Regulation and expression of the metal citrate transporter CitM of Bacillus subtilis

de Jonge - Warner, Jessica

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2002

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
General Introduction

CONTENTS

INTRODUCTION

SOLUTE TRANSPORT SYSTEMS
   ABC transporters
   Secondary transporters
   Bacterial phosphoenolpyruvate phosphotransferase system

INDUCTION OF CATABOLIC PATHWAYS
   Relief of repression
   Activator proteins
   Two-component systems
   PRD-mediated regulators

CARBON CATABOLITE REPRESSION
   Historical overview
   Carbon catabolite repression in Gram-positive bacteria
   Components of catabolite repression in Gram-positive bacteria
     HPr kinase/phosphatase
     HPr and Crh
     CcpA
     Cre site
     CcpB and CcpC

OUTLINE OF THIS THESIS
INTRODUCTION

The content of bacterial cells (cytoplasm) is surrounded by a lipid bilayer, called the cell membrane that protects the cell from outside influences and makes it possible to carefully control the composition of the cytoplasm. In Gram-positive bacteria the cytoplasm is bordered by a single phospholipid bilayer, the cell membrane, and a peptidoglycan layer, the cell wall, which supports firmness of the cell. Gram-negative bacteria possess an additional outer membrane creating a periplasmic space. The outer membrane of Gram-negatives is permeable, allowing small molecules to enter the periplasm. In contrast, the cytoplasmic membrane forms a selective barrier especially for bulky hydrophilic compounds that cannot easily pass the hydrophobic membrane. To get nutrients into and breakdown products out of the cell, membrane embedded proteins, called transport proteins, mediate the translocation of solutes across the cell membrane. Most transport proteins are very selective and only carry one or a few related compounds from one side to the other.

Not all nutrient transport systems need to be present at the same time, since usually limited numbers of carbon and nitrogen sources are present in particular habitats or growth media. To prevent unnecessary consumption of energy and matter for the synthesis of (transport) proteins bacteria have evolved mechanisms to sense the presence of substrates and to react on their availability. In this way, specific transport systems and associated catabolic enzymes are synthesized only when their substrates are present and other preferred carbon and energy sources are lacking. The synthesis of the transport systems and enzymes is regulated at the level of transcription of the coding genes and operons. In the presence of substrate the expression of the respective genes is activated by a mechanism called induction. When more than one substrate is present, bacteria decide which one to utilize first, primarily the one that provides the most energy for growth. Then, only the (transport) proteins required for metabolism of the preferred substrate are synthesized, while synthesis of the proteins required for metabolism of the less favored substrate is unwanted even though the inducer is present. This mechanism is referred to as carbon catabolite repression (CCR). Especially, regulation of the synthesis of transport proteins is essential, since this allows control ‘at the gate’.

This thesis describes the regulation of expression of the Mg-citrate uptake system in Bacillus subtilis. Uptake of the Mg\(^{2+}\)·citrate\(^{-3}\) complex is strongly dependent on the growth conditions. The expression of the transporter involved in the uptake, CitM, is strictly regulated by induction and catabolite repression, which involves the action of several regulatory proteins. In the remainder of this introduction an overview is given of different transport systems and the different mechanisms of induction and catabolite repression in bacteria, with a focus on the situation in B. subtilis.
Different transport systems can be recognized on the basis of the type of energy source that is used to drive solute uptake. Types of metabolic energy include chemical, redox, light and electrochemical energy. Chemical energy is the energy stored in chemical bonds of compounds, which is released upon chemical reactions. In biological systems the hydrolysis of ATP is the major source of metabolic energy. Redox energy is chemical energy released in specific reactions, the redox reactions, which involve the transfer of electrons between molecules. An example is the electron transfer from NADH to molecular oxygen catalyzed by a chain of membrane embedded electron carriers, the electron transport chain, which drives the transport of protons out of the cell. Light is an energy source used by f.i. bacteriorhodopsin, a membrane protein of *Halobacterium* species that absorbs the light and transduces the energy into electrochemical energy. Finally, electrochemical energy is the energy stored in electrochemical gradients of ions and solutes across the cytoplasmic membrane. The electrochemical gradients of protons or sodium ions that exist across the membrane of biological cells lead to a proton motive force (pmf) and sodium ion motive force (smf), respectively. Both forces consist of a chemical component, the pH or Na⁺ concentration gradient, and an electrical component, the membrane potential.

Transporters that are driven by chemical, redox, or light energy are termed primary transporters. The transporters that generate the pmf and smf are usually of this type. They convert chemical (f.i. F₁F₀-ATPase), redox (f.i. electron transfer chain) or light (f.i. bacteriorhodopsin) energy into electrochemical gradients of protons or sodium ions. ATP-driven ABC transporters for the uptake or excretion of solutes belong to the class of primary transporters and will be discussed below. Secondary solute transporters drive the uptake or efflux of solutes by converting the electrochemical energy of one solute into the electrochemical energy of another solute. The transporters are driven by the pmf or the smf or both. A third type of transporters, the group translocation systems, chemically modifies the translocated solute when it passes the membrane. In bacteria the phosphoenolpyruvate-dependent phosphotransferase system (PTS) transports and concomitantly phosphorylates sugars at the expense of the phosphoenolpyruvate (PEP).

*B. subtilis*, the model organism for Gram-positive bacteria with a low GC content, contains a large number of proteins involved in transport processes. Of all transport proteins present 63% is secondary and 27% is primary, whereas 6% belongs to the group translocation systems (Paulsen *et al.*, 2000).
Chapter I

**ABC-transporters**

ATP-driven solute transporters, also called ATP-binding cassette transporters or ABC-transporters for short, use the energy derived from ATP hydrolysis to drive the transport of the solute. The transport process can occur in both directions: into the cell, uptake systems, and out of the cell, excretion systems. The former systems contain an additional component, the binding protein. ABC-transporters are primarily built up of four domains encoded by one to four genes. The domains are present as part of multidomain proteins or as separate proteins. Two domains are located in the membrane and two in the cytoplasmic space. The membrane domains usually consist each of six transmembrane α-helices that together form the translocator. The two cytoplasmic domains called the ATPase subunits, contain ATP-binding sites and are responsible for ATP hydrolysis. The uptake systems contain in addition a binding protein that is located extracellular. The binding protein recognizes the substrate and delivers it to the transport system. In Gram-negative bacteria the solute binding protein moves freely in the periplasmic space, whereas in Gram-positive bacteria, that lacks this space, the solute binding proteins are anchored to the membrane via a lipid or lipoprotein anchor, which allows movement over the membrane surface. More recently, an ABC-transporter was presented in which the solute binding proteins were fused to the integral membrane domains (OpuA of *Lactococcus lactis*; van der Heide & Poolman, 2000). Binding protein-dependent transport systems are found for many growth substrates, like sugars, amino acids, and peptides. The excretion systems serve an important function in protecting the cells from toxic compounds by extruding those compounds from the cell. They can either handle a single substrate or show a broad specificity for multiple substrates. The latter type of excretion system, called multidrug resistance proteins (MDR), has clinical importance, since it leads to the resistance against many antibiotics (toxic compounds) and makes it difficult to treat infections of pathogenic bacteria.

**Secondary transporters**

Secondary transport systems use the electrochemical gradient of protons and sodium ions across the membrane to drive the uptake and excretion of solutes. They couple the transport of a solute to the movement of protons and/or sodium ions. Under physiological conditions, the force on the protons and sodium ions, the sum of the chemical gradients and the membrane potential, is directed inwardly, making the co-ions move from ‘out’ to ‘in’. The solute can be transported in or out depending on the coupling mechanism. Symporters transport solutes and co-ions simultaneously in the same direction, thereby transporting the solute into the cell. Antiporters transport solutes and co-ions in opposite directions and remove the solute from the cell. Uniporters catalyze the transport of a solute without
coupling it to the transport of a co-ion. Transport is in the direction of its electrochemical concentration gradient. The total force in the transport reaction depends on the stoichiometry and the direction of the transport process.

Secondary transporters are mainly single gene products. They roughly contain 11 to 15 α-helices that span the membrane and are connected by intra- and extracellular loops. Interestingly, many secondary transporters are not (fully) functional in their monomeric state, but oligomerize into dimeric or multimeric structures to catalyze uptake, probably thereby increasing the interaction with signal proteins and improving regulatory activities (Veenhoff et al., 2002; Heuberger et al., 2002).

The Mg$^{2+}$-citrate transporter, CitM, of *B. subtilis* is a secondary transporter that belongs to the MeCit family of transporters. CitM catalyzes a symport reaction that couples the uptake of Mg$^{2+}$-citrate$^{3-}$ to the uptake of several protons (Boorsma et al., 1996).

**Bacterial phosphoenolpyruvate-dependent phosphotransferase system**

The phosphoenolpyruvate-dependent phosphotransferase system (PTS) mainly transports and phosphorylates mono- and disaccharides in a limited number of bacteria. In addition, it is involved in multiple regulatory mechanisms that control gene expression. It couples sugar transport to phosphorylation driven by the high phosphate transfer potential of phosphoenolpyruvate (PEP) (Postma et al., 1993). The phosphoryl group of PEP is transferred to the sugar via a phosphotransfer cascade consisting of both cytoplasmic and membrane bound components. Cytoplasmic HPr is phosphorylated by PEP in a reaction catalyzed by enzyme I (EI). The phosphoryl group is covalently linked to a histidine residue. Subsequently, HPr(\text{His}--\text{P}) donates the phosphoryl group to enzyme II (EII) that donates it to the sugar during which process the sugar is transported into the cell. The two cytoplasmic components EI and HPr are not specific for the different sugar substrates and are referred to as general phosphotransfer proteins, while a separate EII exists for each sugar substrate.

EII itself is composed of three to four domains, which can exist as separate individual proteins or can be organized in a multidomain protein (Saier & Reizer, 1992). The EIIA and EIIB domains are involved in the phosphoryl transfer and localized in or protrude into the cytoplasm, whereas EIIC and EIID (if present) are trans-membrane and responsible for the translocation of the sugar across the membrane.

Besides a function in sugar transport, the PTS intermediates generate signals for regulatory functions. It is involved in regulation of expression of catabolic genes via phosphorylation of transcriptional regulators and by generation of co-factors of CcpA, a central component in carbon catabolite repression (Stülke & Hillen, 1998). Such regulation mechanisms are described in the following paragraphs.
Figure 1. Solute transport systems in bacteria. (A) ATP-driven ABC transporters for the uptake and excretion of solutes. The first system shows the architecture of a binding protein-dependent uptake system. The second system shows the architecture of an excretion system. (B) Secondary transporters, depicted are a symporter, an antiporter, and a uniporter, respectively. (C) Phosphoenolpyruvate-dependent phosphotransferase system for the uptake of sugar substrates.
INDUCTION OF CATABOLIC PATHWAYS

Expression of genes and operons, encoding enzymes required for the uptake and metabolism of specific nutrients, is generally dependent on the presence of the respective nutrient in the growth medium (induction) and the absence of carbon sources that are preferably metabolized (carbon catabolite repression; see below). For the induction process the presence of nutrients first of all need to be detected to determine which genes or operons have to be expressed. The detection may take place inside or outside the cell and may sense the nutrient itself or a (catabolic) signal derived thereof. Secondly, following detection of specific nutrients, a signal has to be transduced to the transcription apparatus to start gene expression. The signal is transmitted to transcriptional regulators that act by increasing the rate of transcription initiation or by varying the length of transcripts by premature transcription termination (Steinmetz 1993; Rutberg 1997; Martin-Verstraete et al., 1998). The different modes of detection and signaling result in several modes of induction control ranging from simple ‘road blocks’ to complex phosphorelay systems. The following summarizes different mechanisms of induction.

Relief of repression

In general, in the absence of inducer repressor proteins are bound to operator sequences in the promoter region of the respective gene(s) or operon(s), thereby making transcription impossible (‘road block’). Binding of the inducer in the cell by the repressor protein reduces the affinity of the repressor for the DNA, thereby releasing the regulator from the operator site and allowing transcription to proceed (Fig. two examples repressor action). Transcriptional repressors can be grouped by homology into several families of transcriptional regulators. Common regulator families are GalR/LacI (Weickert & Adhya, 1992), DeoR (van Rooijen & de Vos, 1990) and GntR (Reizer et al., 1991), which are characterized by significant sequence homology extending over several domains of the family members that possess similar functions. The protein sequences of the regulators contain well-conserved motifs such as cofactor binding domains for binding of the inducer, helix-turn-helix and zinc finger DNA-binding motifs for interaction with the operator site, and oligomerization motifs for the formation of functionally active dimers or multimers (Weickert & Adhya, 1992; Gallegos et al., 1997; Bell & Lewis, 2001; Huffman & Brennan, 2002).

The classical example of this mode of induction is the regulation of the lac operon by the lac repressor in E. coli. In B. subtilis GntR, the prototype of the GntR family of regulatory proteins represses the expression of the gluconate operon (gnt). When gluconate or gluconolactone, the inducers of the gnt operon, bind to the cofactor binding domain of
Chapter I

GntR, the repressor loses its affinity for the gnt operator, allowing release from the DNA and progress of gene transcription (Stülke & Hillen, 2000).

Figure 2. Schematic representation of regulation by transcriptional repressor proteins. Situation A represents regulation in the absence of inducer (non-inducing conditions), situation B represents regulation in the presence of inducer (inducing conditions). Indicated are: R, repressor, and RNAp, RNA polymerase.

Activator proteins

Transcriptional activators bind to a region upstream or near the promoter region of the target genes where they activate gene transcription by interaction with the RNA polymerase. Binding affinity of the transcriptional activators to the DNA is controlled in two ways, either by direct binding of their specific inducer inside the cell or by phosphorylation in response to inducer availability. The latter is observed in two-component systems and in the PRD-mediated regulation (PTS regulation domain) that will be described below. Here the activation by regulators, which bind their inducer, is discussed. Most transcriptional activators that bind inducers contain a well-conserved DNA-binding domain (helix-turn-helix) and a less-conserved domain for binding of the inducer. An example of a transcriptional activator in B. subtilis is GutR that activates transcription of the glucitol operon. GutR shares some similarity to the members of the 100-kDA transcriptional activator family and was found to possess a typical helix-turn-helix motif at the N-terminal region that binds the GutR binding site upstream of the gut promoter (Poon et al., 2001a).
Although GutR can bind to this regulatory region in the absence of the inducer glucitol, it was demonstrated that in the presence of glucitol binding of GutR was much stronger. It was suggested that, in addition, GutR has to bind ATP for activation and that the tight binding induced by glucitol allows GutR to be positioned at the proper location for ATP binding to occur (Poon et al., 2001b).

**Figure 3.** Schematic representation of regulation by transcriptional activator proteins.
Situation A represents regulation in the absence of inducer (non-inducing conditions), situation B represents regulation in the presence of inducer (inducing conditions).
Indicated are: A, activator, and RNAp, RNA polymerase.

**Two-component systems**

The transcriptional repressors and activators described above rely on the entrance of the inducer into the cell. In contrast, two-component systems act upon detection of the inducer located at the outside of the cell. The system represents a common mechanism to translate the presence or absence of external inducers into the expression of genes needed for the uptake and metabolism of the respective nutrients. The system senses the presence of the inducer at the outer surface of the cell and signals its presence via phosphotransfer to the transcriptional machinery. Two-component systems are made up of two conserved components, a sensor histidine kinase and a response regulator. The sensor histidine kinase is composed of two domains: a sensor domain that functions to recognize the (external) signal (inducer), and a highly conserved kinase domain that undergoes an ATP-dependent...
autophosphorylation at a conserved histidine residue. The autophosphorylation is a bimolecular reaction, in which one histidine kinase monomer catalyzes the phosphorylation of the second monomer. In turn, the second component, the response regulator catalyzes the phosphoryl transfer from the histidine kinase to a conserved aspartate residue in its own regulatory domain. Phosphorylation of the response regulator is supposed to promote conformational changes affecting its effector domain, which possesses DNA-binding activity and functions to activate transcription of specific genes. Some response regulators are known to act as repressors (Stock et al., 2000).

**Figure 4.** Schematic representation of regulation by a two-component system. Situation A represents regulation in the absence of inducer (non-inducing conditions), situation B represents regulation in the presence of inducer (inducing conditions). Indicated are: HK, histidine kinase, RR, response regulator, and RNAp, RNA polymerase.
In *B. subtilis*, 36 sensor histidine kinases and 35 response regulators have been detected in a genome-wide search (Kobayashi *et al.*, 2001) that in addition to the regulation of catabolic pathways are involved in the regulation of processes like chemotaxis (Falke *et al.*, 1997), aerobic/anaerobic respiration (Nakano *et al.*, 1998), and sporulation (Perego 1998). For the regulation of these global processes, which involve multiple proteins, there are variations on the simple two-step signal transduction scheme described above. For instance, multiple histidine kinases phosphorylate the same response regulator or one histidine kinase phosphorylates multiple response regulators. A more elaborate version of the two-component system, is the four-component signal transduction system, also known as the phosphorelay, which comprises a cascade of phosphorylation reactions, involving a histidine kinase, a phosphorylated aspartate domain of a response regulator, a phosphotransferase and a response regulator (Stock *et al.*, 2000).

Expression of the *citM* gene encoding the Mg-citrate transporter of *B. subtilis* is positively regulated by the CitS-CitT two-component system. The sensor histidine kinase CitS senses the presence of Mg-citrate at the outside of the cell and via autophosphorylation transfers the signal to the response regulator CitT. CitT subsequently binds to the *citT* binding site located upstream of the *citM* promoter region and enhances *citM* gene expression (Yamamoto *et al.*, 2000; this thesis). Most interestingly, all genes and operons encoding TCA cycle enzymes and permeases are positively regulated by associated two-component systems.

**PRD-mediated regulators**

The induction mechanism of many genes encoding proteins involved in the uptake of PTS substrates is indirect and involves intermediates of the phosphotransferase uptake system itself. The transcriptional regulators contain so-called PRD-domains (PTS regulation domain) that are phosphorylated by components of the PTS. Phosphorylation regulates the activity of the regulators. Induction involving PRD-mediated regulators represents a mechanism in which transport and regulation of transcription are integrated.

PRD-mediated regulators can be divided into two groups based on their activation mechanism: activators and antiterminators. Activators bind to activating sequences located upstream (upstream activating sequence, UAS) of the regulated gene or operon and activate transcription by interacting with the RNA polymerase, similarly as described for the activator proteins that directly bind the inducers (Stülke & Hillen, 2000). Antiterminators prevent the premature termination of transcription by interfering with the formation of so-called transcriptional terminators, stem-loop structures in the mRNA that are located between the transcriptional and translational start point (Rutberg 1997). The binding of an
antiterminator protein to a ribonucleic antiterminator sequence (RAT), which is capable of forming a competing stem-loop structure that overlaps the terminator site, precludes formation of the transcription terminator and results in transcription read-through by the RNA polymerase (Henkin 2000; Rutberg 1997). Both types of PRD-mediated regulators are found in the same organism. In *B. subtilis*, transcription of the *lev*, *lic*, *man*, and *mtl* operons is regulated by the PRD-mediated activators LevR, LicR, ManR, and MltR, respectively, whereas the *sac*, *bgl*, and *pts* operons are regulated by the PRD-mediated antiterminators SacT and SacY, LicT, and GlcT, respectively.

All PRD-mediated regulators contain two PRDs that are targets for both positive and negative regulation. Positive regulation is mediated by HPr(His15~P) via phosphorylation of one PRD, whereas negative regulation is mediated by the sugar-specific EII via phosphorylation of the other PRD. When both PRD-domains become phosphorylated negative control overrules the positive control resulting in an inactive regulator. In the case of LevR of *B. subtilis*, this results in the following mechanism of induction.

Besides the two PRD-domains, LevR contains an N-terminal helix-turn-helix motif to bind DNA and a central domain necessary for interaction with the RNA polymerase. Regulation of expression follows the following sequence of events. In the absence of any PTS substrate both HPr and EIIbLev are phosphorylated and donate their phosphoryl group to the PRDs of LevR leading to a double phosphorylated, but inactive regulator. If only the inducer fructose is present EIIbLev becomes dephosphorylated due to phosphoryl transfer to fructose during transport. As a result LevR becomes dephosphorylated at one of its PRDs, but is still phosphorylated at the other PRD by HPr(His15~P). This renders LevR active and results in activation of gene transcription. In the presence of PTS substrates, in- or excluding fructose, the level of HPr(His15~P) will be low and therefore PRD-phosphorylation leading to activation of LevR and subsequently activation of gene transcription is diminished (Stülke *et al.*, 1998).

By nature, the PRD-mediated transcriptional activators and antiterminators regulate operons coding for PTS-mediated sugar transport systems. Two antiterminators are concerned with the regulation of the non-PTS histidine and glycerol operons. These follow deviating activation mechanisms (Glatz *et al.*, 1996; Darbon *et al.*, 2002; Oda *et al.*, 2000).
Figure 5. Schematic representation of regulation by PRD-mediated regulators involving the PTS system. Situation A represents regulation in the absence of inducer (non-inducing conditions), situation B represents regulation in the presence of inducer (inducing conditions). Indicated are the components of the PTS (EIICBA, HPr, EI); PRD, PTS regulation domain; RNAp, RNA polymerase.
Chapter I

CARBON CATABOLITE REPRESSION

Historical overview

In 1942 Monod discovered the phenomenon ‘diauxie’, also referred to as ‘adaptation’ or ‘the glucose effect’. Growth of *E. coli* in the presence of the two carbohydrates glucose and lactose, resulted in a growth curve exhibiting two successive growth cycles separated by a lag-period (Monod 1942), depending on the carbohydrates added (Ullmann 1996). The mechanism to achieve diauxic growth in *E. coli* on glucose and lactose has been the paradigm for carbon catabolite repression (CCR) in bacteria and gene regulation in general. Magasanik launched the name ‘catabolite repression’ after it became clear that not only glucose, but any compound which can serve efficiently as a source of intermediary metabolites (catabolites), could reduce (repress) the synthesis rate of proteins involved in the uptake and metabolism of less efficient substrates (Ullmann 1996).

The mechanisms of CCR differ substantially among bacteria, but in all cases involve the PTS. The preferred carbon source, usually glucose, is taken up by the PTS and the phosphorylation state of one of the phosphotransfer intermediates serves as the signal for a transcriptional regulator that affects transcription of the target gene or operon. In the model organism for Gram-negative enteric bacteria, *Escherichia coli*, the repressive effect of glucose on the metabolism of lactose has been studied extensively and was found to operate at the level of transcription of the lactose operon (*lac*).

The glucose-specific EIIA of the PTS (EIIA\textsubscript{Glc}) is central for glucose repression of the *lac* operon. In the absence of glucose, EIIA\textsubscript{Glc} is phosphorylated and catalyzes the transfer of its phosphoryl group to adenylate cyclase, 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 complex binds a specific DNA sequence near the promoter of the *lac* operon. While bound to the DNA, the cAMP-CRP complex can contact RNA polymerase at the promoter site, leading to increased rates of transcription initiation. In the presence of glucose EIIA\textsubscript{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 *lac* operon (Postma *et al.*, 1993).

In addition to the regulation of transcription, unphosphorylated EIIA\textsubscript{Glc} can inhibit the action of the lactose permease, the secondary transporter that mediates the uptake of lactose, and thereby prevents formation of the intracellular inducer allolactose. This process that excludes the inducer from entering the cell is called inducer exclusion.
General introduction

CCR in Gram-positive bacteria was studied much later in time and similarly involves PTS intermediates. But since Gram-positives lack adenylate cyclase/cAMP CCR turned out to follow quite a different mechanism.

Carbon catabolite repression in Gram-positive bacteria

In contrast to *E. coli*, the PTS component involved in CCR signaling in Gram-positive bacteria is HPr and not EIIA. HPr in Gram-positives is phosphorylated at two sites, a histidine and a serine residue. The histidyl phosphorylated HPr is involved in transport and PRD-mediated regulation, whereas the seryl phosphorylated HPr is involved in CCR. Other components implicated are HPr kinase and the global transcriptional regulator CcpA.

![Figure 6. Outline of the PTS phosphorylation cascade from phosphoenolpyruvate to the sugar.](image)

The sugar is transported from outside the cell (sugar$_{\text{out}}$) into the cell (sugar$_{\text{in}}$). The phosphorylation of HPr at the serine residue links the PTS to carbon catabolite repression.

The regulatory mechanism is discussed in the text.

In the presence of a preferred carbon source like glucose, HPr is mostly in the unphosphorylated state, since it transfers its phosphoryl group 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 so-called glycolytic intermediates, like fructose-1,6-bisphosphate (FBP). These glycolytic intermediates activate HPr kinase that catalyzes the phosphorylation of HPr at the serine residue (Ser46 in *B. subtilis* HPr) in an ATP-dependent manner. Subsequently, seryl-phosphorylated HPr interacts with the catabolite control protein CcpA to make binding to a cis-acting catabolite repressive element (cre) in the promoter region of the target gene possible. Binding of this CcpA-HPr(Ser-P) protein complex at this site prevents
transcription of genes, i.e. it functions as a ‘road block’ for transcription by RNA polymerase. Like in *E. coli*, the phenomenon of inducer exclusion exists in *B. subtilis*, but, again, differs considerably. In the absence of glucose EIIA$^{\text{Glc}}$ is in the phosphorylated state. Certain PTS sugar EII permeases that lack an own sugar-specific EIIA domain (Reizer *et al.*, 1999), employ the energized EIIA$^{\text{Glc}}$P for transport. Hence, when glucose is present, EIIA$^{\text{Glc}}$ is involved in glucose uptake and not available to serve as part of other EIIIs. In this way, the uptake of other PTS sugars is excluded (inducer exclusion) (Dahl 2002; Yamamoto *et al.*, 2001). Second, but slightly different, inducer exclusion is found with the utilization of glycerol (*glp* operon). The inducer of the *glp* operon, including the glycerol transporter (GlpF) and kinase (GlpK), is glycerol-3-phosphate that is the product of glycerol kinase. The activity of glycerol kinase is stimulated by HPr(How-P), which is diminished in the presence of repressing carbon sources, linking glucose availability to the abolishment of glycerol uptake (Reizer *et al.*, 1984).

In addition to inducer exclusion, a phenomenon called inducer prevention has been described in *B. subtilis*. Inducer prevention interferes with the process of induction, via an anti-inducer, thereby preventing activation of transcription. An example of an anti-inducer would be glucose-6-phosphate. The trehalose and xylose repressor proteins, TreR and XylR respectively, normally function as repressors of the respective operons and are released from their operator sites when the inducers trehalose-6-phosphate and xylose are present. However, glucose-6-phosphate, a product of trehalose degradation and glucose translocation, is able to bind the transcriptional regulators, thereby preventing binding of the inducers (Dahl 2002; Kraus *et al.*, 1994).
Figure 7. Carbon catabolite repression in Gram-positive bacteria. HPr phosphorylated at the histidine by enzyme I in a PEP-dependent manner has a catalytic function in the PTS. HPr phosphorylated at the serine by HPr kinase in an ATP-dependent manner has a regulatory function in carbon catabolite repression. The binding of a CcpA dimer and two HPr(Ser-P)s to the catabolite responsive element (cre) results in repression of transcription.
Chapter I

Components of catabolite repression in Gram-positive bacteria

HPr kinase/phosphatase

HPr kinase is the primary sensor in CCR. It is a bifunctional enzyme with both kinase and phosphatase activities (HPrK/P) and belongs to a new family of ATP-dependent protein kinases (Reizer et al., 1998; Galinier et al., 1998; Kravanja et al., 1999). HPrK/P phosphorylates HPr and the HPr-like protein, Crh (see below), at a serine residue in an ATP-dependent manner. Fructose-1,6-bisphosphate (FBP) or other glycolytic intermediates promote kinase activity, while inorganic phosphate (Pi) inhibits the activity. It has been reported that HPrK/P forms homo-oligomers (octamers) and displays cooperative binding of ATP and FBP (Jault et al., 2000). Complementary, the HPr phosphatase activity is stimulated by inorganic phosphate, but is not affected by the presence of ATP (Reizer et al., 1998; Galinier et al., 1998; Kravanja et al., 1999). So, the physiological state of the cell is sensed by changes of the kinase and phosphatase activities. HPrK/P is constitutively expressed, irrespective of the presence of glucose in the medium (Reizer et al., 1988; Hanson et al., 2002).

Two highly conserved regions were identified in the primary structure of HPrK/P, an ATP/GTP binding motif, a Walker A box, and a signature sequence, which is likely to be involved in interaction with HPr and HPr(Ser-P). Mutations in the Walker A box resulted in reduced kinase and phosphatase activities, suggesting that both activities are linked to each other (Hanson et al., 2002). ATP and Pi compete for binding to the same site in the Walker A box, explaining the inhibition of kinase activity by high concentrations of Pi (Monedero et al., 2001). In an hprK null mutant, phosphorylation of HPr and Crh at the serine residue is restrained, resulting in loss of catabolite repression of several genes (f.i. xynB and bgIP in B. subtilis) (Reizer et al., 1998). It was also demonstrated that HPrK/P binds HPr and Crh with similar affinity and that the presence of magnesium is necessary for the interactions (Lavergne et al., 2002).

HPr and Crh

In the phosphotransfer cascade resulting in the uptake of a PTS sugar HPr is phosphorylated at a histidine residue (His15 in B. subtilis) by PEP catalyzed by EI and, subsequently, transfers its phosphoryl group to the sugar-specific EII. In CCR, HPr is phosphorylated at a regulatory phosphorylation site, a serine residue (Ser46 in B. subtilis). This reaction is carried out by the ATP-dependent kinase, HPrK/P. Phosphorylation on the serine inhibits the PEP-dependent phosphorylation on the histidine. Seryl-phosphorylated HPr interacts with the global regulator CcpA to give rise to repression of gene transcription. Mutant strains in which the serine was replaced by a non-phosphorylatable residue (Eisermann et
General introduction

Like alanine, showed loss of repression of several catabolic genes. However, the mutation exhibits only minor effects on, for instance, the repression of xynB expression (Galinier et al., 1997), as a result of the presence of a second corepressor for CcpA in B. subtilis.

A novel gene encoding an HPr-like protein, Crh (Catabolite repression HPr), was discovered during the genome sequencing project of Bacillus subtilis (Kunst et al., 1997; Galinier et al., 1997). Crh exhibits 45% sequence identity with HPr and, like HPr, possesses a serine residue at position 46 that can be phosphorylated by the ATP-dependent HPrK/P. Crh(Ser-P) acts as a corepressor of CcpA (Galinier et al., 1997, 1999; Turinsky et al., 1998; Zalieckas et al., 1998; Martin-Verstraete et al., 1999a). Remarkably, Crh lacks the active site histidine of HPr, where it is replaced with a glutamine, i.e. Crh is not able to carry out the catalytic and regulatory functions mediated by HPr(His~P). However, replacement of the glutamine with a histidine, enables Crh to carry out most of the functions exerted by HPr (Martin-Verstraete et al., 1999b). For example, it could restore fructose-dependent induction of the lev operon (see above) in an HPr mutant strain and also efficiently phosphorylate and activate the PRD-containing antiterminator LicT, when it is expressed in monocopy. In contrast, phosphorylation and activation of glycerol kinase and growth of an HPr deletion strain on the PTS carbohydrates, glucose, fructose and mannitol required expression in multicopy (Darbon et al., 2001), suggesting inefficient transport function.

The structures of both HPr and Crh are determined by NMR studies and reveal similar overall topology. Both proteins contain a hydrophobic patch that is involved in binding of CcpA. In HPr, the same hydrophobic area is involved in interaction with EI, EIIA, and PRD-containing regulators. A major difference between Crh and HPr is that the former exists as a mixture of monomers and dimers in solutions, whereas HPr only exists as monomers. Regions in Crh that are implicated in dimer formation show structural deviations from the corresponding regions in HPr, which could explain the observation that HPr does not dimerize. The implication of Crh dimerization in regulation of gene expression is not clear yet (Penin et al., 2001; Favier et al., 2002).

CcpA

In B. subtilis the catabolite control protein A (CcpA) was first implicated in the regulation of α-amylase synthesis and revealed a specific interaction with a cis-acting palindromic sequence, called cre (catabolite responsive element), located in the promoter region of the amyE gene (Henkin et al., 1991; Kim & Chambliss, 1997). CcpA, is a member of the GalR/LacI family of transcriptional regulators and affects the expression of many catabolic genes and operons in response to glucose availability (Weickert & Adhya, 1992; Henkin
Chapter I

1996). It contains a helix-turn-helix DNA-binding domain, an effector binding site and a
dimerization domain (Henkin et al., 1991). Dimerization of two CcpA monomers allows the
formation of two recognition helices in the DNA binding domains of the CcpA dimer, which
bind the cre site at two sites (Kim & Chambliss, 1997). The final complex that results in
CCR consists of a CcpA dimer, two HPr(Ser-P)s or Crh(Ser-P)s and the cre sequence (Jones
et al., 1997). CcpA mutant proteins deficient in cre binding lost the ability to exert
catabolite repression.

Synthesis of CcpA differed not much during growth in the presence or absence of glucose,
indicating that CcpA activity in response to glucose availability is not due to variations in
CcpA protein levels, but is controlled at the post-transcriptional level (Henkin 1996)
depending on the binding of corepressors at the effector domain. Binding of seryl-
phosphorylated HPr or Crh to CcpA enhanced the interaction with the cre sequence,
illustrated by the fact that non-phosphorylated HPr has a decreased affinity for CcpA by
three orders of magnitude and resulted in diminished cre binding. Furthermore, CcpA first
binds HPr(Ser-P) or Crh(Ser-P) and then recognizes the cre sequence.

Some reports suggest that the interaction between CcpA and HPr(Ser-P) is enhanced by
fructose-1,6-bisphosphate (FBP) independent of the effect of the latter on HPrK/P (Galinier
et al., 1999), but this has been disputed by others (Martin-Verstraete et al., 1999a). Studies
on amyE gene repression revealed that in vitro NADP functions as a cofactor of CcpA (Kim
et al., 1998). Also, CcpA can bind to some extent to the cre site in the absence of any
cofactor.

Besides its function as repressor, CcpA can also serve as an activator of genes. While
operons involved in the degradation of secondary carbon sources (f.i. TCA cycle
intermediates) are derepressed in strains deficient for CcpA, genes involved in glycolysis,
overflow metabolism, and ammonium assimilation were not expressed and led to impaired
growth (Tobisch et al., 1999; Faires et al., 1999). However, in some cases distinct pathways
for growth defects and catabolite repression may be present (Küster et al., 1999).

A recent whole-transcriptome analysis of B. subtilis demonstrated that additional ways of
CcpA-mediated gene repression and activation may exist. Besides the classical CcpA-
mediated glucose-dependent repression or activation, there is evidence for CcpA-mediated
but glucose-independent repression and activation, as well as glucose-dependent but CcpA-
independent repression (Moreno et al., 2001).

Cre site

The operator sequence cre (catabolite responsive element) that is recognized by the
CcpA/HPr(Ser-P) protein complex is a palindromic 14-bp consensus sequence (Weickert &
Chambliss, 1990; Miwa et al., 2000). Deviation from the consensus leads to decreased function, which can be explained by a change in cre binding strength of CcpA. Since the cre is (partially) palindromic the sequence on the sense and anti-sense strand are more or less similar. Nevertheless, mismatches in the sense strand are more critical for binding then mismatches in the anti-sense strand and a more palindromic nature improves the function of cre (Miwa et al., 2000). AT-rich flanking regions enhance the function of cre since it improves CcpA binding, illustrated by actual contacts between CcpA and the base pairs adjacent to the cre site (Kim & Chambliss, 1997). Furthermore, it was demonstrated that the position of the cre site relative to the transcription start point was of importance. Cre sites located within or close to the promoter region prevent transcription initiation by blocking the RNA polymerase (as is the case for amyE, bglPH, and acu of B. subtilis). The cre of the lev operon lies between the upstream activating sequence of LevR and the promoter, which prevents interaction of the transcriptional activator LevR and the RNA polymerase (Martin-Verstraete et al., 1995). Cre sites downstream of the transcriptional start point may block the elongation RNA polymerase (examples are acsA, xyl, gnt, hut, and xyn of B. subtilis). For the activation of genes, cre sites are located upstream of the transcriptional start point (Presecan-Siedel et al., 1999; Turinsky et al., 1998). Deletion of the cre or mutations of the conserved base pairs in the consensus sequence have been shown to lead to the loss of catabolite repression or activation.

A search for cre sequences in the Bacillus genome picked out at least 126 cre-like sequences of which 22 were found to function in vivo, including the cre of CitM, the Mg-citrate transporter (Miwa et al., 2000; Yamamoto et al., 2000; this thesis). Some operons contain more than one cre site suggesting regulation of expression under special conditions.
Chapter I

CcpB and CcpC

In addition to CcpA, two other transcriptional regulators have been described in *B. subtilis* that were shown to be involved in CCR. CcpB, the product of the *ccpB* gene, displays 30% identity to CcpA and, thus belongs to the GalR/LacI family of transcriptional regulators. It possesses a highly conserved helix-turn-helix motif, suggesting that CcpB binds to DNA. It was proposed that CcpB binds to the same *cre* sites as CcpA and in a similar fashion (i.e. in complex with HPr(Ser-P)). Mutations in *ccpB* partially relieved the sensitivity of the *xyl* and *gnt* operons of *B. subtilis* to repression by glucose, mannitol and sucrose. However, this only was true when cells were grown on solid media or in liquid media with low agitation rates. The mechanism and exact function of CcpB action still remains unknown (Chauvaux *et al.*, 1998).

In spite of its name, the second transcriptional regulator, CcpC, belongs to another family of transcriptional regulators, namely the LysR family. CcpC has been shown to bind to palindromic DNA sequences (helix-turn-helix motif) in the promoter region of target genes (Kim *et al.*, 2002a). The *citB* gene of *B. subtilis*, encodingaconitase, is fully repressed when cells are grown in minimal medium containing glucose and glutamine. In addition, CcpC contributes to the regulation of *citZ*, encoding the major citrate synthase (Jourlin-Castelli *et al.*, 2000). Interaction of CcpC with the *citB* promoter region takes place at an interrupted dyad symmetry element (palindromic sequence) and a second copy of one arm of the dyad, which is located 32 bp downstream of the first dyad. The presence of citrate disrupts the interaction between CcpC and this dyad sequence, showing that citrate acts as the inducer of the *citB* gene. To further complicate things, the synthesis of CcpC itself is regulated by CcpA, resulting in regulation of *citB* expression by CcpA in an indirect way (Kim *et al.*, 2002b).

OUTLINE OF THIS THESIS

Metal ions are essential for life, for instance as structural or catalytic co-factors in metalloproteins. For this reason bacteria make use of various transport systems for the uptake and, subsequently, usage of metals from the environment. Besides uptake systems for free metal ions, bacteria also possess specific transport proteins that mediate the uptake of metals complexed to so-called chelating agents (Paulsen & Saier, 1997; Thomas *et al.*, 1998; Krom *et al.*, 2000). *B. subtilis* contains both types of transport systems to acquire essential elements, like calcium and magnesium (Silver *et al.*, 1975; Silver & Lusk, 1999; Krom *et al.*, 2000).
The metal citrate transport protein, CitM, of *B. subtilis* mediates the uptake of citrate complexed to divalent metal ions, like magnesium. It is a member of a small distinct family of secondary transporters, the MeCit family (Krom *et al.*, 2002), which use the free energy stored in transmembrane electrochemical gradients of protons to drive the uptake of the substrate (Poolman & Konings, 1993; Boorsma *et al.*, 1996). A second metal citrate transport protein, called CitH (distinct from CitH, the malate dehydrogenase), also belongs to this family and shares 52% sequence identity with CitM (Boorsma *et al.*, 1996). For both proteins the availability of divalent metal ions was found to be essential for functioning in citrate uptake. Most interestingly, the proteins have complementary metal ion specificity with CitM transporting citrate complexed to Mg$^{2+}$, Mn$^{2+}$, Ni$^{2+}$, Co$^{2+}$, and Zn$^{2+}$ and CitH citrate complexed to Ca$^{2+}$, Sr$^{2+}$, and Ba$^{2+}$ (Krom *et al.*, 2000). The main physiological function of CitM is to obtain citrate, whereas the physiological function of CitH is not clear (Krom *et al.*, 2002; this thesis).

The main topic of this thesis concerns the regulation of transcription of the *citM* gene of *B. subtilis*, encoding the major metal citrate transporter. The work focuses on the regulation of *citM* at the transcriptional level and predominantly presents data on the influence of growth media and components therein on expression. The effect on transcription of so-called primary carbon sources that are preferably utilized was investigated in Chapter 2. In addition, Chapter 2 describes the involvement of the catabolite repression system in the regulation of *citM* expression and evaluates its effect on CitM activity. The use of the lacZ reporter gene in the determination of gene expression is subject of debate in Chapter 3, since the level of reporter gene expression appears to change during the exponential growth phase and was shown to be dependent on the growth rate of the cell culture. A kinetic model is presented to explain some obscure points. The expression of *citM* is induced when citrate is present in the growth medium (Chapter 2), but further investigations revealed that also isocitrate is an inducer of the system (Chapter 4). Furthermore, Chapter 4 describes that isocitrate, and more specifically D-isocitrate, is a substrate for the CitM transporter and that transport is dependent on the availability of divalent metal ions. Chapter 5 deals with regulation of *citM* expression in a distinct way, independent of the global regulatory protein CcpA. The key effector in this novel regulation turned out to be arginine and the metabolic route leading to ornithine is shown to be involved. In Chapter 6 an overview is given of the distribution of HPr-like proteins in bacteria. The presence of Crh-mediated signal transduction, as it is described for Gram-positive bacteria, was discovered for some Gram-negative bacteria. Furthermore, Crh-specific and HPr-independent regulation of *citM* expression is demonstrated. Finally, Chapter 7 summarizes the results of this thesis.