGTTGACAGTATTTAGCATACCTTTTATAATAAAG
TACAGATAAATATATGCTGACAACGAAGTCATAGAGGTACGCCGTCTCGTGACTTCGTTTTTTTTGTAAATTTATGCA
TTATTTATAAAATTTTTTCATTACAATATTATATACAAAATTACAGAATGGAGCATAGTATAATGTTAAATAAAATTT
ATATAATAAAGGATATATATTGTAATAAAAAATGAGACTATAAATTTTTAAAATGGAAAAATAATCCTGATAACTTTA
GAATCTAAAGAAAAACGTATCAAATTATTTATATTGTAAAGTTTTTAGTACGAAAATATTTATTTACTAAAAAAATCA
CTTGATTTCTTAAAAAAAAAGGGAAGTATAATTTAGTTATTATATGAATTGAGCAAAGGCGTTCGATAGAAGTTTATAC
TTGTCTATTGTGCGCTTTTTTTAAAAATGTAAAGGGGGAGTTGTAAAATG.GCA.AAA.TAC.ACA.AAG.GAA.GAT
.ATC.ATT.AAT.TTA.GTA.AAA.GAA.AAC.GGC.GTA.AAA.TTT.ATA.AGA.CTT.CAA.TTT.ACT.GAT.A
TC.TTT.GGA.ACA.TTA.AAG.AAT.GTG.GCG.ATT.ACA.GAC.AAA.CAA.TTG.GAA.AAA.GCA.TTA.GAT
.AAT.GAA.TGT.ATG.TTT.GAT.GGA.TCA.TCT.ATT.GAT.GGT.TTT.GTT.AGA.ATT.GAA.GAA.TCA.G
AC.ATG.AAT.TTA.AGA.CCT.AAT.TTA.GAT.AGT.TTT.GTT.ATT.TTT.CCA.TGG.AGA.CCA.CAA.CAA
.GGC.AAA.GTT.GCA.AGA.TTA.ATT.TGT.GAT.GTA.TAT.AAG.CCA.GAT.GGT.ACA.CCA.TTT.GAA.G
GA.GAT.CCA.AGA.CAT.GTC.CTA.AAG.AGA.GCA.AAT.GCA.GAT.GCA.AAA.GAA.TTA.GGA.TAT.ACA
.ATG.AAT.GTT.GGA.CCA.GAA.TGT.GAA.TTC.TTC.TTA.TTT.GAA.ACA.GAT.GAA.AAT.GGC.AGA.G
CA.ACA.ACT.AAT.ACT.CAA.GAT.AAA.GCT.GGA.TAT.TTT.GAT.TTA.GCA.CCT.ACA.GAT.TTA.GGA
.GAA.AAT.GCT.AGA.AGA.GAC.ATG.ACT.TTA.GCT.TTA.GAA.GAA.ATG.GGA.TTC.GAA.ATT.GAA.G
CA.TCT.CAT.CAT.GAA.GTT.GCT.GAG.GGT.CAA.AAT.GAA.ATT.GAC.TTT.AAG.TAT.GGA.GAT.GCA
.TTA.ACT.ACA.GCT.GAT.AAT.ATT.ATG.ACA.TTT.AAA.CTA.GTT.GTT.AAA.TCT.ATT.GCA.CAA.A
GA.CAT.GGA.TTA.CAT.GCA.TCG.TTT.ATG.CCA.AAA.CCA.ATA.TTC.GGA.ATT.AAT.GGT.TCT.GGA
.ATG.CAT.GTA.AAT.ATG.TCA.TTA.TTC.AAA.GAT.GGA.AAA.AAT.GCT.TTT.GTT.GAT.GAA.AAT.G
AT.AAA.AAT.GGA.TTA.AGT.AAA.GTT.GCT.TAT.CAA.TTT.ATC.GCT.GGA.TTA.TTA.AAG.AAC.ATT
.AAA.GGA.ATG.GCT.GCT.GTA.ACT.AAT.CCA.TTA.GTT.AAC.TCA.TAT.AAG.AGA.TTA.GTA.CCA.G
GA.TAT.GAA.GCA.CCA.GTA.TAT.CTA.GCT.TGG.TCT.TGC.AAA.AAC.AGA.ACA.GCA.TTA.ATA.AGA
.GTA.CCA.GCA.GCT.AGA.GGT.GCA.GGG.ACT.AGA.GTA.GAA.TTA.AGA.TGT.CCA.GAT.CCA.AGT.T
CT.AAT.CCA.TAT.TTA.GTA.TTA.GCA.TGT.TTA.TTA.CAA.GCA.GGA.TTA.GAT.GGA.ATT.AAA.AAT
.AAT.TTA.CAA.CCA.CCT.GCT.GAA.GTT.GAA.GCA.AAC.ATT.TTT.GCA.ATG.ACT.GAG.CAA.GAA.A
GA.AAA.GAA.AAT.GGA.ATT.GAT.AAT.CTG.CCA.AAC.AAT.TTA.TAT.GAA.GCA.GTA.AAT.TAC.ATG
.AAA.GAA.AAT.GAA.TTA.GCT.AAA.AAA.GCT.TTA.GGA.GAT.CAT.GTA.TAT.GGA.AAT.TAT.GTT.G
CA.GGA.AAA.GCA.GCA.GAA.TGG.GAT.GAT.TAC.AGA.ACA.AAA.GTA.CAT.GAT.TGG.GAA.TTA.GAG
.AAT.TAT.CTA.AAT.AAA.TAC.TAA.GGCACAATCAAAAAAATAATGGACCAGTATGCTTGTCTGTTTTGCTTCA
TACTGCGTCGGAACTTTCAGTTAATTGCTTGCTATGAGCTAAAAGTTCCTACTTCCATGACACAAAATATACGTCGCA
TCTTTGGTCTATTATTTTCTTTCATGTGCCTAATCTAGTAATGTAATTACTTTTAGTTAAGGCGGGGATTGCAATGGT
CCAAAGTGAAAAAGGAAAACTTATTATAGCATTAAGCAATGTTGAGATAGCAAAAAAATTAAAAACTTCGTAACGCAA
GAAGGGTTTGATATCATAGCCCTATGTACTTCAGGAAATGAATTAATTAGATTAGTTATGCAATATTCCCCAGACCTG
GTTTTAGTAGGATATAAGTTTAAGGACATGAGTTTACTGGATGTGTATGAAAACTTAGTAGATTTAACTAGCTTTTTG
GCTATTGTAAATGAGCCATATAGATCCTTTATAGAAGAAGATACAGATATATATTTGTATTGGTACTAAAATTTCAAC
TGTACTTTTTAACTAATGCCATTGATTTGATATTTCAATGTAAAAGAAGAATCAAAAANTTAAANNAGCAGGTATAAA
ANTTTAGGAACCATTACTCTAGA
```

Fig. D.1.: Complete nucleotide sequence of the *BglII-XbaI* fragment, originated from pHZ200. The *BglII*-site is lost if this fragment is subcloned in the *BamHI/XbaI* sites of vector pLK34. The SD and AUG elements of the *C. acetobutylicum glnA* are underlined.

Appendix F

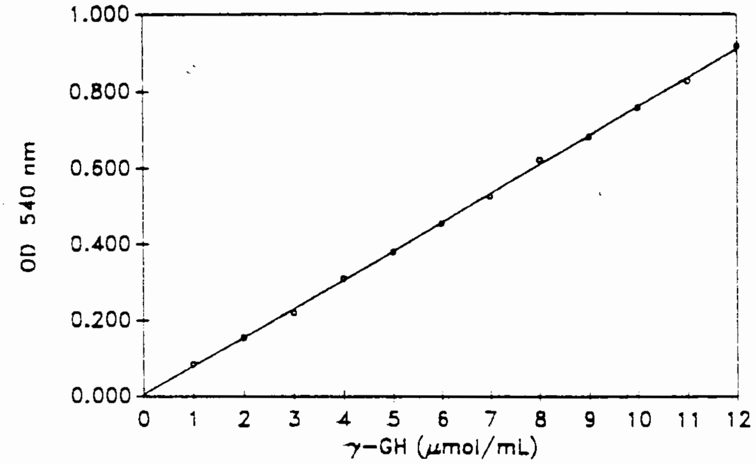
One- and three-letter codes used for amino acids

Amino acid	Codes		Code	Amino acid
Alanine	Ala	A	A	Alanine
Arginine	Arg	R	C	Cysteine
Asparagine	Asn	N	D	Aspartic acid
Aspartic acid	Asp	D	E	Glutamic acid
Cysteine	Cys	C	F	Phenylalanine
Glutamine	Gln	Q	G	Glycine
Glutamic acid	Glu	E	H	Histidine
Glycine	Gly	G	I	Isoleucine
Histidine	His	H	K	Lysine
Isoleucine	Ile	I	L	Leucine
Leucine	Leu	L	M	Methionine
Lysine	Lys	K	N	Asparigine
Methione	Met	M	P	Proline
Phenylalanine	Phe	F	Q	Glutamine
Proline	Pro	P	R	Argenine
Serine	Ser	S	S	Serine
Threonine	Thr	T	T	Threonine
Tryptophan	Trp	W	V	Valine
Tyrosine	Tyr	Y	W	Tryptophan
Valine	Val	V	Y	Tyrosine

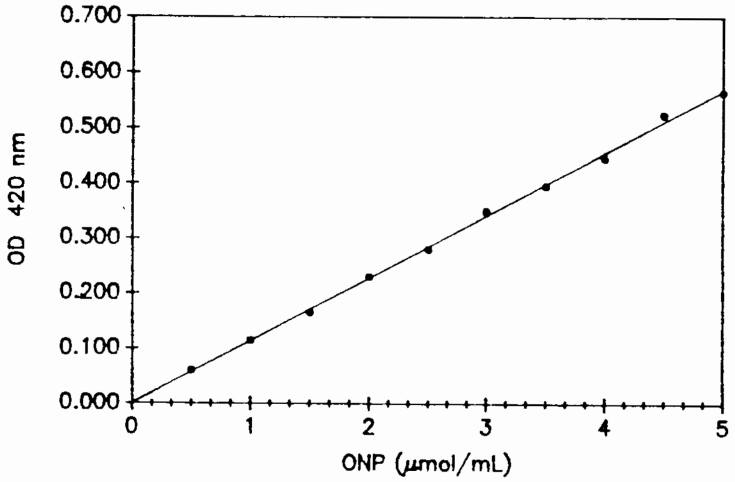
Appendix G

Standard curves used for the calculation of α-galactosidase and γ-glutamyl transferase activities and for the determination of protein concentrations.

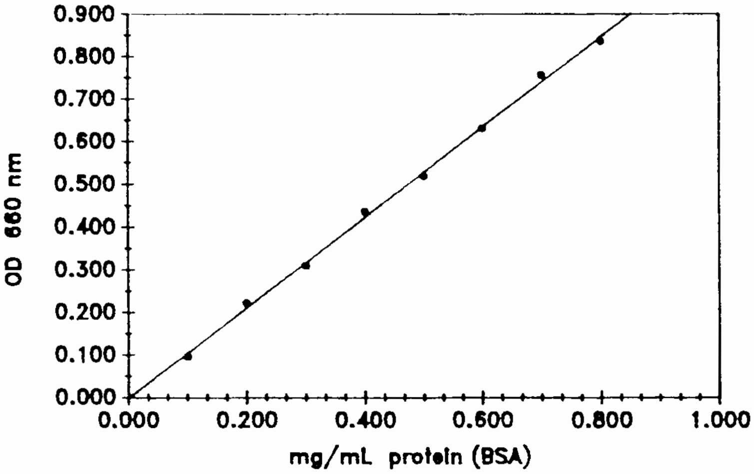
Standard curve for γ-GT assay



Standard curve for β-gal assay



Standard curve for protein determination



Appendix H

Allignment of the amino acid sequences of 11 bacterial GS enzymes

```
Ec     MSAEH-VLTM-LN-EHEVKFVDLRFTDTKGKEQHVTIPAHQVNAEFFEEGKMFDGSSIGGWKGINESDMVLNPDASTAVID
St     MSAEH-VLTM-LN-EHEVKFVDLRFTDTKGKEQHVTIPAHQVNAEFFEEGKMFDGSSIGGWKGINESDMVLMPDASTAVID
Va     MSVEN-VLSL-IQ-ENEVKFVDLRFTDTKGKEQHISIPAHQIDADFFEEGKMFDGSSVAGWKGINESDMVMMPDASSAVLD
Rl     MATASEILKQ-IK-ENDVKFVDLRFTDRRASLQHVTMDVVCVDEDMFADGVMFDGSSIGGWKAINESDMVLMPDTETVHMD
Tf     MGYSPSDVVKLIQ-EKDIKFIDFRFTDTKGKEQHVSFPGMSSKKTPLPKAS-FDGSSIAGWKGINESDMILLPDPDSAVLD
Ab     MSDISKVFDL-IK-EHDVKYVDLRFTDPRGKLHHTAQHVSTIDEDVFEDGIMFDGSSIAGWKAINESDMILQLDPTTAVMD
An     MTTPQEVLKR-IQDEK-IELIDLKFIDTVGTWQHLTLYQNQIDESSFSDGVPFDGSSIRGWKAINESDMTMVLDPNTAWID
Sc     MFQNADDVKKFIADE-DVKFVDVRFCDLPGVMQHFTLPATAFDPDAEQ-A--FDGSSIRGFQAIHESDMSLRPDLSTARVD
Bs     MAKYTREDIEKLVLEEDVKYIRLQFTDILGTIKNVEIPVSQLGKALDNK-VMFDGSSIEGFVRIEESDMYLYPDLNTFVIF
Bc     MARYTKEDIFRLAKEENVKYIRLQFTDLLGVIKNVEIPVSQLTKALDNK-MMFDGSSIEGFVRIEESDMYLYPDLDTWVIF
Ca     MAKYTKEDIINLVKENGVKFIRLQFTDIFGTLKNVAITDKQLEKALDNEC-MFDGSSIDGFVRIEESDMNLRPNLDSFVIF


Ec     PFFADSTLIIR---CDILEPGTLQGYDRDPRSIAKRAEDYLRSTGIADTVLFGPEPEFFLFDDIRFGSSISGSHVAIDDIE
St     PFFADSTLIIR---CDILEPGTLQGYDRDPRSIAKRAEDYLRATGIADTVLFGPEPEFFLFDDIRFGASISGSHVAIDDIE
Va     PFTEDATLNIR---CDILEPATMQGYDRDPRSIAKRAEDFMRSTGVADTVLIGPEPEFFLFDDVKFATDMSGSFFKIDDVE
Rl     PFFAQSTMVIV---CDILDPVSGEAYNRDPRGTAKKAEAYLKASGIGDTVFVGPEAEFFVFDDVKYKADPYNTGFKLDSTE
Tf     PFMDETTLLLR---CDVIEPAYGQGYERDPRSVAKRAGIYLKSTGIADTSFFGP--EFFVFDSVTWNIDMSGCAYLVDAEE
Ab     PFSAQPTLNIL---CDVYEPSTGQPYARCPRGIALAAEKYMASAGIADTAYFGPEAEFFVFDDVKFKVEMNKVSYEFDSEE
An     PFMEVPTLSIV---CSIKEPRTGEWYNRCPRVIAQKAIDYLVSTGIGDTAFFGPEAEFFIFDSARFAQNANEGYYFLDSVE
Sc     PFRRDKTLNINFFIHDPI---TGEQYSRDPRNVAKKAEAYLASTGIADTAFFGPEAEFYVFDSVRFATRENESFYHIDSEA
Bs     PWTAEKGKVAR-FICDIYNPD-GTPFEGDPRNNLKRILKEMEDLGFSDFN-LGPEPEFFLFKLDEKGEPTLELN---D---
Bc     PWTAEKGKVAR-LICDIYNAD-GTPFEGDPRNNLKRVLKEMEALGFSDFN-LGPEPEFFLFKVDEKGNPTLELN---D---
Ca     PWRPQQGKVAR-LICDVYKPD-GTPFEGDPRHVLKRANADAKELGYTMNV--GPECEFFLFETDENGRATTNTQ---D---


Ec     GAWNSSTQYEGGNKGHRPAV--KGGYFPVPPVDSAQDIRSEMCLVMEQMGLVVEAHHHEVATAGQLEVATRFLIMTKKADE
St     GAWNSSTKYEGGNKGHRPGV--KGGYFPVPPVDSAQDIRSEMCLVMEQMGLVVEAHHHEVATAGQNEVATRFLTMTKKADE
Va     AAWNTGSDYEEGNKGHRPGV--KGGYFPVAPVDSSQDIRSAMCLVMEEMGLVVEAHHHE-ATAGQNEIATRFNTLTTKADE
Rl     LPSNDDTDYETGNLGHRPRV--KGGYFPVPPVDSAQDMRSEMLTVLSEMGVVVEKHHHEVAAA-QHELGIKFDTLVRNADK
Tf     AAWNSGKEYEDGNMGHRLGV--KGGYFPVPPVDSAQDLRSAMCLAMEEMGLKVEVHHHEVATAGQHEIGVGFDTLTPRRMR
Ab     GPYTSDKDYEDGNLGHRPGV--KGGYFPVAPVDSGSDLRAEMLSVLAEMGVPVEKHHHEVA-ASQHELGIKFDTLVRTGDN
An     GAWNSGKEGTADKPNLAYKPRFKEGYFPVSPTDSFQDIRTEMLLTMAKLGVPIEKHHHEVATGGQCELGFRFGKLIEAADW
Sc     GAWNTGAL-ED-NRE--YKVRYKGGYFPVPPVDHFADLRAEISLELERSGLQVERQHHEVGTAGQAEINYKFNTLLAAADD
Bs     ----------------------KGGYFDLAPTDLGENCRRDIVLELEEMGFEIEASHHEVAP-GQHEIDFKYAGAVRSCDD
Bc     ----------------------NGGYFDLAPMDLGENCRRDIVLELEEMGFEIEASHHEVAP-GQHEIDFKYANAIRSCDD
Ca     ----------------------KAGYFDLAPTDLGENARRDMTLALEEMGFEIEASHHEVAE-GQNEIDFKYGDALTTADN


Ec     IQIYKYVVHNVVRNRFGKTATFMPKPMFGDNGSGMHCHMSLSKNGVNLFAGDK--YAGLSEQALYYIGGVI-KQPKAINAL
St     IQIYKYVVHNVAH-RFGKTATFMPKPMFGDNGSGMHCHMSLAKNGTNLFSGDK--YAGLSEQALYYIGGVI-KHAKAINAL
Va     IQIYKYVVHNVAH-AFGKTATFMPKPLVGDNGSGMHVHQSLAKDGVNLFAGDK--YGGLSEMALYYIGGII-KHARAINAF
Rl     MQIYKYVVHQVA-NAYGKTATFMPKPIFGDNGSGMHVHQSIWKGGKPTFAGDE--YAGLSESCLFYIGGII-KHAKAINAF
Tf     CKILKYVVHNVA--AVRQTATFMPKPVVGDNGSGMHVHQSLGKDGKNIFAGDL--YGGLSEIALYYIGGII-KHAKAVNAL
Ab     MQYYKYVVHNVAH-AYGKTATFMPKPVFGDNGSGMHMHQSIWKEGQPLFAGNQ--YADLSELALYYIGGII-KHAKALNAF
An     LMIYKYVIKNVAK-KYGKTVTFMPKPIFGDNGSGMHCHQSIWKDGKPLFAGDQ--YAGLSEMGLYYIGGLI-KHAPALLAI
Sc     LQLFKYIVKNVAW-KNGKTATFMPKPIFGDNGSGMHVHQSLWSGGEPLFY-DEQGYAGLSDTARYYIGGGILKHAPSLLAF
Bs     IQTFKLVVKTIAR-KHGLHATFMPKPLFGVNGSGMHCNLSLFKNGVNAFF-VENAALQLSETAKHFIAGIV-KHATSFTAV
Bc     IQTFKLVVKTIAR-KHGLHATFMPKPLYGVNGSGMHCNLSLFKNGENVFY-DQNGDLQLSDDARHFIAGI-LKHAPAFTAV
Ca     IMTTKLVVKSIAQ-RHGLHASFMPKPIFGINGSGMHVNMSLFKDGKNAFV-DENDKNGLSKVAYQFIAGL-LKNIKGMAAV
```

```
Ec      ANPTTNSYKRLVPGYEAPVMLAYSARNRSASIRIPVV-S-SPKARRIEVRFPDPAANPYLCFAALLMAGLDGIKNKIH-PG
St      ANPTTNSYKRLVPGYEAPVMLAYSARNRSASIRIPVV-A-SPKARRIEVRFPDPAANPYLCFAALLMAGLDGIKNKIH-PG
Va      ANPSTNSYKRLVPGFEAPVMLAYSARNRSASIRIPVV-P-SPKARRIEVRFGDPAANPYLCFASMLMAGLDGIKNKIH-PG
Rl      TNPSTNSYKRLVPGYEAPVLLAYSARNRSASCRIPF--GSNPKAKRVEVRFPDPTANPYLAFAAMLMAGLDGIKNKIH-PG
Tf      TNPSTNSYKRLVPHFEAPVLLAYSAKNRSASIRIPYV-N-SPKARRIEVRFPDSTANPYLAFSAMLMAGLDGIQNKIH-PA
Ab      TNPTTNSYKRLVPGYEAPVLLAYSARNRSASCRIPYV-A-SPKGKRVEVRFPDPSANPYLAFAALLMAGLDGIQNKIH-PG
An      TNPSTNSYKRLVPGYEAPVNLAYSQGNRSASIRIPLS-GTNPKAKRLEFRCPDATSNPYLAFAAMLCAGIDGIKNKIH-PG
Sc      TNPTVNSYHRLVPGFEAPVNLVYSQRNRSAAMRIPIT-GSNPKAKRVEFRAPDASGNPYLAFSALLLAGLDGIKNKIE-PA
Bs      TNPTVNSYKRLVPGYEAPCYVAWSAQNRSPLIRIPASRGIST---RVEVRSVDPAANPYLALSVLLAAGLDGIKNKLEAPA
Bc      ANPTVNSYKRLVPGYEAPCYVAWSAQNRSPLVRIPASRGIST---RVEVR-VDPAANPYLVMATLLAAGLDGIKNKLTPPA
Ca      TNPLVNSYKRLVPGYEAPVYLAWSCKNRTALIRVPAARGAGT---RVELRCPDPSSNPYLVLACLLQAGLDGIKNNLQPPA


Ec      EAMDKNLYDLPPEEAKEIPQVAGSLEEALNELDLDREFLKAGGVFTDEAIDA-YALRREE--DDRVRMTPHPVEFELYYSV
St      EPMDKNLYDLPPEEAKEIPQVAGSLEEALNALDLDREFLKAGGVFTDEAIDA-YIALRPE-EDDRVRMTPHPVEFELYYSV
Va      EAMDKDLYDLPAEESAEIPTVAYSLKDALAELDADREFLTAGGVFSDDFID-SYIELKSQ-DVERVNMTTHPVEFELYYSV
Rl      KAMDKDLYDLPPKELKKIPTVCGSLRQALESLDKDRKFLTAGGVFDDDQIDA-FIELKMA-EVMRFEMTPHPVEYDMYYSA
Tf      TAMDKNLYDLPAEEQANIPGVAASLEEALRALEADHDFLMKGGVFSESWLQG-YLDVKWA-EVQTLRVTTHPVEFQ-YYSL
Ab      EAMDKNLYDLPAEELAKVPTVCGSR-EALDSLKADSAFLQKGDVFTKDMI-ESYIDLRTE-ELLAFEM-PHPIEYKMYYSV
An      EPLDKNIYELSPEELAKVPSTPGSLELALEALENDHAFLTDTGVFTEDFIQN-WIDYKLANEVKQMQLRPHPYEFSIYYDV
Sc      EPIDKDLYELAPEEHANVAQVPTSLGAVLDRLEADHEFLLQGDVFTPDLI-ETWIDFKRANEIAPLQLRPHPHEFEMYFDV
Bs      -PIDRNIYVMSKEERMENGIVDLPATLA-EALEEFKSNEVMV-KALGEHLFEHFIEAKEI-EWDMFRTQVHPWEREQYHSQY
Bc      -AVDRNIYVMTKEEREEAGIVDLPATLAQ-ALVTLQSNEVISN-ALGDHLLEHFIEAKEF-EWDIFRTQVHQWERDQYMSLY
Ca      E-VEANIFAMTEQERKENGIDNLPNNLY-EA-VNYMKENELAKKALGDHVYGNYVAGKAA-EWDDYRTKVHDWELENYLNKY
```

Fig. H.1.: Optimal allignment of the aa sequences of 11 bacterial GS enzymes. References are as given in Table 1.1, abbrevations are as in Fig. 2.5.

LITERATURE CITED

Adachi, K., and I. Suzuki, 1977. Purification and properties of glutamate synthase from *Thiobacillus thioparus*. J. Bacteriol. **129:**1173-1182.

Adler, S. P., D. Purich, and E. Stadtman. 1975. Cascade control of *Escherichia coli* glutamine synthetase: properties of the P_{II} regulatory protein and the uridylyltransferase-uridylyl-removing enzyme. J. Biol. Chem. **250:**6264-6272.

Aiba, H., S. Adhya, and B. de Crombrugghe. 1981. Evidence for two functional *gal* promoters in intact *Escherichia coli* cells. J. Biol. Chem. **256:** 11905-11910.

Aiba, H., S.-I. Matsuyama, T. Mizuno, and S. Mizushima. 1987. Function of *micF* as an antisense RNA in osmoregulatory expression of the *ompF* gene in *Escherichia coli*. J. Bact. **169:**3007-3012.

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

Allf-Steinberger C. 1984. Evidence for a coding pattern on the non-coding strand of the *E. coli* genome. Nucl. Ac. Res. **12:**2235-2241.

Almassy, R. J., C. A. Janson, R. Hamlin, N.-H. Xuong, and D. Eisenberg. 1986. Novel subunit-subunit interactions in the structure of glutamine synthetase. Nature (London) **323:**304-309.

Alvarez-Morales, A., R. Dixon, and M. Merrick. 1984. Positive and negative control of the *glnA ntrBC* regulon in *Klebsiella pneumoniae*. EMBO J. **3:**501-507.

Andersch, W., H. Bahl, and G. Gottschalk. 1982. Acetone -butanol production by *Clostridium acetobutylicum* in an ammonia-limited chemostat at low pH values. Biotechnol. Lett. **4:**29-32.

Andersch, W., H. Bahl, and G. Gottschalk. 1983. Level of enzymes involved in acetate, butyrate and butanol formation by *Clostridium acetobutylicum*. Eur. J. Appl. Microbiol. Biotechnol. **18:**327-332.

Andersen, J., N. Delihas, K. Ikenaka, P. Green, O. Pines, O. Ilercil, and M. Inouye. 1987. The isolation and characterization of RNA coded by the *micF* gene in *Escherichia coli*. Nucl. Ac. Res. **15:**20892101.

Andersen, J., S. A. Forst, K. Zhao, M. Inouye, and N. Delihas. 1989. The function of *micF* RNA. J. Biol. Chem. **264:**17961-17970.

Auble, D. T., and P. L. deHaseth. 1988. Promoter recognition by *Escherichia coli* RNA polymerase. J. Mol. Biol. **202**:471-482

Ausubel, F. M. 1984. Regulation of nitrogen fixation genes. Cell **37**:5-6.

Ayers, D. G., D. T. Auble, and P. L. deHaseth. 1989. Promoter recognition by *Escherichia coli* RNA polymerase. J. Mol. Biol. **207**:749-756.

Backman, K. C., Y.-M. Chen, S. Ueno-Nishio, and B. Magasanik. 1983. The product of *glnL* is not essential for regulation of bacterial nitrogen assimilation. J. Bacteriol. **154**:516-519.

Backman, K., Y.-M. Chen, and B. Magasanik. 1981. Physical and genetic characterization of the *glnA-glnG* region of the *Escherichia coli* chromosome. Proc. Natl. Acad. Sci. USA**78**:3743-3747.

Bahl, H., W. Andersch, and G. Gottschalk. 1982b. Continuous production of acetone and butanol by *Clostridium acetobutylicum* in a two-stage phosphate limited chemostat. Eur. J. Appl. Microbiol. Biotechnol. **15**:201-205.

Bahl, H., W. Andersch, K. Braun, and G. Gottschalk. 1982a. Effect of pH and butyrate concentration on the production of acetone and butanol by *Clostridium acetobutylicum* grown in continuous culture. Eur. J. Appl. Microbiol. Biotechnol. **14**:17-20.

Barber, J. M., F. T. Robb, J. R. Webster, D. R. Woods. 1979. Bacteriocin production by *Clostridium acetobutylicum* in an industrial fermentation proces. Appl. Environ. Microbiol. **37**:433-437.

Barros, M. E. C., D. E. Rawlings, and D. R. Woods. 1985. Cloning and expression of the *Thiobacillus ferrooxidans* glutamine synthetase gene in *Escherichia coli*. J. Bacteriol. **164**:1386-1389.

Barros, M. E., D. E. Rawlings, and D. R. Woods. 1986. Purification and regulation of a cloned *Thiobacillus ferooxidans* glutamine synthetase. J. Gen. Microbiol. **132**:1989-1995.

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

Belasco, J. G., and C. F. Higgins. 1988. Mechanisms of mRNA decay in bacteria: a perspective. Gene **72**:15-23.

Belin, D., E. A. Mudd, P. Prentki, Y. Yi-Yi, and H. M. Krisch. 1987. Sense and antisense transcription of bacteriophage T4 gene 32. J. Mol. Biol. **194:**231-243.

Bender, R. A., K. A. Janssen, A. D. Resnick, M. Blumenberg, F. Forrest, and B. Magasanik. 1977. Biochemical parameters of glutamine synthetase from *Klebsiella aerogenes*. J. Bacteriol. 129:1001-1009

Bender, R. A., P. M. Snyder, R. Bueno, M. Quinto, and B. Magasanik. 1983. Nitrogen regulation system of *Klebsiella aerogenes*: the *nac* gene. J. Bacteriol. **156:**444-446.

Bertram, J., and P. Dürre. 1989. Conjugal transfer and expression of streptococcal transposons in *Clostridium acetobutylicum*. Arch. Microbiol. **151:**551-557.

Beynon, J., M. Cannon, V. Buchanan-Wollaston, and F. Cannon. 1983. The *nif* gene promoters of *Klebsiella* have a characteristic primary structure. Cell **34:**665-671.

Beyon, W.-H., and B. Weisblum. 1984. Post-transcriptional regulation of chloramphenicol acetyl transferase. J. Bacteriol. **158:**543-550.

Bhatnagar, L., J. G. Zeikus, and J.-P. Aubert. 1986. Purification and characterization of glutamine synthetase from the archaebacterium *Methanobacterium ivanovi*. J. Bacteriol. **165:**638-643.

Birkmann, A., R. G. Sawers, and A. Böck. 1987. Involvement of the *ntrA* gene product in the anaerobic metabolism of *Escherichia coli*. Mol. Gen. Genet. **210:**535-542.

Bodasing, S. J., P. W. Brandt, F. T. Robb, and D. R. Woods. 1985. Purification and regulation of glutamine synthetase in a collagenolytic *Vibrio alginolyticus* strain. Arch. Microbiol. **140:**369-374.

Bohannon, D. E. , and A. L. Sonenshein. 1989. Positive regulation of glutamate biosynthesis in *Bacillus subtilis*. J. Bacteriol. **171:**4718-4727.

Bohannon, D. E., M. S. Rosenkrantz, and A. L. Sonenshein. 1985. Regulation of *Bacillus subtilis* glutamate synthase genes by the nitrogen source. J. Bacteriol. **163:**957-964.

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

Botstein, D. and D. Shortle. 1985. Strategies and applications of in vitro mutagenesis. Science 229:1193-1209.

Bott, K. F., G. Reysset, J. Gregoire, D. Islert, and J.-P. Aubert. 1977. Characterization of glutamine requiring mutants of *Bacillus subtilis*. Biochem. & Biophys. Res. Comm. **79**:996-1003.

Botterman, J., and M. Zabeau. 1987. Versatile plasmid vectors for manipulation and enhanced expression of genes in *Escherichia coli*. DNA **6**:583-591.

Bourret, R. B., J. F. Hess, K. A. Borkovich, A. A. Pakula, and M. I. Simon. 1989. Protein phosphorylation in chemotaxis and two-component regulatory systems of bacteria. J. Biol. Chem. **13**:7085-7089.

Bowles, L. K., and W. L. Ellefson. 1985. Effects of butanol on *Clostridium acetobutylicum*. Appl. Environ. Microbiol. **41**:459-1170.

Bozouklian, H., and C. Elmerich. 1986. Nucleotide sequence of the *Azospririllum brasilense* Sp 7 glutamine synthetase gene. Biochimie **68**:1181-1187.

Brana, A. F., N. Paiva, and A. L. Demain. 1986. Pathways and regulation of ammonium assimilation in *Streptomyces clavuligerus*. J. Gen. Microbiol. **132**:1305-1317.

Bravo, A., and J. Mora. 1988. Ammonium assimilation in *Rhizobium phaseoli* by the glutamine synthetase-glutamate synthase pathway. J. Bacteriol. **170**:980-984.

Briat, J.-F., and M. J. Chamberlain. 1984. Identification and characterization of a new transcriptional terminator factor from *Escherichia coli*. Proc. Natl. Acad. Sc. **81**:7373-7377.

Brunner, M., and H. Bujard. 1987. Promoter recognition and promoter strength in the *Escherichia coli* system. EMBO J. **6**:3139-3144.

Buck, M., W. Cannon, and J. Woodcock. 1987. Transcriptional activation of the *Klebsiella pneumoniae* nitrogenase promoter may involve DNA loop formation. Molecul. Microbiol. **1**:243-249.

Bueno, R., G. Pahel, and B. Magasanik. 1985. Role of *glnB* and *glnD* gene products in regulation of the *glnALG* operon of *Escherichia coli*. J. Bacteriol. **164**:816-822.

Buikema, W. J., W. W. Szeto, P. V. Lemley, W. H. Orme Johnson, and F. M. Ausubel. 1985. Nitrogen fixation specific regulatory genes of *Klebsiella pneumoniae* and *Rhizobium meliloti* share homology with the general nitrogen regulatory gene *ntrC* of *K. pneumoniae*. Nucl. Ac. Res. **13**:4539-4555.

Caballero, F. J., F. J. Cejudo, F. J. Florencio, J. Cardenas, and F. Castillo. 1985. Molecular and regulatory properties of glutamine synthetase from the phototropic bacterium *Rhodopseudomonas capsulata* E1F1. J. Bacteriol. **162**:804-809.

Carlson, T. A., and B. K. Chelm. 1986. Apparent eukaryotic origin of glutamine synthetase II from the bacterium *Bradyrhizobium japonicum*. Nature (London) **322**:568-570.

Carlson, T. A., G. B. Martin, and B. C. Chelm. 1987. Differential transcription of the two glutamine synthetase genes of *Bradyrhizobium japonicum*. J. Bacteriol. **169**:5861-5866.

Carlson, T. A., M.-L. Guerinot, and B. K. Chelm. 1985. Characterization of the gene encoding glutamine synthetase I (*glnA*) from *Bradyrhizobium japonicum*. J. Bacteriol. **162**:698-703.

Carlson,, T. A., M. L. Guerinot, and B. K. Chelm. 1985. Characterization of the gene encoding glutamine synthetase I (*glnA*) from *Brayrhizobium japonicum*. J. Bacteriol. **162**:698-703.

Carter, P., H. Bedouelle, and G. Winter. 1985. Improved oligonucleotide site-directed mutagenesis using M13 vectors. Nucl. Ac. Res. **13**:4431-4442.

Carter-Muenschau, P., and R. E. Wolf, Jr. 1989. Growth-rate dependent regulation of 6-phosphogluconate dehydrogenase level mediated by an anti-Shine-Dalgarno sequence located within the *Escherichia coli gnd* structural gene. Proc. Natl. Acad. Sc. USA **86**:1138-1142

Cary, J. W., D. J. Petersen, E. T. Papoutsakis, and G. N. Bennett. 1988. Cloning and expression of *Clostridium acetobutylicum* phosphotransbutyrylase and butyrate kinase genes in *Escherichia coli*. J. Bacteriol. **170**:4613-4618.

Casadaban, M. J., J. Chou, and S. N. Cohen. 1980. *In vitro* gene fusions that join an enzymatically active β-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. J. Bacteriol. **143**:971-980.

Cashel, M., and K. E. Rudd. 1987. The stringent response, p. 1410-1438. *In* F. C. Neidhardt (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D. C.

Cassan, M., C. Parsot, G. N. Cohen, and J.-C. Patte. 1986. Nucleotide sequence of the *lysC* gene encoding the lysine sensitive aspartokinase II of *Escherichia coli* K12. J. Biol. Chem. **261:**1052-1057.

Castaño, I., and F. Bastarrachea. 1984. *glnF-lacZ* fusions in *Escherichia coli*: studies on *glnF* expression and its chromosomal orientation. Mol. Gen. Genet. **195:**228-233.

Castaño, I., F. Bastarrachea, and A. A. Covarrubias. 1988. The *gltBDF* operon of *Escherichia coli*. J. Bacteriol. **170:**821-827.

Chang, S., and S. N. Cohen. 1979. High frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA. Mol. Gen. Genet. **168:**111-115.

Chang, A. Y. C., and S. N. Cohen. 1978. Construction and characterization of amplifiable multcopy DNA cloning vehicles derived from the p15A cryptic miniplasmid. J. Bact. **134:**1141-1156.

Chen, C.-Y., J. T. Beatty, S. N. Cohen, and J. G. Belasco. 1988. An intercistronic stem-loop structure functions as an mRNA decay terminator necessary but insufficient for *puf* mRNA stability. Cell **52:**609-619.

Chen, G., and J. B. Russell. 1989. Transport of glutamine by *Streptococcus bovis* and conversion of glutamine to pyroglutamic acid and ammonia. J. Bact. **171:**2981-2985.

Chen, N.-Y., F.-M. Hu, and H. Paulus. 1987. Nucleotide sequence of the overlapping genes for the subunits of *Bacillus subtilis* aspartokinase II and their control regions. J. Biol. Chem. **262:**8787-8798.

Chen, Y.-M., K. Backman, and B. Magasanik. 1982. Characterization of a gene, *glnL*, the product of which is involved in the regulation of nitrogen utilization in *Escherichia coli*. J. Bacteriol. **150:**214-220.

Chopra, L. 1986. Genetic and biochemical basis of tetracycline resistance. J. Antimicrob. Chemother. **18:**51-56.

Clewell, D.B. 1972. Nature of ColE plasmid replication in *Escherichia coli* in the presence of chloramphenicol. J. Bacteriol. **110:**667-676.

Colanduoni, J., R. Nissan., and J. J. Villafranca. 1987. Studies on the mechanism of glutamine synthetase utilizing pH-dependent behavior in catalysis and binding. J. Biol. Chem. **262**:3037-3043.

Colombo, G., and J. J. Villafranca. 1986. Amino acid sequence of *Escherichia coli* glutamine synthetase deduced from the DNA nucleotide sequence. J. Biol. Chem. **261**:10587-10591.

Colonna-Romano, S., A. Riccio, M. Guida, R. Defez, A. Lamberti, M. Iaccarino, W. Arnold, U. Priefer and A. Pühler. 1987. Tight linkage of *glnA* and a putative regulatory gene in *Rhizobium leguminosarum*. Nucl. Ac. Res. **15**:1951-1964.

Corfield, V. A., S. J. Reid, J. Bodmer, and J. Thomson. 1984. A modified protoplast regeneration protocol facilitating the detection of cloned exoenzyme gene in *Bacillus subtilis*. Gene **30**:17-22.

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

Crouse, G. F., A. Frischauf, and H. Lehrach. 1983. An integrated and simplified approach to cloning into plasmids and single-stranded phages. Meth. Enzymol. **101**:78-98.

Cummings, C. S., and J. L. Johnson. Taxonomy of the Clostridia: wall composition, DNA homologies in *Clostridium butyricum* and other butyric acid-producing Clostridia. J. Gen. Microbiol. 67:33-46.

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

Darrow, R. A., and R. R. Knotts. 1977. Two forms of glutamine synthetase in free-living root nodule bacteria. Biochem. Biophys. Res. Comm. **78**:554-559.

Dautry-Varsat, A., G. N. Cohen, and E. Stadtman. 1979. Some properties of *Escherichia coli* glutamine synthetase after limited proteolysis by subtilisin. J. Biol. Chem. **254**:3124-3128.

De Block, M., J. Botterman, M. Vandewiele, J. Dockx, C. Thoen, V. Gosselé, N. Rao Movva, and C. Thompson. 1987. Engineering herbicide resistance in plants by expression of a detoxifying enzyme. EMBO J. **6**:2513-2518.

de Bruijn, F. J., and F. M. Ausubel. 1983. The cloning and characterization of the *glnF* (*ntrA*) gene of *Klebsiella pneumoniae*: role of *glnF* (*ntrA*) in the regulation of nitrogen fixation (*nif*) and other nitrogen assimilation genes. Mol. Gen. Genet. **192:**342-353.

de Bruijn, F. J., S. Rossbach, ans J. Schell. 1985. Cloning and characterization of glutamine synthetase genes from *Rhizobium meliloti* 1021, *R. sesbania* ORS571 and *Agrobacterium tumefaciens* C58, p.218. *In* Evans, H. J., P. J. Bottomley, W. E. Newton (Eds.). Nitrogen fixation research progres. Nijhoff, Boston.

Dean, D. R., J. A. Hoch, and A. I. Aronson. 1977. Alteration of the *Bacillus subtilis* glutamine synthetase results in overproduction of the enzyme. J. Bacteriol. **131:**981-987.

Deretic, V., R. Dikshit, W. N. Konyecsni, A. M. Chakrabarty, and T. K. Misra. 1989. The *algR* gene, which regulates mucoidy in *Pseudomonas aeruginosa*, belongs to a class of environmentally responsive genes. J. Bacteriol. **171:**1278-1283.

Deshpande, K. L., and J. F. Kane. 1980. Glutamate synthase from *Bacillus subtilis*: in vitro reconstruction of an active amidotransferase. Biochem. Biophys. Res. Comm. **93:**308-314.

Desphande, K. L., J. R. Katze, and J. F. Kane. 1980. Regulation of glutamate synthase from *Bacillus subtilus* by glutamine. Biochem. & Biophys. Res. Comm. **95:**55-60.

Deuel, T. F., A. Ginsburg, J. Yen, E. Shelton, and E. R. Stadtman. 1970. *Bacillus subtilis* glutamine synthetase: purification and physical characterization. J. Biol. Chem. **245:**5195-5205.

Deuel, T. F., and E. R. Stadtman. 1970. Some kinetic properties of *Bacillus subtilis* glutamine synthetase. J. Biol. Chem. **245:**5206-5213.

Deuel, T. F., and S. Prusiner. 1974. Regulation of glutamine synthetase from *Bacillus subtilis* by divalent cations, feedback inhibitors, and L-glutamine. J. Biol. Chem. **249:**257-264.

Deuschle, U., W. Kammerer, R. Gentz, and H. Bujard. 1986. Promoters of *Escherichia coli*: a hierarchy of *in vivo* strength indicates alternate structures. EMBO J. **5:**2987-2994.

Dingman, D. W., and A. L. Sonenshein. 1987a. Purification of aconitase from *Bacillus subtilis* and correlation of its N-terminal amino acid sequence with the sequence of the *citB* gene. J. Bacteriol. **169:**3062-3067.

Dingman, D. W., M. S. Rosenkrantz, and A. L. Sonenshein. 1987b. Relationship between aconitase gene expression and sporulation in *Bacillus subtilis*. J. Bacteriol. **169:**3068-3075.

Dixon, R. 1986. The *xylABC* promoter from the *Pseudomonas putida* TOL plasmid is activated by nitrogen regulatory genes in *Escherichia coli*. Mol. Gen. Genet. **203:**129-136.

Doi, R. H., and L.-F. Wang. 1986. Multiple procaryotic ribonucleic acid polymerase sigma factors. Microbiol. rev. **50:**227-243.

Donald, R. G., J. Lapointe, and R. A. Ludwig. 1988. Characterization of the *Azorhizobium sesbaniae* ORS571 genomic locus encoding NADPH-glutamate synthase. J. Bacteriol. **170:**1197-1204.

Dreyfus, B. L., C. Elmerich, Y. R. Dommergues. 1983. Free-living *Rhizobium strain* able to grow under N_2 as a sole nitrogen source. Appl. Environ. Microbiol. **45:**711-713.

Dreyfus, M. 1988. What constitutes the signal for the initiation of protein synthesis on *Escherichia coli* mRNAs ? J. Mol. Biol. **204:**79-94.

Drummond, M., P. Whitty, and J. Wootton. 1986. Sequence and domain relationship of NtrC and NifA from *Klebsiella pneumoniae*: Homologies to other regulatory proteins.EMBO J. **5:**441-447.

Ebright, R. H. 1982. Sequence homologies in the DNA of six sites known to bind to the catabolite activator protein of *Escherichia coli*, p. 91-99. *In* J.P Griffin and W. L. Duax (eds.), Molecular Structure and Biological Activity. Elsevier Science Publishing Inc., N.Y.

Ernst, J.F. 1988. Codon usage and gene expression. Trends in Biotechnol. **6:**196-199.

Errington, J. 1984. Efficient *Bacillus subtilis* cloning system using bacteriophage vector ϕ109. J.Gen. Microbiol. **130:**2615-2628.

Espin, G., A. Alvarez-Morales, F. Cannon, R. Dixon, and M. J. Merrick. 1982. Cloning of the *glnA*, *ntrB* and *ntrC* genes of *Klebsiella pneumoniae* and studies of their role in regulation of the nitrogen fixation (*nif*) gene cluster. Mol. Gen. Genet. **186:**518-524.

Ferrari, E., D. Henner, and J. A. Hoch. 1981. Isolation of *Bacillus subtilis* genes from a Charon 4A library. J. Bacteriol. **146:**430-432.

Fersht, A., and J. W. Knill-Jones. 1981. Polymerase accuracy and spontaneous mutation rates: purine-purine, purine-pyrimidine, and pyrimidine-pyrimidine mismatches during DNA replication. Proc. Natl. Acad. Sci. USA **78:**4251-4254.

Fiedler, S., and R. Wirth. 1988. Transformation of bacteria with plasmid DNA by electroporation. Anal. Biochem. **170:**38-44.

Filser, M. M., C. Moscatelli, A. Lamberti, E. Vincze, M. Guida, G. Salzano, and M. Iaccarino. 1986. Characterization and cloning of two *Rhizobium leguminosarum* genes coding for glutamine synthetase activities. J. Gen. Microbiol. **132:**2561-2569.

Fisher, R., R. Tuli, and R. Haselkorn. 1981. A cloned cyanobacterial gene for glutamine synthetase functions in *Escherichia coli,* but the enzyme is not adenylylated. Proc. Natl. Acad. Sc. USA **78:**3393-3397.

Fisher, S. H. 1989a. Glutamate synthesis in *Streptomyces coelicolor.* J. Bacteriol. **171:**2372-2377.

Fisher, S. H., and A. L. Sonenshein. 1977. Glutamine requiring mutants of *Bacillus subtilis.* Biochem. & Biophys. Res. Comm. **79:**987-995.

Fisher, S. H., and A. L. Sonenshein. 1984. *Bacillus subtilis* glutamine synthetase mutants pleiotropically altered in glucose catabolite repression. J. Bacteriol. **157:**612-621.

Fisher, S. H., and B. Magasanik. 1984c. 2-Ketoglutarate and the regulation of aconitase and histidase formation in *Bacillus subtilis.* J. Bacteriol. **158:**379-382.

Fisher, S. H., and L. V. Wray, Jr. 1989b. Regulation of glutamine synthetase in *Streptomyces coelicolor.* J. Bacteriol. **171:** 2378-2383.

Fisher, S. H., M. S. Rosenkrantz, and A. L. Sonenshein. 1984. Glutamine synthetase gene of *Bacillus subtilis.* Gene **32:**427-438.

Fitch W. M., and Margoliash E. 1967. Construction of phylogenetic trees: a method based on mutation distances as estimated from cytochrome C sequences, is of general applicability. Science **155:**279-284.

Forst, S., J. Delgado, and M. Inouye. 1989. Phosphorylation of OmpR by the osmosensor EnvZ modulates expression of the *ompF* and *ompC* genes in *Escherichia coli.* Proc. Natl. Acad. Sc. USA **86:**6052-6056.

Freese, E. 1981. Initiation of Bacterial sporulation. *In* Sporulation and Germination. Levinson, H. S., A. L. Sonenshein, and D. J. Tipper (eds). American Society of Microbiology, Washington D. C., pp. 1-12.

Freese, E., J. E. Heinze, and E. M. Galliers. 1979. Partial purine deprivation causes sporulation of *Bacillus subtilis* in the presence of excess ammonia, glucose and Phosphate. J. Gen. Microbiol. **115**:193-205.

Friedman, D. I. 1988. Integration host factor: a protein for all reasons. Cell **55**:545-554.

Fry, D. C., S. A. Kuby, and A. S. Mildvan. 1986. ATP-binding site of adenylate kinase: mechanisticimplications of its homology with ras-encoded p21, F1-ATPase, and other nucleotide-binding proteins. Proc. Natl. Acad. Sc. USA **83**:907-911.

Fuchs, R. L., and D. L. Keister. 1980a. Identification of two glutamine synthetases in *Agrobacterium*. J. Bacteriol. **141**:996-998.

Fuchs, R. L., D. L. Keister. 1980b. Comparative properties of glutamine synthetases I and II in *Rhizobium* and *Agrobacterium* spp. J. Bacteriol. **144**:641-648.

Gabriel, C. L. 1928. Butanol fermentation process. Ind. Eng. Chem. **20**:1063-1067.

Gawron-Burke, C., and D. B. Clewell. 1984. Regeneration of insertionally inactivated streptococcal DNA fragments after excision of Transposon Tn*916* in *Escherichia coli*: strategy for targetting and cloning of genes from Gram-positive bacteria. J. Bacteriol. **159**:214-221.

Gebhardt, C., J. E. Oliver, B. G. Forde, R. Saarelainen, and B. Miflin. 1986. Primary structure and differential expression of glutamine synthetase genes in nodules, roots and leaves of *Phaseolus vulgaris*. EMBO J. **5**:1429-1435.

Gibbs, D. F. 1983. The rise and fall (....and rise?) of acetone/butanol fermentations. Trends in Biotechnol. **1**:12-15.

Gold, L. 1988. Post-transcriptional regulatory mechanisms in *E. coli*. Annu. Rev. Biochem. **57**:199-233.

Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. S. Singer, and G. Stormo. 1981. Translational initiation in procaryotes. Annu. Rev. Microbiol. **35**:365-403.

Gottschal, J. C., and J. G. Norris. 1981a. The induction of acetone and butanol production in cultures of *Clostridium acetobutylicum* by elevated concentrations of acetate and butyrate. FEMS Nicrobiol. Lett. **12**:385-389.

Gottschal, J. C., and J. G. Norris. 1981b. Non-production of acetone and butanol by *Clostridium acetobutylicum* during glucose- and ammonium-limitation in continuous culture. Biotechnol. Lett **3**:525-530.

Gough, N. M., A. Webb, S. Cory, and J. M. Adams. 1980. Molecular cloning of seven mouse immunoglobulin K chain messenger ribonucleic acids. Biochem. **19**:2702-2710.

Gouy, N., and Gautier. 1982. Codon usage in bacteria: correlation with gene expressivity. Nucl. Ac. Res. **8**:1893-1912.

Gralla, J. D. 1989. Bacterial gene regulation from distant DNA sites. Cell 57:193-195.

Graves, M. C., **and J. C. Rabinowitz**. 1986. *In vivo* and *in vitro* transcription of the *Clostridium pasteurianum* ferredoxin gene. J. Biol. Chem. **261**:11409-11415.

Greenblatt J., M. **McLimont, and S. Hanley**. 1981. Termination of transcription by *nusA* gene protein of *Escherichia coli.* Nature **292**:215-220.

Grosjean, H., and W. Fiers. 1982. Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. Gene **18**:199-209.

Gross, R., B. Arico, and R. Rappuoli. 1989. Families of bacterial signal-transducing proteins. Molecul. Microbiol. **3**:1661-1667.

Gubler, M., and H. Hennecke. 1988. Regulation of the *fixA* gene and *fixBC* operon in *Bradyrhizobium japonicum. J.* Bacteriol. **170**:1205-1214.

Gussin, G. N., C. W. Ronson, and F. M. Ausubel. 1986. Regulation of nitrogen fixation genes. Annu. Rev. Genet. **20**:567-591.

Hager, P. W., and J. C. Rabinowitz. 1985. Translational specificity in *Bacillus subtilis,* pl-32. In Dubnau (ed.), The Molecular Biology of the Bacilli. Academic Press, Inc., Florida.

Harley, C. B., and R. P. Reynolds. 1986. Analysis of *Escherichia coli* promoter sequences. Nucl. Ac. Res. **15**:2343-2361.

Hawkins, A. R., S. J. Gurr, P. Montague, and J. R. Kinghorn. 1989. Nucleotide sequence and regulation of expression of the *Aspergillus nidulans gdhA* gene encoding NADP dependent glutamate dehydrogenase. Mol. Gen. Genet. **218:**105-111.

Hawley, D. K., and W. R. McClure. 1983. Compilation and analysis of *Escherichia coli* prokoter DNA sequences. Nucl. Ac. Res. **11:**2237-2255.

Hayward, B. E., A. Hussain, R. H. Wilson, A. Lyons, V. Woodcock, B. McIntosh, and T. J. R. Harris. 1986. The cloning and nucleotide sequence of cDNA for an amplified glutamine synthetase gene from the Chinese hamster. Nucl. Ac. Res. **14:**999-1008.

Heinrikson, R. L., and H. S. Kingdon. 1971. Primary structure of *Escherichia coli* glutamine synthetase, II. The complete amino acid sequence of a tryptic heneicosapeptide containing covalently bound adenylic acid. J. Biol. Chem. **246:**1099-1106.

Helmann, J. D., and M. Chamberlin. 1988. Structure and function of bacterial sigma factors. Ann. Rev. Biochem. **57:**839-872.

Hemmila, I. A., and P. Mantsala. 1978. Purification and properties of glutamate synthase and glutamate dehydrogenase from *Bacillus megaterium*. Biochem. J. **173:**45-52.

Henikoff, S., G. W. Haughn, J. M. Calvo, and J. C. Wallace. 1988. A large family of bacterial activator proteins. Proc. Natl. Acad. Sci. USA **85:**6602-6606.

Herrero, A. A. 1983. End-product inhibition in anaerobic fermentations. Trends in Biotechnol. **1:**49-53.

Hilgert, U., J. Schell, and F. J. de Bruijn. 1987. Isolation and characterization of Tn5-induced NADPH-glutamate synthase (GOGAT-) mutants of *Azorhizobium sesbaniae* ORS571 and cloning of the corresponding *glt* locus. Mol. Gen. Genet. **210:**195-202.

Hill, R. T., J. R. Parker, H. J. K. Goodman, D. T. Jones, and D. R. Woods. 1989. Molecular analysis of a novel glutamine synthetase of the anaerobe *Bacteroides fragilis*. J. Gen. Microbiol. (In press)

Hirschman, J., P.-K. Wong, K. S. Keener, and S. Kustu. 1985. Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transcription in vitro: evidence that the *ntrA* product is a sigma factor. Proc. Natl. Acad. Sci. USA **82:**7525-7529.

Holtel, A., and M. Merrick. 1988. Identification of the *Klebsiella pneumoniae glnB* gene: nucleotide sequence of wild-type and mutant alleles. Mol. Gen. Genet. **215**:134-138.

Hongo, M., A. Murata, K. Kono, and F. Kato. 1968. Lysogeny and bacteriocinogeny in strains of *Clostridium* species. Agric. Biol. Chem. **32**:27-33.

Horinouchi, S., W.-H. Beyon, and Weisblum. 1983. A complex attenuator regulates inducible resistance to macrolides, lericosamides, and streptogramin type B antibiotics in *Streptococcus sanguis*. J. Bact. **154**:1252-1262.

Howitt, S. M.,and P. M. Gresshoff. 1985. Ammonia regulation of glutamine synthetase in *Rhizobium* sp. ANU289. J. Gen. Microbiol. **131**:1433-1440.

Huala, E., and F. M. Ausubel. 1989. The central domain of *Rhizobium meliloti* NifA is sufficient to activate transcription from the *R. meliloti nifH* promoter. J. Bacteriol. **171**:3354-3365.

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

Hunt, T. P., and B. Magasanik. 1985. Transcription of *glnA* by purified *Escherichia coli* components: core RNA polymerase and the products of *glnF*, *glnG*, and *glnL*. Proc. Natl. Acad. Sci. USA **82**:8453-8457.

Ikemura, T., and H. Ozeki. 1983. Codon usage anf transfer RNA contents: organism specific codon choice patterns in reference to isoacceptor contents. Cold Spring Harbor Symp. Quant. Biol. **47**:1087-1097.

Inouye, M., and N. Delihas. 1988. Small RNAs in the prokaryotes: a growing list of diverse roles. Cell **53**:5-7.

Ishihama, A. 1988. Promoter selectivity of prokaryotic RNA polymerases. Trends in Genet. **4**:282-286.

Jacquet, M.-A., R. Ehrlich, and C. Reiss. 1989. *In vivo* gene expression directed by a synthetic promoter construction restricted to the -10 and -35 consensus hexamers of *E. coli*. Nucl. Ac. Res. **17**:2933-2945.

Janson, C. A., P. S. Kayne, R. J. Almassy, M. Grustein, and D. Eisenberg. 1986. Sequence of glutamine synthetase from *Salmonella typhimurium* and implications for the protein structure. Gene **46**:297-300.

Janssen, P. J., W. A. Jones, D. T. Jones, and D. R. Woods. 1988. Molecular analysis and regulation of the *glnA* gene of the Gram-positive anaerobe *Clostridium acetobutylicum*. J. Bacteriol. **170**:400-408.

Janssen, P. J., D. T. Jones, and D. R. Woods. 1989. Regulation of *Clostridium acetobutylicum glnA* promoters in *Escherichia coli*. Molecul. Microbiol. (Submitted)

Jiang, J., B. Gu, L. M. Albright, B. T. Nixon. 1989. Conservation between coding and regulatory elements of *Rhizobium meliloti* and *Rhizobium leguminosarum dct* genes. J. Bacteriol. **171**:5244-5253.

Jollif, G., P. Béguin, and J.-P. Aubert. 1986. Nucleotide sequence of the cellulase gene *celD* encoding endoglucanase D of *Clostridium thermocellum*. Nucl. Ac. Res. **14**:8605-8613.

Jones, D. T., A. van der Westhuizen, S. Long, E. R. Allcock, S. J. Reid, and D. R. Woods. 1982. Solvent production and morphological changes in *Clostridium acetobutylicum*. Appl. Environ. Microbiol. **43**:1434-1439

Jones, D. T., and D. R. Woods. 1986a. Acetone-butanol fermentation revisited. Microbiol. Rev. **50**:484-524.

Jones, D. T., and D. R. Woods. 1986b. Gene transfer, recombination and gene cloning in *Clostridium acetobutylicum*. Microbiol. Sc. **3**:19-22.

Jones, D. T., and D. R. Woods. 1988. Genetic manipulation of *Clostridium acetobutylicum* and other species of *Clostridium*, p. 137-156. *In* J. A. Thomson (ed.), Recombiant DNA and Bacterial Fermentation. CRC press, Boca Raton, Florida.

Jones, R., and R. Haselkorn. 1989. The DNA sequence of the *Rhodobacter capsulatus ntrA*, *ntrB* and *ntrC* gene analogues required for nitrogen fixation. Mol. Gen. Genet. **215**:507-516.

Kajitami, M, and A. Ishihama. 1984. Promoter selectivity of *Escherichia coli* RNA dependent polymerase. Differential stringent control of the multiple promoters from ribosomal RNA and protein operons. J. Biol. Chem. **259**:1951-1957.

Kammerer, W., U. Deuschle, R. Gentz, and H. Bujard. 1986. Functional dissection of *Escherichia coli* promoters: information in the transcribed region is involved in late steps of the overall process. EMBO J. **5**:2995-3000.

Kanamori, K., R. L Weiss, and J. D. Roberts. 1987a. Ammonia assimilation in *Bacillus polymyxa*. N-15 NMR and enzymatic studies. J. Biol. Chem. **262**:11038-11045.

Kanamori, K., R. L. Weiss, and J. D. Roberts. 1987b. Role of glutamate dehydrogenase in ammonia assimilation in nitrogen-fixing *Bacillus macerans*. J. Bacteriol. **169:**4692-4695.

Kanamori, K., R. L. Weiss, and J. D. Roberts. 1988. Glutamate biosynthesis in *Bacillus azotofixans*. J. Biol. Chem. **263:**2817-2823.

Kanamori, K., R. L. Weiss, and J. D. Roberts. 1989. Ammonia assimilation pathways in nitrogen-fixing *Clostridium kluyverii* and *Clostridium butyricum*. J. Bacteriol. **171:**2148-2154.

Kantengwa, S., and M. Ojha. 1985. Role of glutamine synthetase in the inhibition of sporulation in *Bacillus polymyxa*. Arch. Microbiol. **141:**204-208.

Kantengwa, S., and M. Ojha. 1986. *In vivo* inhibition of glutamine synthetase induces sporulation in a protease deficient asporogenous mutant of *Bacillus polymyxa*. Arch. Microbiol. **145:**311-314.

Katinka, M., P. Cossart, L. Sibilli, I. Saint-Girons, M. A. Chalvignac, and G. Le Bras. 1980. Nucleotide sequence of the *thrA* gene of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **77:**5730-5733.

Keener, J., and S. Kustu. 1988. Protein kinase and phosphoprotein phosphotase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: roles of the conserved amino-terminal domain of NTRC. Proc. Natl. Acad. Sci. USA **85:**4976-4980.

Keilty, S. and M. Rosenberg. 1987. Constitutive function of a positively regulated promoter reveals new sequences essential for activity. J. Biol. Chem. **262:**6389-6395.

Kinnaird, J.H., and J. R. S. Ficham. 1983. The complete nucleotide sequence of the *Neorospora crass am* (NADP specific glutamate dehydrogenase) gene. Gene **26:** 253 -260.

Koduri, R. K., D. M. Bedwell, and J. E. Brenchley. 1980. Characterization of a *HindIII*-generated DNA fragment carrying the glutamine synthetase gene of *Salmonella typhimurium*. Gene **11:**227-237.

Kofoid, E. C., and J. S. Parkinson. 1988. Transmitter and receiver modules in bacterial signaling proteins. Proc. Natl. Acad. Sci. USA **85:**4981-4985.

Köhler, T., S. Harayama, J.-L. Ramos, and K. N. Timmis. 1989. Involvement of *Pseudomonas putida* RpoN Í factor in regulation of various metabolic functions. J. Bacteriol. **171:**4326-4333.

Kölling, R., and H. Lother. 1985. AsnC: an autogenously regulated activator of aparagine synthetase A transcription in *Escherichia coli*. J. Bacteriol. **164:**310-315.

Koningsberg, W., and G. N. Godson. 1983. Evidence for use of rare codons in the *dnaG* gene and other regulatory genes of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **80:**687-691.

Kozak, M. 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes and organelles. Microbiol. Rev. **47:**1-45.

Kramer, W., K. Schugart, and H.-J. Fritz. 1982. Directed mutagenesis of DNA cloned in filamentous phage: influence of hemimethylated GATC sites on marker recovery from restriction fragments. Nucl. Ac. Res. **10:** 6475-6485.

Kramer, W., V. Drutsa, H.-W. Jansen, B. Kramer, M. Plugfelder, and H.-J. Fritz. 1984. The gapped duplex DNA approach to oligonucleotide-directed mutation construction. Nucl. Ac. Res. **12:**9441-9456.

Kreft, J., K. J. Burger, and W. Goebel. 1983. Expression of antibiotic resistance genes from *Escherichia coli* in *Bacillus subtilis*. Mol. Gen. Genet. **190:**384-389.

Krishnan, I.S., R. K. Singhal, and R. D. Dua. 1986. Purification and characterization of glutamine synthetase from *Clostridium pasteurianum*. Biochem. **25:**1589-1599.

Kunkel, T. A. 1985. Rapid and efficient site specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82:**488-492.

Kur, J., N. Hasan, and W. Szybalski. 1989. Physical and biological consequences of interactions between integration host factor (IHF) and coliphage lambda late p'_R promoter and its mutants. Gene. **81:**1-15.

Kustu, S., E. Santero, J. Keener, D. Popham, and D. Weiss. 1989. Expression of σ^{54} (*ntrA*)-dependent genes is probably united by a common mechanism. Microbiol. Rev. **53:**367-376.

Lamond, A. I., and A. T. Travers. 1985a. Genetically separable functional elements mediate the optimal expression and stringent regulation of a bacterial tRNA gene. Cell **40:**319-326

Lamond, A. I., and A. T. Travers. 1985b. Stringent control of bacterial transcription. Cell **41:**6-8.

Lazzarini, R. A., and A. E. Dahlberg. 1971. The control of ribonucleic acid synthesis during amino acid deprivation in *Escherichia coli*. J. Biol. Chem. **246**:420-429.

Lee, S. W., and G. Edlin. 1985. Expression of tetracycline resistance in pBR322 deivatives reduces the fitness of plasmid-containing *Escherichia coli*. Gene **39**:173-180.

Lei, M., U. Aebi, E. G. Heidner, and D. Eisenberg. 1979. Limited proteolysis of glutamine synthetase is inhibited by glutamate and by feedback inhibitors. J. Biol. Chem. **254**:3129-3134.

Leon, P., D. Romero, A. Garciarrubio, F. Bastarrachea, and A. A. Covurrabias. 1985. Glutamine synthetase constitutive mutations affecting the *glnALG* upstream promoter of *Escherichia coli*. J. Bacteriol. **164**:1032-1038.

Levi-Meyrueis, C., K. Fodor, and P. Schaeffer. 1980. Polyethyleneglycol-induced transformation of *Bacillus subtilis* protoplasts by bacterial chromosomal DNA. Mol. Gen. Genet. **179**:589-594.

Lismaa, S. E., and J. M. Watson. 1989. The *nifA* gene product from *Rhizobium leguminosarum* biovar *trifolii* lacks the N-terminal domain found in other NifA proteins. Molecul. Microbiol. **3**:943-955.

Long, S., D. T. Jones, and D. R. Woods. 1983. Sporulation of *Clostridium acetobutylicum* P262 in a defined medium. Appl. Env. Microbiol. **45**:1389-1393.

Long, S., D. T. Jones, and D. R. Woods. 1984. Initiation of solvent production, clostridial stage and endospore formation in *Clostridium acetobutylicum* P262. Appl. Microbiol. Biotechnol. **20**:256-261.

Looman, A. C., J. Bodlaender, L. J. Comstock, D. Eaton, P. Jhurani, H. A. de Boer, and P. H. van Knippenberg. 1987. Influence of the codon following the AUG initiaition codon on the expression of a modified *lacZ* gene in *Escherichia coli*. EMBO J. 6:2489-2492.

Looman, A. C., J. Bodlaender, M de Gruyter, A. vogelaar, and O. H. van Knippenberg. 1986. Secondary structure as primary determinant of the efficiency of ribosomal binding site in Escherichia. NAR **14**:5481-5497.

Losick, R., and P. J. Youngman. 1984. Endospore formation in *Bacillus subtilis*. *In* Microbial Development, R. Losick, L. Shapiro (eds). Cold Spring Harbor, NY: Cold Spring Harbor Lab., pp. 63-88.

Losick, R., P. Youngman, and P. J. Pigot. 1986. Genetics of endospore formation in *Bacillus subtilis*. Ann. Rev. Genet. **20**:625-669.

Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall. 1957. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193**:265-275.

Lozinski, T., W. T. Markiewicz, T. K. Wyrzykiewicz, and K. L. Wierzchowski. 1989. Effect of the sequence-dependent structure of the 17 bp AT spacer on the strength of consensus-like *E. coli* promoters *in vivo*. Nucl. Ac. Res. **17**:3855-3863.

Lozoya, E., R. Sanchez-Pescador, A. A. Covarrubias, I. Vichido, and F. Bolivar. 1983. Tight linkage of the genes that encode the two glutamate synthase subunits of *Escherichia coli* K-12. J. Bacteriol. **144**:616-621.

MacNeil, T., D. MacNeil, and B. Tyler. 1982a. Fine-structure deletion map and complementation analysis of the *glnA glnL-glnG* region in *Escherichia coli*. J. Bacteriol. **150**:1302-1313.

MacNeil, T., G. P. Roberts, D. MacNeil, and B. Tyler. 1982b. The products of *glnL* and *glnG* are bifunctional proteins. Mol. gen. genet. **188**:325-333.

Madonna, M. J., R. L. Fuchs, and J. E. Brenchley. 1985. Fine structure analysis of *Salmonella typhimurium* glutamate synthase genes. J. Bacteriol. **161**:353-360.

Maeda, S., Y. Ozawa, T. Mizuno, and S. Mizushima. 1988. Stereospecific positioning of the *cis*-acting sequence with respect to the canonical promoter is required for activation of the *ompC* gene by a positive regulator, OmpR, in *Escherichia coli*. J. Mol. Biol. **202**:433-441.

Magasanik, B. 1982. Genetic control of nitrogen assimilation in bacteria. Ann. Rev. Genet. **16**:135-168.

Magasanik, B. 1988. Reversible phosphorylation of an enhancer binding protein regulates the transcription of bacterial nitrogen utilization genes. TIBS **13**:475-479.

Magasanik, B. and F. C. Neidhardt. 1987. Regulation of carbon and nitrogen utilization, p. 1318-1325. *In* Neidhardt (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington D. C.

Maharaj, R., E. Rumbak, W. A. Jones, S. M. Robb, F. T. Robb, and D. R. Woods. 1989. Nucleotide sequence of the *Vibrio alginolyticus glnA* region. Arch. Microbiol. **152**:542-549.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1984. Molecular Cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.

Martin, G. B., K. J. A. Chapman, and B. K. Chelm. 1988. Role of the *Bradyrhizobium japonicum ntrC* gene product in differential regulation of the glutamine synthetase II gene (*glnII*). J. Bacteriol. **170:**5452-5459.

Martin, J. R., H. Petitdemange, J. Ballongue, and R. Gay. 1983. Effects of acetic and butyric acids on solvent production by *Clostridium acetobutylicum.* Biotechnol. Lett. **5:**89-94.

Masters, D. S., and A. Meister. 1982. Inhibition by homocysteine sulfonamide of glutamate synthase purified from *Saccharomyces cerevisiae.* J. Biol. Chem. **257:**8711-8715.

McAllister, C. F., and E. C. Achberger. 1989. Rotational orientation of upstream curved DNA affects promoter function in *Bacillus subtilis.* J. Biol. Chem. **264:**10451-10456.

McClure, W. R. 1985. Mechanism and control of transcription initiation in procaryotes. Ann. Rev. Biochem. **54:**171-204

McNeil, B., and B. Kristiansen. 1985. Effect of temperature upon growth rate and solvent production in batch cultures of *Clostridium acetobutylicum.* Biotechnol. Lett. **7:**499-502.

McNeil,B., and B. Kristiansen. 1986. The acetone butanol fermentation. Adv. Appl. Microbiol. **31:**61-92.

McPherson, M. J.,and J. C. Wootton. 1983. complete nucleotide sequence of the *Escherichia coli gdhA* gene. Nucl. Ac. Res. **11:**5257-5266.

Meers, J. L., and D. W. Tempest. 1970. 'Glutamine(amide): 2-oxoglutarate amino transferase oxido-reductase (NADP)', an enzyme involved in the synthesis of glutamate by some bacteria. J. Gen. Microbiol. **64:**187-194.

Mei, B., and R. Jiao. 1988. Purification and properties of glutamate synthase from *Nocardia mediterranei.* J. Bacteriol. **170:**1940-1944.

Merrick, M. J. 1983. Nitrogen control of the *nif* regulon in *Klebsiella pneumoniae*: involvement of the *ntrA* gene and analogies between *ntrC* and *nifA*. EMBO J. **2:**39-44.

Merrick, M. J., and J. R. Gibbins. 1985. The nucleotide sequence of the nitrogen-regulation gene *ntrA* of *Klebsiella pneumoniae* and comparison with conserved features in Bacterial RNA polymerase sigma factors. Nucl. Ac. Res. **13:**7607-7620.

Merrick, M. J., and W. D. P. Stewart. 1985. Studies on the regulation and function of the *Klebsiella pneumoniae ntrA* gene. Gene **35:**297-303.

Messing, J. 1983. New M13 vectors for cloning. Methods Enzymol. **101:**20-78.

Miller, E. S., and J. E. Brenchley. 1984. Cloning and characterization of *gdhA,* the structural gene for glutamate dehydrogenase of *Salmonella typhimurium.* J. Bacteriol. **157:**171-178.

Miller, R. E., and E. R. Stadtman. 1972. Glutamate synthase from *E. coli*: An iron-sulfide flavoprotein. J. Biol. Chem. **109:**275-301.

Miller, S. I., A. M. Kukral, and J. J. Mekalanos. 1989. A two-component regulatory system (*phoP phoQ*) controls *Salmonella typhimurium* virulence. Proc. Natl. Acad. Sci. USA **86:**5054-5058.

Minchin, S. D., S. Austin, and R. A. Dixon. 1989. Transcriptional activation of the *Klebsiella pneumoniae nifLA* promoter by NTRC is face-of-the-helix dependent and the activator stabilizes the interaction of sigma 54-RNA polymerase with the promoter. EMBO J. **8:**3491-3499.

Miranda-Rios, J., R. Sanchez-Pescador, M. Urdea, and A. A. Covarrubias. 1987. The complete nucleotide sequence of the *glnALG* operon of *Escherichia coli K12.* Nucl. Ac. Res. **15:**2757-2770.

Misra, T. K. 1985. A new strategy to create ordered deletions for rapid nucleotide sequencing. Gene **34:** 263-268.

Mitani, T., J. E. Heinze, and E. Freese. 1977. Induction of sporulation in *Bacillus subtilis* by decoyinine or hadacidin. Biochem. Biophys. Res Comm. **77:**1118-1125.

Mitzushima-Sugano, J., and Y. Kaziro. 1985. Regulation of the expression of the *tufB* operon: DNA sequences directly involved in the stringent control. EMBO J. **4:**1053-1058.

Mongkolsuk, S., Y. W. Chiang, R. N. Reynolds, and P. S. Lovett. 1983. Restriction fragments that exert promoter activity during post-exponentional growth of *Bacillus subtilis.* J. Bacteriol. **155:**1399-1406.

Mongkolsuk, S., E. J. Duvall, and P. S. Lovett. 1985. Transcription termination signal for the *cat-86* indicator gene in a *Bacillus subtilis* promoter cloning plasmid. Gene **37:**83-90.

Monot, F., and J. M. Engasser. 1983. Production of acetone and butanol by batch and continuous culture of *Clostridium acetobutylicum* under nitrogen limitation. Biotechnol. Lett. **5:**213-218.

Monroe, D. M., C. M. Noyes, M. J. Griffith, R. L. Lundblad, and H. S. Kingdon. 1985. Structural and enzymatic properties of *Escherichia coli* glutamine synthetase subjected to limited proteolysis. Curr. Top. Cellul. Regul. **27:**361-372.

Monroe, D. M., C. M. Noyes, R. L. Lundblad, H. S. Kingdon, and M. J. Griffith. 1984. Interaction of substrates with glutamine synthetase after limited proteolysis. Biochem. **23:**4565-4572.

Moon, K., and E. L. Smith. 1973. Sequences of bovine liver glutamate dehydrogenase. VIII. Peptides produced by specific chemical cleavages, the complete sequence of the protein. J. Biol. Chem. **248:**3082-3088.

Moran, C. P.,Jr., N. Lang, S. F. J. LeGrice, G. Lee, M. Stephens, A. L. Sonenshein, J. Pero, and R. Losick. 1982. Nucleotide sequences that signal the initiation of transcription and translation in *Bacillus subtilis*. Mol. Gen. Genet. **186:**339-346.

Morgan, W. D., D. G. Bear, B. L. Litchman, and P. von Hippel. 1985. RNA sequence and secondary structure requirements for rho-dependent termination. Nucl. Ac. Res. **13:**3739-3754.

Mountain, A., M. J. McPherson, A. J. Baron, and J. C. Wootton. 1985. The *Klebsiella aerogenes* glutamate dehydrogenase (*gdhA*) gene: cloning, high-level expression and hybrid enzyme formation in *Escherichia coli*. Mol. Gen. Genet. ***199*:**141-145.

Mullin, D. A., A. Newton. 1989. *Ntr*-like promoters and upstream regulatory sequence *ftr* are required for transcription of a developmentally regulated *Caulobacter crescentus* flagellar gene. J. Bacteriol. **171:**3218-3227.

Nagasu, T., and B. D. Hall. 1985. Nucleotide sequence of the GDH gene coding for the NADP-specific glutamate dehydrogenase of *Saccharomyces cerevisiae*. Gene **37:**247-253.

Nagatani, H., M. Shimizu, and C. Valentine. 1971. The mechanism of ammonia assimilation in nitrogen fixing bacteria. Arch. Mikrobiol. **79:**164-171.

Nakajima, R., T. Imanaka, and S. Aiba. 1985. Nucleotide sequence of the *Bacillus stearothermophilus* α-amylase gene. J. Bacteriol. **163:**401-406.

Nakano, Y., C. Kato, E. Tanaka, K. Kimura, and K. Horikoshi. 1989b. The nucleotide sequence of the glutamine synthetase gene (*glnA*) and its upstream region from *Bacillus cereus*. J. Biochem. **106:**209-215.

Nakano, Y., E. Tanaka, C. Kato, K. Kimura, and K. Horikoshi. 1989a. The complete nucleotide sequence of the glutamine synthetase gene (*glnA*) of *Bacillus subtilis*. FEMS Mirobiol. Lett. **57:**81-86.

Newbury, S. F., N. H. Smith, and C. F. Higgins. 1987. Differential mRNA stability controls relative gene expression within a polycistronic operon. Cell **51:**1131-1143.

Nilsson, G., J. G. Belasco, S. N. Cohen, and A. von Gabain. 1984. Growth-rate dependent regulation of mRNA stability in *Escherichia coli*. Nature (London) **312:**75-77.

Ninfa, A. J., and B. Magasanik. 1986. Covalent modification of the *glnG* product, NR_I, by the *glnL* product, NR_{II}, regulates the transcription of the *glnALG* operon in *Escherichia coli*. Proc. Natl. Acad. Sc. USA **83:**5909-5913.

Ninfa, A. J., E. G. Ninfa, A. N. Lupas, and A. Stock. 1988. Crosstalk between bacterial chemotaxis signal transduction proteins and regulators of transcription of the Ntr regulon: evidence that nitrogen assimilation and chemotaxis are controlled by a common phosphotransfer mechanism. Proc. Natl. Acad. Sci. USA **85:**5492-5496.

Ninfa, A. J., D. A. Mullin, G. Ramakrishnan, and A. Newton. 1989. *Escherichia coli* σ^{54} RNA polymerase recognizes *Caulobacter crescentus flbG* and *flaN* flagellar gene promoters *in vitro*. J. Bacteriol. **171:**383-391.

Nixon, B. T., C. W. Ronson, and F. Ausubel. 1986. Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntrB* and *ntrC*. Proc. Natl. Acad. Sci. USA **83:**7850-7854.

Nohno, T., S. Noji, S. Tanigchi, and T. Saito. 1989. The *narX* and *narL* genes encoding the nitrate-sensing regulators of *Escherichia coli* are homologous to a family of prokaryotic two-component regulatory genes. Nucl. Ac. Res. **17:**2947-2957.

Nordström, K., G. H. Wagner, C. Persson, P. Blomberg, and M. Öhman. 1988. Translational control by antisense RNA in control of plasmid replication. Gene **72**:237-240.

Norrander, J., T. Kempe, and J. Messing. 1983. Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. Gene **261**:101-106.

O'Brien, R. W., and J. G. Morris. 1971. Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. J. Gen. Microbiol. **68**:307-318.

Ogasawara, N. 1985. Markedly unbiased codon usage in *Bacillus subtilis*. Gene **40**:145-150.

Ogata, S., and M. Hongo. 1979. Bacteriophages of the genus *Clostridium*. Adv. Appl. Microbiol. **25**:241-273.

Ojha, M., and S. Kantengwa. 1989. Purification and properties of glutamine synthetase from *Bacillus polymyxa*. Arch. Mirobiol. **151**:294-299.

Okamoto, K., and M. Freundlich. 1986. Mechanism for the autogenous control of the *crp* operon: transcriptional inhibition by a divergent RNA transcript. Proc. Natl. Acad. Sc. USA **83**:5000-5004.

Oliphant, A. R., K. Struhl. 1988. Defining the consensus sequences of *E. coli* promoter elements by random selection. Nucl. Ac. Res. **16**:7673-7683.

Oliver, G., G. Gosset, R. Sanchez-Pescador, E. Lozoya, L. M. Ku, N. Flores, B. Becerril, F. Valle, and F. Bolivar. 1987. Determination of the nucleotide sequence for the glutamate synthase structural genes of *Escherichia coli* K-12. Gene **60**:1-11.

Osorio, A. V., L. Servin-Gonzalez, M. Rocha, A. A. Covarrubias, and F. Bastarrachea. 1984. *cis*-Dominant, glutamine synthetase constitutive mutations of *Escherichia coli* independent of activation by the *glnG* and *glnF* products. Mol. Gen. Genet. **194**:114-123.

Oultram, J. D. and M. Young. 1985. Conjugal transfer of plasmid pAMβ1 from *Streptococcus lactis* and *Bacillus subtilis* to *Clostridium acetobutylicum*. FEMS Microbiol. Lett. **27**:129-134.

Oultram, J. D., A. Davies, and M. Young. 1987. Conjugal transfer of a small plasmid from *Bacillus subtilis* to *Clostridium acetobutylicum* by cointegrate formation with plasmid pAMβ1. FEMS Microbiol. Lett. **42**:113-119.

Oultram, J. D., M. Louglin, T.-J. Swinfield, J. K. Brehm, D. E. Thompson, and N. P. Minton. 1988. Introduction of plasmids into whole cells of *Clostridium acetobutylicum* by electroporation. FEMS Microbiol. Lett. **56**:83-88.

Pahel, G., A. D. Zelenetz, and B. Tyler. 1978. *gltB* gene and regulation of nitrogen metabolism by glutamine synthetase in *Escherichia coli*. J. Bacteriol. **133:**139-148.

Pahel, G., and B. Tyler. 1979. A new *glnA*-linked regulatory gene for glutamine synthetase in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **76:**4544-4558.

Pahel, G., D. M. Rothstein, and B. Magasanik. 1982. Complex *glnA-glnL-glnG* operon of *Escherichia coli*. J. Bacteriol. **150:**202-213.

Pan, F. L., and J. G. Coote. 1979. Glutamine synthetase activities during growth and sporulation in *Bacillus subtilis*. J. Gen. Microbiol. **112:**373-377.

Pawlowski, K., P. Ratet, J. Schell, and F.J. de Bruijn. 1987. Cloning and characterization of *nifA* and *ntrC* genes of the stem nodulating bacterium *Azorhizobium sesbaniae* ORS571, the nitrogen fixing symbiont of *Sesbania rostrata*: regulation of nitrogen fixation (*nif*) genes in the free living versus symbiotic state. Mol. Gen. Genet. **206:**207-219.

Peschke, U., V. Beuck, H. Bujard, R. Gentz, and S. Le grice. 1985. Efficient utilization of *Escherichia coli* transcriptional signals in *Bacillus subtilis*. J. Mol. Biol. **186:**547-555.

Petersen, C. 1987. The functional stability of the *lacZ* transcript is sensitive towards sequence alterations immediately downstream of the ribosome binding site. Mol. Gen. Genet. **209:**179-187.

Petersen, C. 1989. Long range translational coupling in the *rplJL-rpoBC* operon in *Escherichia coli*. J. Mol. Biol. **206:**323-332.

Petersen, G. B., P. A. Stockwell, and D. F. Hill. 1988. Messenger RNA recognition in *Escherichia coli*: a possible site of interaction with 16 S ribosomal RNA. EMBO J. **7:**3957-3962.

Pines, O., and M. Inouye. 1986. Antisense RNA regulation in prokaryotes. Trends Genet. **2:**284-287.

Pinkofsky, H. B., A. Ginsburg, I. Reardon, and R. L. Heinrikson. 1984. Lysyl residue 47 is near the subunit ATP-binding site of glutamine synthetase from *Escherichia coli*. J. Biochem. **259:**9616-9622.

Plaskon, R. R., R. M. Wartell. 1987. Sequence distributions associated with DNA curvature are found upstream of strong *E. coli* promoters. Nucl. Ac. Res. **15:**785-796.

Platt, T. 1986. Transcription termination and the regulation of gene expression. Annu. Rev. Biochem. **55:**339-372.

Ponnambalam, S., C. Webster, A. Bringham, and S. Busbby. 1986. Transcription initiation in the *Escherichia coli* galactose operon promoters in the absence of the normal -35 region sequences. J. Biol. Chem. **261:**16043-16048.

Popham, D. L., D. Szeto, J. Keener, and S. Kustu. 1989. Function of a bacterial activator protein that binds to transcriptional enhancers. Science **243:**629-635.

Ptashne, M. 1986. Gene regulation by proteins acting nearby and at a distance. Nature (London) **322:**697-701.

Pulido, D., and A. Jimenez. 1987. Optimization of gene expression in Streptomyces lividans by a transcription terminator. Nucl. Ac. Res. **15:**4227-4240.

Queen, C., and L. J. Korn. 1984. A comprehensive sequence analysis program for the IBM personal computer. Nucl. Ac. Res. **12:**581-599.

Raibaud, O.. 1989. Nucleoprotein structures at positively regulated bacterial promoters: homology with replication origins and some hypotheses on the quaternary structure of the activator proteins in these complexes. Molecul. Microbiol. **3:**455-458.

Rasched, I., H. Jornvall, and H. Sund. 1974. Studies on glutamate dehydrogenase: identification of an amino group involved in the substrate binding. Eur. J. Biochem. **41:**603-606.

Ratti, S., B. Curti, G. Zanetti, and E. Galli. 1985. Purification and characterization of glutamate synthase from *Azospirillum brasilense.* J. Bacteriol. **163:**724-729.

Rawlings, D. E., W. A. Jones, E. G. O'Neil, and D.R. Woods (1987). Nucleotide sequence of the glutamine synthetase gene and its controlling region from the acidophilic autotroph *Thiobacillus ferrooxidans.* Gene **53:**211-217.

Reitzer, L. J., and B. Magasanik. 1983. Isolation of the nitrogen assimilation regulator NR_I, the product of the *glnG* gene of *Escherichia coli.* Proc. Natl. Acad. Sci. USA **80:**5554-5558.

Reitzer, L. J., and B. Magasanik. 1985. Expression of *glnA* in *Escherichia coli* is regulated at tandem promoters. Proc. Natl. Acad. Sci. USA **82:**1979-1983.

Reitzer, L. J., and B. Magasanik. 1986. Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter. Cell **45:**785-792.

Reitzer, L. J., and B. Magasanik. 1987a. Ammonia assimilation and the biosynthesis of glutamine, glutamate, aspartate, asparagine, L-alanine, and D alanine, p. 302-320. *In* Neidhardt (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. Americam Society for Microbiology, Washington D. C.

Reitzer, L. J., B. Movsas, and B. Magasanik. 1989. Activation of *glnA* transcription by nitrogen regulator I (NR_I)-phosphate in *Escherichia coli*: evidence for a long range physical interaction between the NR_I-phosphate and RNA polymerase. J. Bacteriol. **171**:5512-5522.

Reitzer, L.J., R. Bueno, W. D. Cheng, S. A. Abrams, D. M. Rothstein, T. Hunt, B. Tyler, and B. Magasanik. 1987b. Mutations that create new promoters suppress the σ^{54} dependence of *glnA* transcription in *Escherichia coli*. J. Bacteriol. **169**:4279-4284.

Rendina, A. R., and R. W. Orme-Johnson. 1978. Glutamate synthase: on the kinetic mechanism of the enzyme from *Escherichia coli* W. Biochem. **17**:5388-5393.

Reysset, G. 1981. New class of *Bacillus subtilis* glutamine requiring mutants. J. Bacteriol. **148**:653-658.

Reysset, G., and M. Sebald. 1985. Conjugal transfer of plasmid-mediated anibiotic resistance from streptococci to *Clostridium acetobutylicum*. Ann. Microbiol. Inst. Pasteur **136**:275-282.

Reysset, G., J. Hubert, L. Podvin, and M. Sebald. 1988. Transfection and transformation of *Clostridium acetobutylicum* strain N1-4081 proplasts. Biotechnol. Tech. **2**:199-304.

Reysset, G., J. Hubert, L. Podvin, and M. Sebals. 1987. Protoplast transformation and regeneration of *Clostridium acetobutylicum* strain N1-4080. J. Gen. Microbiol. **133**:2595-2600.

Reysset, G., K. F. Bott, and J.-P. Aubert. 1978. Relationship between glutamine synthetase and sporulation, p.271-276. *In* Chambliss and Vary (eds.), Spores VII. American Society for Microbiology, Washington, D.C.

Rhee, S. G., P. B. Chock, and E. R. Stadtman. 1985. Glutamine synthetase from *Escherichia coli*. Methods in enzymology **113**:213-241.

Riccardi G., E. De Rossi, G. Della Valle, and O. Ciferri. 1985. Cloning of the glutamine synthetase gene from *Spirulina platensis*. Plant Molecul. Biol. **4**:133-136.

Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translationwith DNA polymerase I. J. Mol. Biol. **113**:237-251.

Riggs, D. L., R. D. Müller, H.-S. Kwan, and S. W. Artz. 1986. Promoter domain mediates guanosine tetraphosphate activation of the histidine operon. Proc. Natl. Acad. Sci. USA **83**:9333-9337.

Roberts, J. W. 1969. Termination factor for RNA synthesis. Nature (London) **224**:1168-1174.

Rocha, M., M. Vazquez, A. Garciarrubio, and A. A. Covarrubias. 1985. Nucleotide sequence of the *glnA-glnL* intercistronic region of *Escherichia coli*. Gene **37**:91-99.

Rogers, P. 1986. Genetics and biochemistry of *Clostridium* relevant to developments of fermentation processes. Adv. Appl. Microbiol. **31**:1-60.

Römermann, D., J. Warrelmann, R. A. Bender, and B. Friedrich. 1989. An *rpoN*-like gene of *Alcaligenes eutrophus* and *Pseudomonas facilis* controls expression of diverse metabolic pathways, including hydrogen oxidation. J. Bacteriol. **171**:1093-1099.

Ronson, C. W., B. T. Nixon, and F. M. Ausubel. 1987a. Conserved domains in bacterial regulatory proteins that respond to environmental stimuli. Cell **49**:579-581.

Ronson, C. W., B. T. Nixon, L. M. Albright, and F. M. Ausubel. 1987b. *Rhizobium meliloti ntrA* (*rpoN*) gene is required for diverse metabolic functions. J. Bacteriol. **169**:2424-2431.

Ronson, C. W., P. M. Astwood, and J. A. Downie. 1984. Molecular cloning and genetic organization of C_4 dicarboxylate transport genes from *Rhizobium leguminosarum*. J. Bacteriol. **160**:903-909.

Ronson, C. W., P. M. Astwood, B. T Nixon, and F. M. Ausubel. 1987c. Deduced products of C_4-dicarboxylate transport regulatory genes of *Rhizobium leguminosarum* are homologous to nitrogen regulatory gene product. Nucl. Ac. Res. **15**:7921-7934.

Roos, J. W., J. McLaughlin, and E. T. Papoutsakis. 1984. The effect of pH on nitrogen supply, cell lysis, and solvent production in fermentations of *Clostridium acetobutylicum*. Biotechnol. Bioeng. **27**:681-694.

Rosenberg, M., and D. Court. 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. Annu. Rev. Genet. **13**:319-353.

Rossbach, S., J. Schell, and F. J. de Bruijn. 1987. The *ntrC* gene of *Agrobacterium tumefaciens* C58 controls glutamine synthetase (GSII) activity, growth on nitrate, and chromosomal but not Ti-encoded arginine catabolism pathways. Mol. Gen. Genet. **209**:419-426.

Rossbach, S., J. Schell, and F.J. de Bruijn. 1988. Cloning and analysis of *Agrobacterium tumefaciens* V58 loci involved in glutamine biosynthesis: neither the glnA (GSI) nor the glnII (GSII) gene plays a special role in virulence. Mol. Gen. Genet. **212**:38-47.

Rossi, M., R. Defez, M. Chiurazzi, A. Lamberti, A. Fuggi, and M. Iaccarino. 1989. Regulation of glutamine synthetase isoenzymes in *Rhizobium leguminosarum* biovar *viceae*. J. Gene. Microbiol. **135**:629-637.

Rothstein, D. M., G. Pahel, B. Tyler, and B. Magasanik. 1980. Regulation of expression from the *glnA* promoter of *Escherichia coli* in the absence of glutamine synthetase. Proc. Natl. Acad. Sc. USA **77**:7372-7376.

Sales, M., and J. E. Brenchley. 1982. The regulation of the ammonia assimilatory enzymes in rel^+ and rel^- strains of *Salmonella typhimurium*. Mol. Gen. Genet. **186**:263-268.

Salser, W. 1977. Globin mRNA sequences: analysis of basepairing and evolutionary implications. Cold Spring Harbor Symp. Quant. Biol. **42**:985-1002.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sc. USA **74**:5463-5467.

Santero, E., T. Hoover, J. Keener, and S. Kustu. 1989. *In vitro* activity of the nitrogen fixation regulatory protein NIFA. Proc. Natl. Acad. Sc. (In Press)

Sasse-Dwight, S., and J. D. Gralla. 1988. Probing the *Escherichia coli glnALG* upstream activation mechanism i*n vivo*. Proc. Natl. Acad. Sci. USA **85**:8934-8938.

Schauder, B., and J. E. G. McCarthy. 1989. The role of bases upstream of the Shine-Dalgarno region and in the coding region in the control of gene expression in *Escherichia coli* translation and stability of mRNAs in vivo. Gene **78**:59-72.

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

Schmitz, G., K. Nikaido, and G. Ferro-Luzzi Ames. 1988. Regulation of a transport operon promoter in *Salmonella typhimurium*: Identification of sites essential for nitrogen regulation. Mol. Gen. Genet. **215**:107-117.

Schofield, P. R., and J. M. Watson. 1986. DNA sequence of *Rhizobium trifolii* nodulation genes reveals a reiterated and potentially regulatory sequence preceeding *nodABC* and *nodFE*. Nucl. Ac. Res. **14**:2891 -2903.

Schreier, H. J., and A. L. Sonenshein. 1986. Altered regulation of the *glnA* gene in glutamine synthetase mutants of *Bacillus subtilis*. J. Bacteriol. **167**:35-43.

Schreier, H. J., and R. W. Bernlohr. 1984. Purification and properties of glutamate synthase from *Bacillus licheniformis*. J. Bacteriol. **160**:591-599.

Schreier, H. J., S. H. Fisher, and A. L. Sonenshein. 1985. Regulation of expression from the *glnA* promoter of *Bacillus subtilis* requires the *glnA* gene product. Proc. Natl. Acad. Sci. USA **82**:3375-3379.

Schreier, H. J., S. W. Brown, K. D. Hirschi, J. F. Nomellini, and A. L. Sonenshein. 1989. Regulation of *Bacillus subtilis* glutamine synthetase gene expression by the product of the *glnR* gene. J. Mol. Biol. **210**:51-63.

Schreier, H. J., T. M. Smith, T. J. Donohue, and R. W. Bernlohr. 1981. Regulation of nitrogen metabolism and sporulation in *Bacillus licheniformis*, p.138-149.. *In* Levinson, Sonenshein, and Tipper (eds.), Sporulation and Germination. American Society for Microbiology, Washington, D.C.

Scolnik, P. A., J. Virosco, and R. Haselkorn. 1983. The wild-type gene for glutamine synthetase restores ammonia control of nitrogen fixation to Gln$^-$ (*glnA*) mutants of *Rhodopseudomonas capsulata*. J. Bacteriol. **155**:180-185.

Shand, R. F., P. H. Blum, R. D. Mueller, D. L. Riggs, and S. W. Artz. 1989. Correlation between histidine operon expression and guanosine 5'-diphosphate-3'-diphosphate levels during amino acid downshift in stringent and relaxed strains of *Salmonella typhimurium*. J. Bact. **171**:737-743.

Shapira, S. K., J. Chan, F. V. Richard, and M. J. Casadaban. 1983. New versatile plasmid vectors for expression of hybrid proteins coded by a cloned gene fused to *lacZ* gene sequences encoding an enzynatically active carboxy-terminal portion of β-galactosidase. Gene **25**:71-82.

Shapiro, B. M., and E. R. Stadtman. 1968. 5'adenyl-o-tyrosine, the novel phosphodiester of adenylylated glutamine synthetase from *Escherichia coli*. J. Biol. Chem. **243**:3769-3771.

Shapiro, B. M., and E. R. Stadtman. 1970. The regulation of glutamine synthesis in microorganisms. Annu. Rev. Microbiol. **24**:501-524.

Shimotsu, H., M. I. Kuroda, C. Yanofsky, and D. J. Henner. 1986. Novel form of transcription attenuation regulates expression of the *Bacillus subtilis* tryptophan operon. J. Bacteriol. **166**:461-471.

Shine, J., and L. Dalgarno. 1974. The 3' terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementary to nonsense triplets and ribosomal binding sites. Proc. Natl. Acad. Sci. USA **71**:1342-1346.

Siebenlist, U., and W. Gilbert. 1980. Contacts between *Escherichia coli* RNA polymerase and an early promoter of phage T7. Proc. Natl. Acad. Sci. **77**:122-126.

Siegele, D. A., J. C. Hu, W. A. Walter, and C. A. Gross. 1989. Altered promoter recognition by mutant forms of the σ^{70} subunit of *Escherichia coli* RNA polymerase. J. Mol. Biol. **206**:591-603.

Simons, R. W. 1988. Naturally occuring RNA control - a brief review. Gene **72**:35-44.

Simons, R. W., and N. Kleckner. 1983. Translational control of IS*10* transposition. Cell **34**:683-691.

Somerville, J. E., and M. Kahn. 1983. Cloning of the glutamine synthetase I gene from *Rhizobium meliloti*. J. Bacteriol. **156**:168-176.

Son, H. S., and S. G. Rhee. 1987. Cascade control of *Escherichia coli* glutamine synthetase: purification and properties of P_{II} protein and nucleotide sequence of its structural gene. J. Biol. Chem. **262**:8690-8695.

Sonenshein, A. L. 1989. Metabolic regulation of sporulation and other stationary phase phenomena. In Regulation of Prokaryotic Development. Setlow, P., R. Slepecky, and I. Smith (eds). Washington, D.C.: American Society for Microbiology, pp. 109-130.

Sorensen, M. A., C. G. Kurland., and S. Pedersen. 1989. Codon usage determines translation rate in *Escherichia coli*. J. Mol. Biol. **207**:365-377.

Southern, J. A., J. R. Parker, and D. R. Woods. 1987. Novel structure, properties and inactivation of glutamine synthetase cloned from *Bacteroides fragilis*. J. Gen. Microbiol. **133**:2437-2446.

Southern, J. A., J. R. Parker, and D. R. Woods. 1987. Novel structure, properties and inactivation of glutamine synthetase cloned from *Baceroides fragilis*. J. Gen. Microbiol. **133**:2437-2466.

Southern, J.A., J. R. Parker, and D. R. Woods. 1986. Expression and purification of glutamine synthetase cloned from *Bacteroides fragilis*. J. Gen. Microbiol. **132**:2827-2835.

Spivey, M. J. 1978. The acetone/butanol/ethanol fermentation. Proc. Biochem. **13**:2-5.

Stadtman, E. R., and A. Ginsburg. 1974. The glutamine synthetase of *Escherichia coli*: structure and control, p. 755-807. *In* Boyer (ed.), The Enzymes (X). Academic Press, Inc., New York.

Stanley, J., J. van Slooten, D. N. Dowling, T. Finan, and W. J. Broughton. 1989. Molecular cloning of the *ntrA* gene of the broad host-range *Rhizobium* sp. NGR234 and phenotypes of a site-directed mutant. Mol. Gen. Genet. **217**:528-532.

Stanssens, P., C. Opsomer, Y. M. McKeown, W. Kramer, M. Zabeau, and H.-J. Fritz. 1989. Efficient oligonucleotide-directed construction of mutations in expression vectors by the gapped duplex DNA method using alternating selectable markers. Nucl. Ac. Res. **17**:4441-4454.

Stanssens, P., E. Remaut, and W. Fiers. 1985. Alterations upstream from the Shine-Dalgarno region and their effect on bacterial gene expression. Gene **36**:211-223.

Stock, A., D. E. Koshland, Jr., and F. Stock. 1985. Homologies between the *Salmonella typhimurium* CheY protein and proteins involved in the regulation of chemotaxis, membrane protein synthesis, and sporulation. Proc. Natl. Acad. Sci. **82**:7989-7993.

Stock, A. M., D. C. Wylie, J. M. Mottonen, A. N. Lupas, E. G. Ninfa, A. J. Ninfa. 1988. Phosphoproteins involved in bacterial signal transduction. Col Sprin Harbor Symp. Quant. Biol. **8**: 49-56.

Stoker, K., W. N. M. Reijnders, L. F. Oltmann, and A. H. Stouthamer. 1989. Initial cloning and sequencing of *hydHG*, an operon homologous to *ntrBC* and regulating the labile hydrogenase activity in *Escherichia coli* K-12. J. Bact. **171**:4448-4456.

Strauch, M. A., A. I. Aronson, S. W. Brown, H. J. Schreier, and A. L. Sonenshein. 1988. Sequence of the *Bacillus subtilis* glutamine synthetase gene region. Gene **71**:257-265.

Streicher, S. L., and B. Tyler. 1981. Regulation of glutamine synthetase activity by adenylylation in the Gram-positive bacterium *Streptomyces cattleya*. Proc. Natl. Acad. Sc. USA **78**:229-233.

Struhl, K. 1985. A rapid method for creating recombinant DNA Molecules. Biotechniques **3**:452-453.

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

Szeto, W. W., B. T. Nixon, C. W. Ronson, and F. M. Ausubel. 1987. Identification and characterization of the *Rhizobium meliloti ntrC* gene. *R. Meliloti* has separate regulatory pathways for activation of nitrogen fixing genes in free-living and symbiotic cells. J. Bacteriol. **169**:1423-1432.

Taylor, J. W., J. Ott, and F. Eckstein.1985. The rapid generation of oligonucleotide mutations at high frequency using phosphorothioate-modified DNA. Nucl. Ac. Res. **13**:8765-8785.

Terracciano, J. S., and E. R. Kashket. 1986. Intracellular conditions required for initiation of solvent production by *Clostridium acetobutylicum*. Appl. Environ. Microbiol. **52**:86-91.

Thanaraj, T. A., and M. W. Pandit. 1989. An additional ribosome-binding site on mRNA of highly expressed genes and a bifunctional site on the colicin fragment of 16 S rRNA from *Escherichia coli*: important determinations of the efficiency of translation-initiation. Nucl. Ac. Res. **17**:2973-2985.

Tischer, E., S. DasSarma, and H. M. Goodman. 1986. Nucleotide sequence of an *alfalfa* glutamine synthetase gene. Mol. Gen. Genet. **203**:221-229.

Toukdarian, A., and C. Kennedy. 1986. Regulation of nitrogen metabolism in *Azotobacter vinelandii*: isolation of *ntr* and *glnA* genes and construction of *ntr* mutants. EMBO J. **5**:399-407.

Trach, K. A., J. W. Chapman, P. J. Piggot, and J. A. Hoch. 1985. Deduced product of the stage 0 sporulation gene *spOF* shares homology with the SpoOA, OmpR, and SfrA proteins. Proc. Natl. Acad. Sc. USA **82**:7260-7264.

Trifonov, E. N., and V. Brendel. 1987. Gnomic: A dictionary of genetic codes, p. 256-257. VCH Publishers, New York.

Trotta, P., K. Platzer, R. Haschemeyer, and A. Meister. 1974. Glutamine-binding subunit of glutamate synthase and partial reactions catalyzed by this glutamine amido transferase. Proc. Natl. Acad. Sci. USA **71**:4607-4611.

Tuli, R., R. Fisher, and R. Haselkorn. 1982. The *ntr* genes of *Escherichia coli* activate the *hut* and *nif* operons of *Klebsiella pneumoniae*. Gene **19**:109-116.

Tumer, N. E., S. J. Robinson, and R. haselkorn. 1983. Different promoters for the *Anabaena* glutamine synthetase gene during growth using molecular or fixed nitrogen. Nature (London) **306**:337-342.

Tyler, B. 1978. Regulation of the asimilationof nitrogen compounds. Annu. Rev. Biochem. **47**:1127-1162.

Ueno-Nishio, S., K. C. Backman, and B. Magasanik. 1983. Regulation at the *glnL*-operator-promoter of the complex *glnALG* operon of *Escherichia coli*. J. Bacteriol. **153**:1247-1251.

Ueno-Nishio, S., S. Mango, L. J. Reitzer, and B. Magasanik. 1984. Identification and regulation of the *glnL* operator-promoter of the complex *glnALG* operon of *Escherichia coli*. J. Bacteriol. **160**:379-384.

Usdin, K. P., H. Zappe, D. T. Jones, and D. R. Woods. 1986. Cloning, expression, and purification of glutamine synthetase from *Clostridium acetobutylicum*. Appl. Environ. Microbiol. **52**:413-419.

Valle, F., B. Becerril, E. Chen, P. Seeburg, H. Heyneker, and F. Bolivar. 1984. Complete nucleotide sequence of the glutamate dehydrogenase gene from *Escherichia coli* K-12. Gene **27**:193-199.

Verhasselt, P., F. Poncelet, K. Vits, A. van Gool, and J. Vanderleyden. 1989. Cloning and expression of a *Clostridium acetobutylicum* α-amylase gene in *Escherichia coli*. FEMS Microbiol. Lett. **59**:135-140.

Viera, J., and J. Messing. 1982. The pUC plasmids, an M13mp7 derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259-268.

Vold, B. 1985. Structure and organization of genes for transfer ribonucleic acid in *Bacillus subtilis*. Microbiol. Rev. **49**:71-80.

Von Hippel, P. H., D. G. Bear, W. D. Morgan, and J. A. McSwiggen. 1984. Protein-nucleic acid interactions in transcription: a molecular analysis. Annu. Rev. Biochem. **53**:389-446.

Wada, A., and A. Suyama. 1986. Local stability of DNA and RNA secondary structure and its relation to biological functions. Prog. Biophys. Molecul. Biol. **47:**113-157.

Walker, J. E., M. Saraste, M. J. Runwick, and N. J. Gay. 1982. Distantly related sequences in the α- and β subunits of ATP synthetase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1:**945-951.

Wang, Y., M. J. Beach, and V. W. Rodwell. 1989. (S)-3-Hydroxy-3-methylglutaryl coenzyme A reductase, a product of the *mva* operon of *Pseudomonas mevalonii*, is regulated at the transcriptional level. J. Bact. **171:**5567-5571.

Wanner, B. L., M. R. Wilmes, and D. C. Young. 1988. Control of bacterial alkaline phosphotase synthesis and variation in an *Escherichia coli* K12- *phoR* mutant by adenyl cyclase, the cyclic AMP receptor protein, and the *phoM* operon. J. Bacteriol. **170:**1092-1102.

Wardham, H., M. J. McPherson, and G. R. Sastry. 1989. Identification, cloning, and sequence analysis of the nitrogen regulatory gene ntrC of Agrobacterium tumefaciens C58. Molecul. Plant-Microbe Interactions **2:**241:248.

Wedler, F. C., D. S. Shreve, R. M. Kenney, A. E. Ashour, J. Carfi, and S. G. Rhee. 1980. Two glutamine synthetases from the *Bacillus caldolyticus*, an extreme thermophile. J. Biol. Chem. **255:**9507-9516.

Weglenski, P., A. J. Ninfa, S. Ueno-Nishio, and B. Magasanik. 1989. Mutations in the *glnG* gene of *Escherichia coli* that result in increased activity of nitrogen regulator I. J. Bacteriol. **171:**4479-4485.

Williams, D. M., E. Duvall, and P. S. Lovett. 1981. Cloning restriction fragments that promote expression of a gene in *Bacillus subtilis*. J. Bacteriol. **146:**1162-1165.

Winship, P. R. 1989. An improved method for direclty sequencing PCR amplified material using dimethyl sulphoxide. Nucl. Ac. Res. **17:**1266

Wirth, R., A. Friesenegger, and S. Fiedler. 1989. Transformation of various species of gram-negative bacteria belonging to 11 different genera by electroporation. Mol. Gen. Genet. **216:**175-177.

Woods, D. R., and D. T. Jones. 1986. Physiological responses of *Bacteroides* and *Clostridium* strains to environmental stress factors. Adv. Microbiol. Physiol. **27:**1-64.

Woolfolk, C. A., and E. R. Stadtman. 1964. Cumulative feedback inhibition in the multiple end product regulation of glutamine synthetase in *Escherichia coli.* Biochem. Biophys. Res. Comm. **17**:313-319.

Wray, Jr., L. V., and S. H. Fisher. 1988. Cloning and sequence of the *Streptomyces coelicolor* gene encoding glutamine synthetase. Gene 71:247-256.

Wu, T.-H., S.-M. liao, W. R. McClure, and M. M. Susskind. 1988. Control of gene expression in bacteriophage P22 by a small anti-sense RNA, II. Characterization of mutants defective in repression. Genes. Dev. **1**:204-212.

Yager, T. D., and P. H. von Hippel. 1987. Transcript elongation and termination in *Escherichia coli.*, p.1241 -1275. *In* Neidhardt (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.

Yamashita, M. M., R. J. Almassy, C. A. Janson, D. cascio, and D. Eisenberg. 1989. Refined atomic model of glutamine synthetase at a 3.5 Å resolution. J. Biol. Chem. **264**: 17681-17690.

Yanisch-Perron, C., J. Vierra, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp8 and pUC19 vectors. Gene **33**:103-119.

Young, M., N. P. Minton, and W. L. Staudenbauer. 1989. Recent advances in the genetics of the clostridia. FEMS Microbiol. Rev. (In press).

Youngleson, J. S., J. D. Santangelo, D. T. Jones, and D. R. Woods. 1988. Cloning and expression of a *Clostridium acetobutylicum* alcohol dehydrogenase gene in *Escherichia coli*. Appl. Environ. Microbiol. **54**:676-682.

Youngleson, J. S., W. A. Jones, D. T. Jones, and D. R. Woods. 1989a. Molecular analysis and nucleotide sequence of the *adh1* gene encoding a NADPH-dependent butanol dehydrogenase in the Gram-positive anaerobe *Clostridium acetobutylicum*. Gene (In Press)

Youngleson, J. S., W. A. Jones, D. T. Jones, and D. R. Woods. 1989b. Homology between hydroxybutyryl and hydroxy acyl coenzyme A dehydrogenase enzymes from *Clostridium acetobutylicum* fermentation and vertebrate fatty acid β- oxidation pathways. J. Bacteriol. **171**:6800-6807.

Yu, P. L., and L. E. Pearce. 1986. Conjugal transfer of streptococcal antibiotic resistance plasmids into *Clostridium acetobutylicum*. Biotechnol. Lett. **8**:469-474.

Zabeau, M., and K. K. Stanley. 1982. Enhanced expression of cro-beta-galactosidase fusion proteins studied under the control of the Pr promoter of bacteriophage lambda. EMBO J. **1**:1217-1224.

Zacharias, M.,H. U Soringer, and R. Wagner. 1989. Influence of the GCGC disrcrimator motif introduced into the ribosomal RNA p2- and *tac* promoter on growth-rate control and stringent sensitivity. EMBO J. **8**:3357-3363.

Zakin, M. M., N. Duchange, P. Ferrara, and G. N. Cohen. 1983. Nucleotide sequence of the *metL* gene of *Escherichia coli*. J. Biol. Chem. **258**:3028-3031.

Zappe, H., D. T. Jones, and D. R. Woods. 1986. Cloning and expression of *Clostridium acetobutylicum* endoglucanase, cellobiase and amino acid biosynthesis genes in *Escherichia coli*. J. Gen. Microbiol. **132**:1367-1372.

Zappe, H., D. T. Jones, and D. R. Woods. 1987. Cloning and expression of a xylanase gene from *Clostridium acetobutylicum* P262 in *Escherichia coli*. Appl. Micriobiol. Biotechnol. **27**:57-63.

Zappe, H., W. A. Jones, D. T. Jones, and D. R. Woods. 1988. Structure of an endo-β-1,4-glucanase gene from *Clostridium acetobutylicum* P262 showing homology with endoglucanases from *Bacillus* species. Appl. Environ. Microbiol. **54**:1289-1292.

Zell, R., and H.-J. Fritz. 1987. DNA mismatch-repair in *E. coli* counteracting the hydrolytic deamination of 5'-methyl-cytosine residues. EMBO J. **6**:1809-1815.

Zhang, J., M. Strauch, and A. I. Aronson. 1989. Glutamine auxotrophs of *Bacillus subtilis* that overproduce glutamine synthetase antigen have altered conserved amino acids in or near the active site. J. Bacteriol. **171**:3572-3574.