Review of Literature
REVIEW OF LITERATURE

1.1. Prokaryotic expression systems for recombinant proteins

The expression of heterologous proteins in microorganisms using genetic recombination is still the high point in the development and exploitation of modern biotechnology (Yin et al., 2007). Bacterial expression systems for heterologous protein production are attractive because of their ability to grow rapidly and at high density on inexpensive media, their often well-characterized genetics and the availability of an increasingly large number of cloning vectors and mutant host strains. *Escherichia coli* and *Bacillus subtilis* are the most widely used hosts for expression. Other bacterial hosts that have been used for the production of recombinant proteins include *Lactococcus lactis* (Nouaille, 2003), *Bacillus megaterium*, *Bacillus brevis*, *Streptomyces*, *Pseudomonas fluorescens*, etc. (Terpe, 2006).

Nevertheless many bacterial systems are not capable of post-translational modifications of proteins such as glycosylation. If the posttranslational modification is essential for bioactivity, bacterial expression systems should not be used for heterologous protein production. Alternative hosts such as yeasts (*Saccharomyces cerevisiae* and *Yarrowia lipolytica*), filamentous fungi (*Aspergillus* and *Trichoderma reesei*), or insect and mammalian cell cultures are available for this application (Terpe, 2006).

1.1.1. *Escherichia coli* expression systems

*E. coli* has been used extensively as the cellular host for foreign protein expression due to its rapid growth rate, capacity for continuous fermentation, relatively low cost and extensive knowledge in its genetics, molecular biology and fermentation biology. Many strains are available which are adapted for the production of proteins either in the cytoplasm or periplasm, and a large number of expression vectors with differently regulated promoters and tags for efficient protein purification have been constructed (Yin et al., 2007).

Several strains of *E. coli* which are protease-deficient (e.g. BL21), or facilitate cytoplasmic disulphide bond formation (e.g. Origami) have been developed. Similarly, strains such as Rosetta, which supply additional tRNAs for rarely used codons to facilitate the expression of eukaryotic proteins have also been employed (Terpe, 2006). The most widely used promoters for high level expression of genes in *E. coli* include the isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible *τac*, *τrc* and *lac* promoters, thermally induced bacteriophage lambda promoter *P_L* and the T7 promoter systems. The IPTG-inducible
hybrid promoters *tac* or *trc* are powerful and widely used for basic research. However, the cost and toxicity of IPTG have precluded its use for the large-scale production of human therapeutic proteins (Makrides, 1996). In addition, the *lac, tac, trc* and T7 promoter systems are regulated by catabolite repression (Terpe, 2006). Other promoters that have been employed for overexpression include the *tetA* promoter, the cold-inducible *cspA* promoter (Makrides, 1996) and the L-arabinose or L-rhamnose-inducible *P_{BAD} or rhaP_{BAD}* promoters (Terpe, 2006).

However, *E. coli* expressions systems are being overlooked as tools to produce recombinant products due to certain disadvantages. The outer membrane of *E. coli* contains lipopolysaccharides (LPS or endotoxins), which are pyrogenic to humans and animals. In commercial applications, *E. coli* is not ideal due to its inability to secrete proteins into the surrounding medium since they localize proteins within the cells, where they aggregate, resulting in the formation of inclusion bodies (Zweers et al., 2008). The formation of protein aggregates within the cell imposes metabolic burdens on the host and also complicates downstream processing which result in poor yield of proteins (Li et al., 2004). Another drawback of *E. coli* is the limited ability to facilitate extensive disulphide bond formation (Makrides, 1996). Disulfide bonds are crucial for the activity and stability of many proteins of biotechnological or pharmaceutical interest.

### 1.1.2. Bacillus expression systems

Gram-positive bacteria such as *Bacillus* species have a long history of safe commercial application in the food, pharmaceutical and agricultural industries. *Bacilli* account for 60% of the commercially available proteins synthesized on an economical scale. The majority of these proteins are homologous proteins that are naturally secreted into the growth medium, such as alkaline proteases and amylases (Westers et al., 2004). Certain species from the genus are applied in the development of expression systems for recombinant protein production (Schallmey et al., 2004). Among them, the soil bacterium, *Bacillus subtilis* has been developed as an attractive host because of several reasons. Like most of its closest relatives e.g. *B. licheniformis, B. subtilis* is non-pathogenic and is considered as a GRAS (generally regarded as safe) organism by the US Food and Drug Administration. Its excellent fermentation capacities that are equal to, or if not, better than *E. coli*, have made it an attractive cell factory. It is capable of producing and secreting gram/litre quantities of functional extracellular proteins directly into the culture medium thereby facilitating downstream processing (Schallmey et al., 2004). Moreover, the absence of significant bias
in its codon usage allows the production of a wide variety of proteins. The protein disulfide isomerase (thiol disulfide oxido-reductase) has been shown to assist in the folding pathway of disulfide-containing proteins both in vitro and in vivo (Noiva, 1994). Genes encoding four proteins with similarities to thiol disulfide oxido-reductases (BdbA, BdbB, BdbC, BdbD) have been identified in B. subtilis (Tjalsma et al., 2004). Furthermore, not all Gram-positive bacteria seem to be equally capable in catalyzing the formation of disulfide bridges. Therefore, correct choice of the Gram-positive host organism is crucial, if the desired heterologous protein requires disulfide bonds for activity. Furthermore, the sequencing projects of a number of Bacillus species, Bacillus subtilis (Kunst et al., 1997) Bacillus licheniformis (Rey et al., 2004); Bacillus cereus (Ivanova et al., 2003), etc., have made a large body of information concerning transcription, translation, protein folding, secretion mechanisms and genetic manipulation available.

Several of the bottlenecks that were initially presented by the use of Bacillus species for the production of commercial proteins have been largely overcome. The prolific production of proteases inherent in Bacilli, which rendered heterologous proteins vulnerable and therefore minimized their yield, as well as their ability to form spores when environmental conditions become unfavourable have been addressed by the construction of protease deficient and asporogenous strains. Moreover, the phenomenon of plasmid instability which has been previously reported for Bacilli is almost non-existent due to the introduction of plasmids derived from pTA1060 (Uozumi et al., 1980), or those from pAMβ1 (Janniere, 1990) and pBS72 (Titok et al., 2003). Successful strategies for engineering of B. subtilis to improve protein production include the knockout of extracellular and/or intracellular proteases, overexpression of chaperones and folding catalysts, overexpression of components of the secretion machinery, and/or modification of the cell wall microenvironment. Besides engineering the host, the expression system used to produce the protein can be modified in order to improve production and/or secretion, for example by the use of strong or inducible promoters. Another strategy is to modify the protein that is being produced itself, for example by selecting an optimal signal peptide or by rendering the protein less sensitive to degradation through site-specific mutagenesis (Zweers et al., 2008). The latter protein modification approach has the disadvantage that it can affect the functionality and folding of the protein.
1.1.3. Expression systems used in *Bacillus subtilis*

1.1.3.1. Plasmid-based systems

Plasmids are important tools for studying bacterial functions and protein expression. For convenience purposes, the majority of the *Bacillus* plasmids are shuttle vectors that can replicate in two dissimilar organisms. Such vectors are usually furnished with a selection marker for selective pressure; the expression cassette and the elements for maintenance in the host cell. Since natural plasmids in *B. subtilis* are usually cryptic, molecular cloning in this organism has mainly been carried out with high-copy-number plasmids from *Staphylococcus aureus*, such as pC194 (Cm\(^\text{r}\)) (Ehrlich, 1977), pE194 (Em\(^\text{r}\)) and pUB 110 (Km\(^\text{r}\)) (Gryczan and Dubnau, 1978). While these plasmids replicate stably in *B. subtilis*, addition of recombinant DNA can confer mainly structural and sometimes segregational instability, too (Ehrlich *et al.*, 1986; Bron and Luxen, 1985; Bron *et al.*, 1988). The molecular basis for the structural instability has been related to their rolling circle mode of replication. The production of single-stranded DNA as an intermediate, and short direct repeats within this single-stranded DNA may lead to the deletion of one of the two repeats and the intervening DNA. Cloning vectors without structural instability have already been described, such as those based on the natural plasmids pAMβ1 and pBS72 (Janniere, 1990; Titok *et al.*, 2003) which replicate via the θ-mode as well as *B. subtilis* pTA1060-derived plasmids (Uozumi *et al.*, 1980; Bron *et al.*, 1987). Although pTA1060-based vectors replicate via a single stranded DNA-generating rolling-circle mechanism, they are far superior to the commonly used *Staphylococcus aureus* vectors with regard to cloning efficiencies and structural and segregational stability. The reason for their stability has been suggested to be their low copy number, 5 per chromosome, as compared to 20-50 for staphylococcal plasmids, which results in low amounts of ssDNA replication intermediates and replication intermediates of high molecular weight DNA that cause plasmid instability. The introduction of the *ori*- in plasmids like pHPS9, a pTA1060 derivative, reduces even further, the production of ssDNA thereby improving the stability properties of this plasmid (Haima *et al.*, 1990).

1.1.3.2. Chromosomal integration-based systems

The structural instability of autonomously replicating plasmids led to the development of vectors that integrate into chromosome. Several *Bacillus* integrative vectors have been extensively applied in the development of expression and secretion systems for both heterologous and homologous protein expression with efficient results (Meima *et al.*, 1990).
Examples of integrative vectors that have been used to date include pMUTIN (Vagner et al., 1998) which integrates by a single-cross over event, and vectors based on the amyE and thrC locus of the host cell (Kim et al., 1996; Guerout-Fleury et al., 1996), that integrate by double-crossover events.

1.2. Salient features of an expression secretion vector

Expression systems are comprised of host cells and genetic elements, such as promoters, plasmids and genes. The construction of an expression plasmid requires several elements whose configuration must be carefully considered to ensure the highest levels of protein. The essential and basic features of an expression vector are shown in Fig. 1.

![Fig. 1. A schematic representation of the salient features of a Bacillus expression-secretion vector](adapted from Makrides, 1996). The promoter (P) consists of the -35 and -10 sequences separated by a 17-base spacer. The arrow indicates the direction of transcription. The RBS consists of the Shine Dalgarno (SD) sequence 5′ AAAGGAGG 3′ followed by an A+T-rich translational spacer that has an optimal length of 7-9 bps. The SD sequence interacts with the 3′ end of the 16S rRNA during translational initiation, as shown. The three start codons (AUG, UUG, GUG) and stop codons (UAA, UGA, UAG), along with the frequency of their usage in B. subtilis (Rocha et al., 1999) are shown. The regulatory gene (R) encodes a repressor that may be present on the vector itself or may be integrated in the host chromosome, and it modulates the activity of the promoter. The transcription terminator (TT) serves to stabilize the mRNA and the vector. In addition, an antibiotic resistance gene, e.g., for erythromycin (Em), facilitates phenotypic selection of the vector, and the origin of replication (Ori) determines the vector copy number. The various features are not drawn to scale.

1.2.1. Promoters used for expression in Bacillus subtilis

One of the most important strategies for constructing an expression system is the selection of a suitable promoter and its regulators. Strong promoters are usually applied to improve the level of gene expression (Keasling, 1999). The strength of the promoter refers to the frequency of the promoter to initiate transcription which is dependent on the presence of conserved consensus sequences. Moreover, the use of a strong promoter that produces large amount of mRNA can, to a certain extent, compensate for mRNA instability, poor
translation efficiency or an unstable protein (Sawers and Jarsch, 1996). The suitability of promoters for high-level gene expression is controlled by several criteria. First, the promoter needs to be strong and at the same time, the promoter should also be compatible with the target gene so that a high level of expression can be attained (Provvedi et al., 2005). Second, the promoter should exhibit a minimal level of basal transcription; a highly repressible promoter is particularly important for cases in which the protein is toxic or detrimental to the growth and development of the host cell. Third, an ideal promoter should also be capable of induction in a simple and cost-effective manner, and the extent to which the promoter can be induced (the induction ratio) is also important (Sawers and Jarsch, 1996).

One of the main problems associated with the delay in the development of stable, efficient Bacillus expression systems was the lack of well-characterized, strong controllable promoters. With the exception of promoters derived from Bacillus species and other Gram-positive bacteria, majority of foreign promoters are not well utilized in Bacillus expression systems (Patek et al., 2003). It is therefore imperative that well-characterized promoters compatible with Bacillus expression systems are employed to drive the expression of both heterologous and endogenous genes.

Most of the promoters described require trans- and cis-acting regulatory factors for optimal function. Many of the cis elements that participate in controlling gene expression are a considerable distance away from the actual RNA polymerase binding site and such remote control regions are usually not utilized in plasmid constructs. In a high-copy number plasmid there is a risk of saturating the limited amount of endogenous regulatory proteins, which might result in a high basal activity of the promoter. Similarly, the titering out of endogenous positive regulatory factors affects the performance of a promoter when present in a high-copy plasmid. Therefore, regulation of a given promoter on a high-copy number plasmid can be considerably different from that which takes place in a single copy on the chromosome (Stader, 1995). Thus, not only is the promoter an important factor, but it is also necessary to consider regulatory factors when designing and using expression systems for Bacillus. However, the problems encountered when using high-copy number plasmids can be circumvented by using low-copy number plasmids and/or chromosomal expression systems (Keasling, 1999). Some of the promoters that have been utilized in Bacillus expression systems are shown in Table 1.
<table>
<thead>
<tr>
<th>Promoter type</th>
<th>Examples</th>
<th>Features</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporally regulated</td>
<td><em>amy, npr, apr</em></td>
<td>Most widely used in several expression-secretion systems, moderately inducible stationary phase promoters, not regulated, subject to glucose repression</td>
<td>Palva et al 1982, Wells et al., 1983, Honjo et al., 1984</td>
</tr>
<tr>
<td>Constitutive</td>
<td><em>P43, P59</em></td>
<td>Expression during vegetative phase results in less degradation by proteases produced in stationary phase in <em>Bacillus</em>, does not require expensive inducers High-level expression exerts metabolic pressure on the cells, strong transcriptional read-through affects stability of plasmids</td>
<td>Wang and Doi, 1984, Ye et al., 1999, Kim et al., 2000, Brockmeier et al., 2006</td>
</tr>
<tr>
<td>Artificial/hybrid</td>
<td><em>Pspac,</em></td>
<td>Fusion of phage SPO1 promoter and <em>E. coli lac</em> operator, IPTG-inducible Not tightly regulated, expensive induction, repressed by glucose</td>
<td>Yansura and Henner, 1984, Bhavsar et al., 2001</td>
</tr>
<tr>
<td>Sporulation</td>
<td><em>Btl, BtII</em></td>
<td>Induced during sporulation</td>
<td>Park et al., 1998</td>
</tr>
<tr>
<td>Phosphate starvation-inducible</td>
<td><em>phoA, pst</em></td>
<td>Strong, tightly regulated, highly induced by phosphate starvation, not repressed by glucose, regulated by PhoPR two component system</td>
<td>Lee et al., 1991, Kerovuo et al., 2000</td>
</tr>
<tr>
<td>Citrate-inducible</td>
<td><em>citM</em></td>
<td>Positively regulated by citST two-component system, induced by citrate, cre site mutation renders it insensitive to catabolite repression</td>
<td>Yamamoto et al., 2000, Fukushima et al., 2002</td>
</tr>
<tr>
<td>Sugar-inducible</td>
<td><em>Xyl, Sac</em></td>
<td>High levels of induction with relatively cheap substrates (xylose and sucrose), low basal expression, repressed by glucose</td>
<td>Kim et al., 1996, Bhavsar et al., 2001, Heng et al., 2005, Ye et al., 1999</td>
</tr>
<tr>
<td>Heat-inducible</td>
<td><em>T7, phi105</em></td>
<td>Tight regulation, relatively lower maximum induction</td>
<td>Conrad et al., 1996, Liu et al., 2004</td>
</tr>
<tr>
<td>Antibiotic-inducible</td>
<td><em>Tet</em></td>
<td>High induction with tetracycline, relatively lower maximum induction</td>
<td>Geissendorfer and Hillen, 1990</td>
</tr>
</tbody>
</table>
Of special interest in this study, is the phosphate starvation promoter of the *pst* operon. A brief description of the operon, its regulation in relation to the general phosphate deprivation response in *Bacillus* and its application in biotechnology is presented in the following section.

**1.2.1.1. Phosphate starvation-inducible promoters**

Many of the expression systems described so far, require inducers such as xylose or IPTG which become economically unattractive in large scale or repeat experiments. Therefore, it is desirable to have a vector in which gene expression can be controlled by the level of nutrients such as phosphate. Phosphate is one of the most important nutrients for all organisms. It is found in lipids, nucleic acids, proteins and sugars, and is involved in many biochemical reactions that rely on transfer of phosphoryl groups (Lamarche *et al.*, 2008). Soil-dwelling bacteria, including *B. subtilis*, have evolved complex regulatory systems for utilizing this limited nutrient which is often present at levels 2-3 orders of magnitude lower than those of other required ions (Hulett, 1996).

![Fig. 2. Schematic representation of the phosphate starvation response in *B. subtilis*. (Adapted from Hulett, F.M., 2002)](image-url)

The *pst* operon is a member of the *pho* regulon that is induced by PhoP–P. PhoPR is the main regulator. It regulates the expression of the *pho* regulon genes to allow the cell to use alternative organic phosphate sources. ResD is required for full expression of *pho* genes. Spo0A–P represses the transcription of *phoPR* through inhibiting the expression of AbrB and ResDE. Two component systems are labelled as histidine kinase (HK) and response regulator (RR). Positive regulation is indicated by a plus sign (+) and repression by a minus sign (−).
Phosphate (pho) regulons of bacteria include all genes whose expression rates respond to phosphate limitation and are controlled directly by the two-component regulatory system PhoP-PhoR (Fig. 2). The pho regulon of *B. subtilis* is comprised of at least eight operons, and five monocistronic genes (Allenby *et al.*, 2005). Among these is the *pst* operon (Qi *et al.*, 1997) which is induced when inorganic phosphate (Pi) concentrations in the extracellular environment fall below 0.1 mM (Allenby *et al.*, 2005).

During phosphate deprivation, *B. subtilis* ceases to synthesize the major cell wall component, teichoic acid, which is a glycerol or ribitol phosphate polymer, and replaces it with teichuronic acid, another anionic phosphate-free polymer. The cell also secretes multiple alkaline phosphatases (APases), alkaline phosphodiesterases and other hydrolytic enzymes to recycle the available phosphorus from teichoic acid. The inorganic phosphate (Pi) produced in this process is then transferred into the cell by an efficient high affinity phosphate transport system viz., the *pst* operon. It is composed of five genes, *pstS*, *pstC*, *pstA*, *pstB1* and *pstB2*. These genes encode proteins similar to ABC transporters (Ames, 1986), which are composed of a substrate binding periplasmic lipoprotein (PstS), two integral inner membrane permease-type proteins that facilitate the passage of Pi across the membrane (PstC and PstA) and the cytoplasmic ATP-binding proteins (PstB1 and PstB2) that energize this transport (Qi *et al.*, 1997).

During Pi limitation, the membrane-bound sensor histidine kinase PhoR turns on the pho regulon genes by phosphorylating PhoP, a response regulator. The activated PhoP~P in turn activates transcription by binding specifically to pho box sequences: direct repeats of TT(A/T/C)ACA with a 5 ± 2 bp spacer (Eder *et al.*, 1999). For efficient binding, four TT(A/T/C)ACA-like sequences with an 11-bp periodicity are required. The pho box is part of the promoter of each pho regulon gene, including the *pst* operon. The PhoP-binding sites are located on the coding strand of the promoter region of genes induced by PhoP~P and on the noncoding strand of the promoter regions of genes repressed by PhoP~P (Hulett, 2002; Pragai *et al.*, 2004).

The response to phosphate starvation is rapid, and maximum expression of genes is achieved in as little as 30-60 minutes (Schweizer and Boos, 1985). Furthermore, the genes of the pho regulon are expressed at very high levels, suggesting very strong promoter activity. Thus, phosphate regulated promoters are attractive for recombinant protein expression. The other advantages of expressing protein during phosphate starvation when cells are in the stationary phase is that many components of the secretion apparatus are
maximally expressed in the early stationary growth phase (Vitikainen et al., 2001). Thus, proteins meant for secretory production are better expressed during this period. Also, the cells will have reached a good cell density which would increase the yield of the target protein. Conversely, protein overexpression during growth can impose a metabolic burden on the cells and ultimately decrease the cell density and thereby, the yield of the target protein (Donovan et al., 1996).

Several phosphate regulated promoters such as those of phoA and ugpBAECQ operon in E. coli (Oka et al., 1985; Carter et al., 1992; Su et al., 1990), acid phosphatase pho5 promoter in yeast (Kramer et al., 1984) have been used for successful production of recombinant proteins. Proteomic analyses of B. subtilis (Eymann et al., 1996; Antelmann et al., 2000) indicate that the most abundant protein induced in response to phosphate starvation is PstS, indicating the presence of a very strong promoter. In addition Qi et al., 1997 reported a very high induction level of 5000 for the pst promoter. Kerovuo et al., 2000 have successfully applied this promoter for recombinant phytase expression. The fact that the pst promoter is tightly regulated and not repressed by glucose also makes it an ideal promoter for high level protein production.

1.2.2. Translational signals

For efficient expression of foreign or native genes in any organism it is indispensable to use optimal translational signals in addition to strong promoters. These include the ribosome binding site (RBS), initiation codon, the spacer between the Shine Dalgarno (SD) sequence in the RBS and the initiation codon, and the secondary structure around the translation initiation site.

1.2.2.1. Ribosome binding site

Since most heterologous and homologous genes are expressed in Bacillus species by direct fusion of the gene of interest to the transcription and translational signals of a native Bacillus species gene, not much effort has been invested in optimizing the RBS conditions. Unlike in E. coli, the translation machineries of Bacillus species are quite specific and require a "stringent" RBS for efficient translation initiation. The S1 protein which helps to align the SD sequence with the free 3' end of 16S rRNA of the ribosome complex is absent in Bacillus, therefore magnifying the importance of a strong RBS. The consensus SD sequence in Bacillus is AAAGGAGG (Rocha et al., 1999).
In addition, the RBS plays an important role in mRNA stability. Mutations in the RBS that interfered with ribosome binding significantly destabilized mRNA, whereas those that improved ribosome binding prolonged mRNA longevity, indicating that RBS occupancy by ribosomes helps to protect mRNAs from ribonuclease attack (Deana and Belasco, 2005).

1.2.2.2. Initiation codons
AUG is the most commonly used initiation codon followed by UUG and then GUG (Vellanoweth and Rabinowitz 1992) in Bacillus, and therefore it is advisable to use AUG as an initiation codon for heterologous gene expression in Bacillus species. The employment of AUG as the initiation codon simplifies hybrid gene construction since most prokaryotic and eukaryotic genes initiate with the AUG codon (Ganoza et al., 1987).

1.2.2.3. Spacer region
The sequence (spacer region) between SD sequence and initiation codon ranges between 5 to 11 nucleotides in Bacillus with an optimal spacing of 7-9 bp (Vellanoweth and Rabinowitz, 1992). The sequence of this spacer region is A-rich and C-poor to reduce the formation of secondary structure in the mRNA (Rocha et al., 1999). Occlusion of the translation initiation region by stem-loop structure can affect gene expression (Makrides, 1996).

1.2.3. Other factors influencing increased expression
Although much less attention has been given to transcription termination than to transcription initiation in heterologous and homologous protein expression, efficient transcription terminators are indispensable elements of expression vectors because they serve several important functions. There is well documented evidence indicating that transcription read-through from strong promoters into the replication region can destabilize plasmids (Stueber and Bujard, 1982; Ehrlich et al., 1986; Mountain et al., 1984; Mountain, 1989; Gentz et al., 1981) as a result of overproduction of the ROP protein involved in control of the plasmid copy number. Thus proper placement of transcription terminators does not only increase plasmid stability, but also decreases metabolic load by reducing the transcription (and translation) of other plasmid-encoded genes whose high level expression is not essential for the maintenance and expression of the gene of interest (Mountain, 1989). A number of Escherichia coli ρ-independent terminators were shown to be able to
function in *Bacillus* species (Peschke *et al*., 1985). The *rrn*B (an operon for constitutively expressed rRNA genes) terminators were utilized in the construction of the expression vector pWE1, which was used for the expression and secretion of human atrial natriuretic α-factor (Wang *et al*., 1988). Similarly, the native *B. licheniformis* α-amylase gene terminator has been used successfully for the construction of an efficient *B. subtilis* shuttle vector (Li *et al*., 2006).

It has also been reported that certain terminator sequences can increase the level of expression by increasing the mRNA stability, probably due to the formation of a stable stem-loop structure that protects the mRNA from exonuclease attack. mRNA stability has a huge impact on prokaryotic gene expression. The 5′ and 3′ UTRs of mRNAs seem to be important mediators of mRNA stability. A 5′ secondary structure and ribosome binding, but not actual translation were shown to be important for the stability of the *B. subtilis* aprE mRNA (Hambraeus *et al*., 2002). The presence of a strong ribosome-binding site (RBS) in the *B. subtilis* gsiB mRNA was critical for its long half-life (Jurgen *et al*., 1998).

### 1.2.4. Protein secretion in *Bacillus*

One of the reasons for the extensive use of *Bacillus* species in industry is their enormous secretion potential of proteins (Behnke, 1992). In contrast to Gram-negative bacteria with an additional outer membrane, secretory proteins produced by Gram-positive bacteria only need to traverse a single cytoplasmic membrane to enter the extracellular environment. Therefore, it is advantageous to use Gram-positive bacteria for secretory protein production since exported proteins usually maintain their native conformation, whereas, intracellular production in many cases results in aggregation of the proteins (inclusion bodies). Another advantage is that secretion facilitates and simplifies downstream processing (Bron *et al*., 1998). The vast majority of the extracytosolic proteins are transported in an unfolded state by the Sec pathway. In contrast, completely folded proteins are exported by the Tat systems (Gellissen, 2002). Despite its high secretion capacity for homologous proteins, the protein secretion machinery of *Bacillus* has certain limitations, which become evident when it is challenged with heterologous proteins - eukaryotic proteins with multiple disulfide bonds in particular. Bottlenecks in the secretion pathway of *B. subtilis* that have been reported so far relate to different stages in the secretion process (Bolhuis *et al*., 1999).
1.2.4.1. Signal peptides

Proteins destined for secretion are synthesized as precursors equipped with a signal peptide (15-30 amino acids long) that is proteolytically removed by signal peptidases during or shortly after translocation. For the production of heterologous proteins in *B. subtilis*, it is necessary to use a signal peptide that directs the protein very efficiently to the translocase and that is cleaved efficiently by the signal peptidases. In most approaches, the signal sequences of α-amylases and proteases have been utilized. The combination of the signal peptide and mature protein also appears to affect the efficiency of export in both *Bacillus* species and *E. coli* (Nakamura *et al.*, 1989). A signal peptide that can support efficient export for one protein need not do so for another (Gellissen, 2002).

1.2.4.2. Degradation by cell-associated and secreted proteases

*Bacillus* species produce large amounts of secreted proteases mainly neutral proteases (or metalloproteases) and alkaline proteases (or serine proteases). In addition, several so-called minority proteases are also exported by *Bacillus* species (Simonen and Palva, 1993) (Fig. 3). Proteolytic degradation of heterologous proteins was one of the major obstacles in developing *Bacillus* into efficient secretory expression systems suitable for biotechnological applications. The development of protease-deficient *B. subtilis* hosts (Kawamura and Doi, 1984; Fahnestock and Fisher, 1989; Wu *et al.*, 1991, Wu *et al.*, 2002), and/or the use of *Bacillus* species, such as *B. brevis* which has very low extracellular protease activity have significantly improved recombinant protein production in *Bacillus* hosts. Another alternative to avoid degradation is to produce the foreign proteins in the exponential growth phase, when relatively small amounts of proteases are secreted (Edelman *et al.*, 1988; Simonen and Palva, 1993). Therefore chemostats, in which the cells are kept continuously in the exponential phase, are essential for producing these proteins for production purposes. In addition to soluble proteases in the supernatant, the membrane-cell wall interface contains numerous proteases, such as the *wprA*-encoded CWBP52 (Stephenson and Harwood, 1998) and homologues of the *E. coli* HtrA and Tsp proteases (Bolhuis *et al.*, 1999), which appear to play a role in the quality control of secreted proteins. A rapid degradation of newly secreted proteins is observed by cell wall-associated proteases, if the intrinsic folding of the target protein is inefficient (Jacobs *et al.*, 1993; Meens *et al.*, 1997).
Recent studies suggest that, in addition to at least six secreted proteases, proteases residing at the membrane-cell-wall interface can be particularly problematic in this respect. For example, it was shown that the wall-bound serine protease CWBP52, specified by the \textit{wprA} gene, is active at the site of preprotein translocation, and that CWBP52 depletion results in increased yields of secreted \(\alpha\)-amylase (Stephenson and Harwood, 1998). Thus, proteases of the latter type are likely to be involved in the quality control of secreted proteins by removing incorrectly folded proteins in order to prevent jamming of the secretory pathway.

\textbf{1.2.4.3. Overloading of the cell secretion machinery}

Multicopy plasmids expressing \textit{B. stearothermophilus} \(\alpha\)-amylase in \textit{B. subtilis} were unstable until a spontaneous copy-number mutant (1/10 of the original) was obtained (Diderichsen and Christiansen, 1988), showing that overproduction of a secretory protein is deleterious to the bacterium and strongly selected against. When \(\alpha\)-amylase was produced...
in small amounts, it was processed very rapidly, but when the production was increased by increasing the gene dosage, processing became slower and cell-associated precursors started to accumulate. At a copy number of about 10 to 20, the secretion machinery became saturated and an increase of gene copies no longer notably increased the amount of enzyme secreted. Addition of the enhancer mutation degU9 (DegU is the response regulator of the two component system DegS-DegU that positively regulates exoprotease production) to the strain carrying the α-amylase gene in a multicopy plasmid resulted in a very rapid loss of the plasmid from the culture (Vehmaanpera and Korhola, 1986), suggesting that the secretion capacity of the strain had been exceeded. *Staphylococcus aureus* protein A provides another example of overloading in *Bacillus* cells. Protein A is both efficiently synthesized and secreted in *B. subtilis*, but it could not be expressed in multicopy plasmids with its own promoter. Expression and secretion occurred only when it was integrated into the chromosome as a single copy or when the promoter was changed to a weaker one (Fahnestock *et al.*, 1986).

1.3. Carbon catabolite repression

1.3.1. Overview

When exposed to more than one carbon source, bacteria assimilate the carbohydrate that provides the most rapid and successful growth, first. For many heterotrophic bacteria, glucose is the preferred carbon source, and in its presence, the genes required for the utilization of secondary carbon sources are not expressed. This phenomenon, initially described as the glucose effect is found in virtually all living organisms (Lorca *et al.*, 2005; Deutscher, 2008; Singh *et al.*, 2008). Magasanik launched the name ‘catabolite repression’ after it became clear that not only glucose, but any compound that can serve as a source of intermediary metabolites (catabolites), could repress the synthesis of proteins involved in the uptake and metabolism of less efficient substrates (Ullmann, 1996). Thus carbon catabolite repression (CCR) can be defined as a regulatory mechanism by which the expression of genes required for the transport and metabolism of less favourable carbon sources is prevented by the presence of preferred carbon sources such as glucose.

1.3.2. The mechanism

The regulatory mechanisms underlying CCR have since been intensively studied mostly in the model organisms *E. coli* and *B. subtilis*. In both species, CCR involves a global
mechanism and several operon specific regulatory mechanisms, such as inducer exclusion and induction prevention. In *E. coli* and *B. subtilis*, the regulatory outcomes of the global mechanisms of CCR are similar: the genes that enable the use of secondary carbon sources are not expressed when glucose, or another preferred carbon source, is available. However, the molecular mechanisms by which this global regulation is achieved are completely different in these organisms. In *E. coli*, CCR is mediated by the prevention of transcriptional activation of catabolic genes in the presence of glucose. By contrast, in *B. subtilis*, CCR is mediated by negative regulation through a repressor protein in the presence of glucose. Although the mechanisms of CCR differ in these two organisms, the phosphoenolpyruvate–carbohydrate phosphotransferase system (PTS) is important in both organisms in the signal-transduction pathways that lead to CCR (Gorke and Stulke, 2008). However, some bacteria such as *Rhizobium, Bradyrhizobium, Sinorhizobium* and *Pseudomonas*, prefer acetate or tricarboxylic acid (TCA) cycle intermediates such as succinate over glucose or fructose and CCR in these organisms is possibly mediated by additional mechanisms (Deutscher, 2006).

1.3.3. *The phosphoenolpyruvate-dependent carbohydrate phosphotransferase system*

In bacteria, the phosphoenolpyruvate-dependent carbohydrate phosphotransferase system (PTS) is the main carbohydrate uptake system and, in addition, plays an important role in the induction and carbon catabolite repression (CCR) of numerous catabolic genes and operons. During carbohydrate uptake, the phospho-carrier protein HPr is phosphorylated at the catalytic histidine residue by enzyme EI at the expense of PEP. The phosphoryl group is then serially transferred to the cytoplasmic protein EIIA (*E. coli*) or part of the EIIABC complex (*B. subtilis*) to EIIB and finally to the integral membrane transporter domain EIIC. The carbohydrate molecule is transported into the cell through EIIC and at the same time phosphorylated by EIIB (Fig. 4).
Fig. 4. The phosphoenolpyruvate-dependent sugar phospho-transferase system (PTS). The PTS is the main sugar uptake system that couples the nutritional conditions and the metabolic state of the cell to regulate the expression of numerous catabolic genes and operons (Adapted from Gorke and Stulke, 2008).

This dynamic modulation of the phosphorylation states of PTS proteins in response to nutritional conditions and the metabolic state of the cell provides the basis for PTS-mediated signalling and regulation (Gorke and Stulke, 2008).

1.3.4. CCR in Escherichia coli

In *E. coli*, the EIIA domain of the glucose transporter (EIIA^{Glc}) is the central processing unit in CCR. Regulation of CCR is brought about by the modulation of the phosphorylation state of EIIA^{Glc}. In the absence of glucose, EIIA^{Glc} is phosphorylated and catalyzes the transfer of its phosphoryl group to adenylate cyclase (AC), leading to its activation. This results in the production of cyclic AMP (cAMP), which binds to the cAMP receptor (or regulatory) protein (CRP), also known as catabolite activator protein (CAP). The resulting cAMP-CRP complex binds a specific DNA sequence near the promoter of the catabolic gene or operon, where it can contact RNA polymerase at the promoter site, leading to increased rates of transcription initiation. In the presence of glucose, EIIA^{Glc} is mainly unphosphorylated, since it donates the phosphoryl group to the incoming glucose. Consequently, adenylate cyclase is inactive, resulting in the loss of transcription activation and repression of the catabolic gene or operon. In addition to the regulation of transcription, unphosphorylated EIIA^{Glc} binds and inactivates metabolic enzymes and transporters of secondary carbon sources, such as GlpK, LacY and other proteins. This mechanism is called inducer exclusion because it prevents the intracellular formation of the inducer of the respective catabolic system (Gorke and Stulke, 2008) (Fig. 5a).
1.3.5. **CCR in B. subtilis**

In gram-positive bacteria, the main signalling intermediate is HPr and not EIIA\textsubscript{Glc} (Warner and Lolkema, 2003, Lorca \textit{et al.}, 2005). HPr can be phosphorylated either at histidine residue (His-15) by E1 in the presence of PEP or at a serine residue (Ser-46) by the bifunctional Hpr kinase/phosphorylase in the presence of ATP. Hpr-(His-P) is a poor substrate for HPr kinase while HPr-(Ser-P) is a poor substrate for E1. In the presence of a preferred carbon source like glucose, HPr is mostly in the unphosphorylated state, since it transfers its phosphoryl group from His-15 to EII of the PTS to complete the uptake of the sugar. When glucose is taken up, it is metabolized via glycolysis leading to the formation of glycolytic intermediates, such as fructose-1,6-bisphosphate (FBP) and glucose-6-phosphate (Glu-6-P). These glycolytic intermediates activate HPr kinase that catalyzes the phosphorylation of HPr at the serine residue (Ser46 in \textit{B. subtilis} HPr) in an ATP-dependent manner. Subsequently, seryl-phosphorylated HPr interacts with the catabolite control protein CcpA and binds to a \textit{cis-}acting catabolite repressive element (\textit{cre}) within or downstream of the promoter region of the target gene and prevents transcription of the gene by RNA polymerase (\textit{Fig. 5b}). Although Hpr-(Ser-P), FBP and Glu-6-P may stimulate the affinity of CcpA for \textit{cre}s, they are not required for CcpA binding to \textit{cre}s of \textit{amyE, gnt, xyl, hut, lev,} and \textit{ackA in vitro} (Kim \textit{et al.}, 1998).
Fig. 5. Schematic representation of CRP-dependent and CcpA-dependent carbon catabolite repression pathways in *E. coli* (a) and *B. subtilis* (b) respectively. (from Gorke and Stulke, 2008). In *E. coli* the degree of phosphorylation of EIIA determines the activity of adenylate cyclase and, consequently, the concentration of cAMP in the cell. Binding of cAMP to CRP results in a complex that stimulates transcription of target genes (positive regulation). In *B. subtilis* (B), fructose-1,6-phosphate (FBP) produced from glucose-6-phosphate in glycolysis activates HPrK that phosphorylates HPr or Crh at the regulatory-site serine at the expense of ATP or PPI. Binding of HPr-Ser-P or Crh-Ser-P to CcpA results in a complex that inhibits the transcription of target genes (negative regulation). The degree of phosphorylation of HPr at His-15 and Ser-46 is used as a signal for the availability of PTS substrates and the metabolic activity of the cell. The absence of the phosphate from His-15 indicates the presence of a PTS sugar, and the presence of a phosphate at Ser-46 reflects a high glycolytic activity. These signals are recognized by the carbon catabolite control protein CcpA, which is the central regulatory protein for carbon catabolite repression (CCR) in Gram-positive bacteria (Kraus *et al.*, 1998).
1.4. Significance of carbon catabolite repression and attempts to surmount it with special reference to amylases

Catabolite repression (CR) controls many cellular and developmental processes including, sporulation, enzyme transport systems, and extracellular enzyme synthesis in *Bacillus* species. Significant in this aspect, is the production of extracellular enzymes in industrial fermentations, where CR is a most important consideration for the fermentation technologist. All the commercial enzymes such as amylases of *B. licheniformis* and *B. amyloliquefaciens* and the proteases of *B. licheniformis*, *B. amyloliquefaciens* and *B. subtilis* are subject to catabolite repression (Priest and Sharp, 1989). It has been reported that amylase expression is inhibited in the presence of glucose by 10-15 fold in *B. licheniformis* and *B. subtilis* (Laoidé et al., 1989, Weickert and Chambliss, 1990).

To overcome CR, the approaches that have been used in industrial fermentations include the preparation of mutants and/or the use of slowly metabolized carbon sources such as starch, starch hydrolysates etc., as well as cloning. In most cases, the increase in fold production of the target gene has only been mediocre.

1.4.1. Use of mutants to overcome CCR

1.4.1.1. cre mutants

Extensive site-directed mutagenesis studies with the *cre* site of amylase in *B. subtilis* was carried out to identify bases important for glucose regulation. Although many mutations reduced the repression ratio 3- to 11-fold, some caused only a 2-2.9 fold increase in amylase production. Other mutations caused hyperproduction, without affecting catabolite repression (Weickert and Chambliss, 1990). On the contrary, site-directed mutagenesis of the dyadic symmetric element (DSE) present in the amylase promoter region of *Streptomyces* strain sp. TO1 still resulted in catabolite repression by glucose or glycerol (Mellouli *et al.*, 1999). However, repression was overcome when the amylase gene was cloned onto a high copy number vector, wherein the negative transcriptional regulator would have been titrated out by the DSE (Mellouli *et al.*, 2002).

1.4.1.2. CcpA mutants

Mutations of the pleiotropic transcriptional regulator CcpA in *B. subtilis* and *B. megaterium* resulted in loss of CCR in many catabolic genes and operons (Stulke and Hillen, 2000). A CcpA mutant generated by transposon insertion, showed loss of glucose
repression of amylE gene expression in B. subtilis, but the increase in amylase production was only about 4-fold in glucose. Repression by other sugars, such as sucrose and fructose, was also reduced, while glycerol still fully repressed amylase synthesis (Kenkin et al., 1991). Although the deletions or disruptions of cepA result in CCR-resistance, CepA mutant strains exhibit a severe growth defect on minimal media. This could be explained by the fact that CepA is necessary for the full expression of glycolytic enzymes and for efficient ammonium assimilation (Stulke and Hillen, 2000). Thus such mutants would not be suitable for growth on cheap, industrial carbon sources such as starch for the commercial production of enzymes.

1.4.1.3. Hpr mutants

In a B. subtilis Hpr mutant strain (ptsH1) in which Ser-46 of HPr was replaced with a nonphosphorylatable alanyl residue (ptsH1 mutation), synthesis of gluconate kinase, glucitol dehydrogenase, mannitol-1-P dehydrogenase, mannitol-specific PTS permease (Deutscher et al., 1994) and β-glucanase (Kruger et al., 1993) was relieved of CR whereas synthesis of inositol dehydrogenase was partially relieved from catabolite repression. However, a ptsH1 mutation in B. subtilis SA003 did not prevent glucose repression of the α-amylase gene suggesting that Hpr(Ser-P) may not be involved in CepA mediated repression of α-amylase (Voskuil and Chambliss, 1996).

1.4.2. Other approaches to overcome CR

Starch or starch-based agro residues such as wheat bran, being readily available, cheap substrates are used for large-scale commercial production of amylases. However, accumulation of starch degradation products such as maltose and glucose in batch cultures, ultimately results in the cessation of enzyme synthesis due to CCR. Continuous cultures to maintain non-repressible levels of preferred carbon sources have improved yield of amylases compared to batch cultivation (Emanuilova and Toda, 1984). However, conditions have to be very carefully monitored to ensure high production. It was found that with increasing growth rate, α-amylase production decreased considerably in chemostat experiments (Emanuilova and Toda, 1984; McMahon et al., 1997). Alternatively, solid state fermentation (SSF) has also been employed to overcome catabolite repression of α-amylases (Ramesh and Lonsane, 1991; Babu and Satyanarayana, 1995; Mellouli et al., 2002). However, in some cases the expression of amylase was found to be repressed by
glucose during SSF (Kunamneni et al., 2005; Gangadharan et al., 2006). Although SSF appears to be superior to submerged fermentation with regard to higher fermentation productivity, higher end-concentration of products, higher product stability, lower catabolic repression, lower demand on sterility due to the low water activity used in SSF, these advantages have been observed only at the laboratory scale (Holker et al., 2004). Scale-up of the SSF technique for viable commercial production has faced severe engineering problems due to the build-up of temperature, pH, O₂, substrate and moisture gradients (Holker et al., 2004). In addition, the diffusion of products through the solid media complicates both extraction processes (in which concentrated impurities could also be obtained) and purification steps. This contributes to an increase in recovery costs (Perez-Guerra et al., 2003).

1.4.3. Cloning of α-amylases

Alternatively, to improve the yield, several bacterial heat stable α-amylases have been expressed on multicopy plasmids in E. coli and Bacillus species (Table 2). Most of them describe the use of the native promoter and/ or signal sequence, or a heterologous bacterial amylase promoter on a plasmid for overexpression. However, the increase in production has mostly only been modest, since the amylase promoter contains the cre sequence and multi-copies of the gene on plasmids do not necessarily titrate out the regulatory factors implicated in CR (Laoide et al., 1989; Weickert & Chambliss, 1989).
Table 2. List of cloned thermostable $\alpha$-amylases in *E. coli* and *Bacillus* species

<table>
<thead>
<tr>
<th>Amylase donor</th>
<th>Host species</th>
<th>Origin of promoter and signal sequence</th>
<th>Secretion</th>
<th>Yield and/ or comments *</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>B. lich</em></td>
<td><em>B. subtilis</em></td>
<td>apr (B. subtilis)</td>
<td>+</td>
<td>0.5-1 g/l</td>
<td>Sloma <em>et al.</em>, 1988</td>
</tr>
<tr>
<td><em>B. lich</em></td>
<td><em>B. subtilis</em></td>
<td>amy P and SS (B. amylo)</td>
<td>&gt; 95 %</td>
<td>Several 100X B. lich</td>
<td>Sibakov 1986</td>
</tr>
<tr>
<td><em>B. lich</em></td>
<td><em>B. brevis</em></td>
<td>intact</td>
<td>+</td>
<td>300 U/ml</td>
<td>Yamagata <em>et al.</em>, 1987</td>
</tr>
<tr>
<td><em>B. lich</em></td>
<td><em>E. coli</em></td>
<td>T7 promoter, amy SS (B. lich)</td>
<td>+</td>
<td>0.235 $\mu$mol/min/mg protein</td>
<td>Shahhoseini <em>et al.</em>, 2003</td>
</tr>
<tr>
<td><em>B. lich</em></td>
<td><em>B. brevis</em></td>
<td>mwp (B. brevis)</td>
<td>+</td>
<td>44,000 U/ml high copy number vector</td>
<td>Yamagata <em>et al.</em>, 1987</td>
</tr>
<tr>
<td><em>B. lich</em></td>
<td><em>B. subtilis</em></td>
<td>P43, sacB SS</td>
<td>+</td>
<td>4700 U/ml</td>
<td>Liu <em>et al.</em>, 2008</td>
</tr>
<tr>
<td><em>B. stearo</em></td>
<td><em>B. brevis</em></td>
<td>intact</td>
<td>+</td>
<td>0.5 g/l</td>
<td>Tsukagoshi <em>et al.</em>, 1989</td>
</tr>
<tr>
<td><em>B. stearo</em></td>
<td><em>B. subtilis</em></td>
<td>intact</td>
<td>+</td>
<td>0.1 g/l</td>
<td>Tsukagoshi <em>et al.</em>, 1989</td>
</tr>
<tr>
<td><em>B. stearo</em></td>
<td><em>B. subtilis</em></td>
<td>intact</td>
<td>50 %</td>
<td>30 U/ml</td>
<td>Sen and Oriel, 1982</td>
</tr>
<tr>
<td><em>B. stearo</em></td>
<td><em>B. stearo</em></td>
<td>intact</td>
<td>+</td>
<td>5 X more than B. stearo</td>
<td>Aiba <em>et al.</em>, 1983</td>
</tr>
<tr>
<td><em>B. amylo</em></td>
<td><em>B. subtilis</em></td>
<td>amy (B. amylo)</td>
<td>&gt; 95 %</td>
<td>1-3 g/l</td>
<td>Palva 1982, Ulmanen <em>et al.</em>, 1985</td>
</tr>
<tr>
<td><em>B. amylo</em></td>
<td><em>B. subtilis</em></td>
<td>P$\alpha$ (lambda), amy SS (B. amylo)</td>
<td>+</td>
<td>13 mg/l at OD of 1</td>
<td>Breitling <em>et al.</em>, 1990</td>
</tr>
<tr>
<td><em>Pyrococcus furiosus</em></td>
<td><em>E. coli</em></td>
<td>intact</td>
<td></td>
<td>0.2 U/ml</td>
<td>Laderman <em>et al.</em>, 1993</td>
</tr>
</tbody>
</table>

*B. lich* = *B. licheniformis*, *B. stearo* = *B. stearothermohilus*, *B. amylo* = *B. amylo liquefaciens* P = promoter; SS = signal sequence, amy = $\alpha$-amylase; mwp = middle wall protein; apr = alkaline serine protease; penP = $\beta$-lactamase. Intact = the gene is expressed from its own promoter and signal sequence. *The units used to express the yield are not always comparable.
The present study
1.5. The present study: Engineering a strong, catabolite repression-free expression system for a thermophilic amylase

High level gene expression is of particular importance in industrial applications where commercially relevant production titres are required. Out of several available expression systems, *Bacillus subtilis* is an attractive host because of its high secretory capacity, non-pathogenicity and high genetic amenability. A successful *Bacillus* expression system will require a host lacking extracellular proteases, strong transcription and translation elements and nutritional conditions that allow expression of foreign genes during growth. With all the available knowledge about promoters, plasmids, mutant host strains and fermentations, an optimal production system needs to be rationally chosen as the choice of each component can be critical for the successful expression of the gene of interest. The present work is in this effort using a thermophilic *B. licheniformis* α-amylase as the protein of interest.

Thermophilic amylases account for about a quarter of the world’s enzyme consumption and have a wide variety of applications in starch processing, baking, brewing, textile, paper, detergent, processing of animal feeds, waste water treatment and pharmaceutical industries. Thermophilic enzymes are preferred in various processes because of the specificity of the reaction, reduced risk of contamination, faster reaction rates and tolerance to organic solvents and elevated temperatures often used in raw material pretreatments. Due to the increasing demand for these enzymes, there is enormous interest in enhancing enzyme production/yield, thermostability or activity.

The synthesis of α-amylase in *B. licheniformis* and *B. subtilis* is repressed 10-15 fold in the presence of glucose (Laoide *et al.*, 1989; Weickert & Chambli ss, 1990) as well as by other readily metabolized carbon sources. Consequently, carbon catabolite repression (CCR) becomes an important consideration in the large-scale commercial production of amylase, where sugars accumulate and repress enzyme production in starch containing media. Although solid state fermentation, cloning and mutagenesis of regulatory sequences have been used to improve the yield of amylases, the increase in fold production of amylase has not been significant.

A thermophilic, maltogenic and cyclodextrin-hydrolyzing α-amylase with optimum activity at 94°C and a wide pH range of 5-9.5 is produced by *B. licheniformis* MSG that was isolated from thermal hot springs at Unai in South Gujarat (M. Nair, unpublished
results). The availability of an enzyme with potential industrial application, as well as the demand for strong and better expression systems led to the conception of this project.

The objectives of this work were:

(i) To engineer an expression system to overcome catabolite repression of *B. licheniformis* α-amylase:

The expression of *B. licheniformis* α-amylase (BLA) is transcriptionally repressed, in the presence of readily metabolized carbon sources, by the CcpA protein when it binds to an operator sequence immediately upstream, but partially overlapping the structural gene. Overcoming CR by avoiding this 5' negative regulatory sequence would allow sustained expression of amylase even in the presence of accumulated sugars arising from the degradation of starch thereby improving amylase yield.

(ii) To express amylase from a strong, heterologous promoter that is induced under growth conditions without added inducers:

By cloning the amylase gene under a compatible, strongly induced and tightly regulated promoter that is not repressed by glucose, expression of α-amylase could be further enhanced. The phosphate starvation-inducible promoter, *pst* of the *B. subtilis* *pst* operon (Qi *et al.*, 1997; Kerovuo *et al.*, 2000) was chosen for these reasons, and also because the calcium-dependent amylase used in this study has a longer shelf-life in medium/buffers containing low phosphate (Nair M., unpublished results).

(iii) To analyze the repression-free expression, secretion of BLA, and influence of promoter on expression in a protease minus *B. subtilis* host:

As even, *B. licheniformis* α-amylase is subject to considerable proteolytic degradation in *B. subtilis* (Stephenson & Harwood, 1998; Jensen *et al.*, 2000), the eight protease-deficient *B. subtilis* WB800 was chosen as the host for overexpression of amylase.