Transcriptional regulation of central amino acid metabolism in Lactococcus lactis
Larsen, Rasmus

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:
2005

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.
Chapter 1

General introduction

Rasmus Larsen
Chapter 1
Bacteria are unicellular organisms, unique in their ability to live and thrive in almost any environmental niche of the world, from deep-sea vents to the mammalian gut. Whereas the make-up of housekeeping genes ultimately determines the environment in which a bacterial species can live, it is the ability to adapt that determines their ability to cope with sudden changes in that environment. To this end a range of regulatory mechanisms has evolved, continuously adjusting the activity of cellular processes in response to external as well as internal stimuli. Strict control of metabolic processes is essential, since the expression of unnecessary proteins would pose an energetically and competitively unsustainable strain on the organism.

The lactic acid bacteria (LAB) comprise a diverse group of organisms in the Clostridium branch of prokaryotes, including species of the genera *Carnobacterium*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Oenococcus*, *Pediococcus*, *Streptococcus*, *Tetragenococcus*, *Vagococcus* and *Weissella*. Common features are their Gram-positive cell wall, the low-G+C content of their genome, their non-spore forming, non-motile, microaerophilic life styles, and importantly, the production of lactic acid as the main end-product of sugar fermentation, which has earned them their name. LAB reside in many different natural habitats, including soil, plant material, and the oral cavity and intestinal tract of animal species. Whereas many LAB are known for their beneficial use in food production and preservation, like strains of *Lactococcus* and *Lactobacillus*, others, e.g. in the genus *Streptococcus* are renowned for their less appreciated ability to inflict human and animal diseases.

Their relatively simple metabolism and small genomes (1.8 Mbp for *O. oeni* to 3.3 Mbp for *Lb. plantarum* (129)) have made LAB good model organisms of “minimal living” of low-G+C Gram-positive bacteria. Additionally, new areas of research warrant promising near-future applications for LAB, such as the production of health products and bacteriocins, investigation into bacterial multidrug resistance, and as vehicles for delivery of vaccines.

The genus *Lactococcus* includes the species *L. lactis*, *L. raffinolactis*, *L. garviæ*, and *L. piscium*. *L. lactis* is by far the best studied of the food-related LAB, which is largely due to its major industrial importance as starter strain in the manufacture of fermented dairy products. The capacity for lactate and antibiotic (bacteriocin) production of *L. lactis* has been beneficial for preservation of food-
factor) binds to the promoter, forming a closed promoter complex. An open promoter complex is formed by ‘melting’ of the DNA double strand just downstream of the promoter region. Binding of nucleotide triphosphates to the single stranded DNA in the unwound region results in the initiation complex and transcription elongation can start by the release of the $\sigma$ factor.

Different $\sigma$ factors are known to respond to different signals and to activate specific genes accordingly, e.g. $\sigma^B$ of B. subtilis and $\sigma^{54}$ of E. coli are activated in response to environmental stress, after which they activate specific subsets of genes (106, 204). The genome sequence of L. lactis IL1403 contains three putative $\sigma$ factor genes: rpoD, comX and sigX (24), of which only the first has been shown to be functional in L. lactis; the functions of ComX and SigX remain to be investigated (3). Amino acid sequence and gene neighbourhood comparisons show that L. lactis ComX is similar to the competence $\sigma$ factor of another LAB, Streptococcus pneumoniae (100). However, competence development in L. lactis has so far not been observed. Vegetative promoters of L. lactis (recognised by RpoD, also called $\sigma^{39}$), generally contain a $-10$ sequence or Pribnow box (TATAAT) and a $-35$ sequence (TTGACA) (55), common to all bacteria (Fig. 1.1). Additionally, in some L. lactis promoters, a so-called extended $-10$ motif (TG) is located upstream of $-10$ with one bp spacing, and in strongly expressed promoters a $-44$ AGTT motif is present upstream of $-35$ (Fig. 1.1) (55, 187). Finally, several promoters lack the $-35$ sequence, and it has been proposed that transcription from such promoters is dependent on specific trans-acting factors (see below) to direct the L. lactis RNAP holoenzyme to these promoters. In the remainder of the thesis, this region, containing the $-10$ and $-35$ motifs will be referred to as the ‘core promoter’ (Fig. 1.1).

**Trans-acting transcriptional regulation**

Whereas the degree to which a promoter sequence resembles the consensus sequence influences the strength of expression, it obviously does not allow differential regulation, and especially not so, if only a single $\sigma$ factor is active, as might be the case in L. lactis. Thus, regulation of transcription initiation must be carried out by other trans-acting factors.

A very complex, extensively studied mechanism is the stringent response. The intracellular level of guanosine-3’,5’-bisphosphate (ppGpp) increases when the cells’
translation machinery is affected (slowed down) as a consequence of amino acid limitation or starvation. ppGpp directly interferes with, and destabilises, the open RNAP-promoter complex, leading to interruption of transcription initiation. This effect seems to mainly occur for promoters expressing components of the translation machinery, such as ribosomal-RNA (rRNA) genes. Stringent response has not been studied in any detail in the food-related LAB, but has been shown to be in effect in *Streptococcus* (167, 168). However, an increase of ppGpp in *L. lactis* has been observed to be related to the development of multi-stress tolerance (198, 199). A possible discriminator (GTACTGTT) motif is well-conserved at the transcription start site of rRNA genes of *L. lactis* (187), and may be involved in regulation via the stringent response.

![Figure 1.1](image)

**Figure 1.1.** The *L. lactis* RNA polymerase and promoter sequence. Schematic representation of the RNAP holoenzyme, consisting of the large RNAP complex and the sigma factor $\sigma^{39}$. The positioning of the RNAP and $\sigma^{39}$ fits with the interaction sites in the promoter region shown below. The common $-35$ and $-10$ sequences are shown on black background and ‘ext.’ denotes the extended $-10$ sequence (TG). The less common $-44$ and ‘discriminator’ sequences, only found in strong promoters, are shown on white background. The transcription start site is indicated by ‘+1’ and a bent arrow, showing the direction of transcription. The nucleotide distance between the $-35$ and $-10$, and between the $-10$ and transcription start site is indicated in bp.

Several bacterial genes encode proteins that are responsible for up- or down-regulation of gene expression. Some regulators, called dedicated regulators, are involved in the control of one or several genes that are all related to a certain function, *e.g.* the same metabolic pathway. In contrast, so-called global regulators are responsible for the control of genes related to different cellular functions, and can in
principle regulate other regulators, forming a regulatory hierarchy. The sum of genes regulated by a single regulator is generally referred to as a ‘regulon’.

Transcriptional regulators usually have a DNA-binding domain (DBD) and a sensing domain, allowing binding in or near core promoters in response to a signal. An environmental signal can be converted into a regulator-mediated response in a number of ways (Fig. 1.2). First, the mere quantity of the regulator can be modulated either via changed expression of the regulator gene, or by proteolytic degradation of the regulator protein. Second, small ligands, e.g. intracellular metabolites can bind to the regulator sensing domain, resulting in a change of conformation or multimerization, affecting the DNA affinity of the DBD. Third, the regulator can be sequestered by another protein, preventing access of the DBD to DNA. Finally, the regulator can be covalently modified as is generally seen for two-component systems, where a membrane-embedded sensor kinase transduces an extracellular signal to the response regulator by phosphorylating the latter.

Regulators recognise DNA regions, called operator sites, which are specific for the individual regulator. Depending on the location of the operator site, the regulator interferes with binding of the RNAP holoenzyme to the promoter and thereby with expression of the respective genes. In short, if the operator overlaps the core promoter, steric hindrance will prevent RNAP from binding to the promoter, and thus inhibit initiation of transcription (Fig. 1.2). This mechanism is referred to as repression. In contrast, the regulator can interact directly with RNAP, or bind to an operator upstream of the promoter, thereby increasing the affinity of RNAP for the promoter (Fig. 1.2). The result, an increase in transcription and gene expression, is referred to as activation. An example of the latter is MleR of *L. lactis*, involved in malolactic fermentation, and the first activator found in Gram-positive bacteria (205). Other regulators affect the bending or coiling of DNA in the vicinity of the promoter, which can weaken or strengthen the interaction of RNAP to the promoter, or the formation of the open complex (Fig. 1.2). As may be expected, a range of variations on these basic concepts exists, e.g. with several different regulators repressing or activating the same promoter, and single regulators acting as activators or repressors on separate promoters (Fig. 1.2).
Figure 1.2. Schematic examples of transcriptional regulation at prokaryotic promoters. Oval circles represent the RNAP holoenzyme. Black boxes indicate the core promoter, consisting of the –35 and –10 sequences, and the transcription start site is indicated by a bent arrow, pointing in the direction of transcription (a crossed-over arrow indicates inhibition of transcription). Activator (light) and repressor (dark) proteins are shown to interact with operator regions presented as oblong, white boxes. (A) The activator is bound to an operator upstream of the promoter, making direct interaction with the RNAP, thereby strengthening the RNAP-promoter complex and increasing the level of transcription. (B) The activator binds in or near the promoter, inducing a conformational change of the DNA that can either improve alignment of the –35 and –10 sequences, or affect the open promoter complex formation during transcription initiation steps. (C) Cooperative interaction between two different activators, leading to RNAP-promoter stabilisation (as (A)), through direct protein-protein interaction between the regulators and RNAP. The co-activator does not necessarily interact directly with both the other activator and RNAP, but may be just one of them. (D) The repressor binds to an operator site overlapping the core promoter, preventing RNAP-promoter binding by steric hindrance. (E) The repressor binds to operators which are located at a distance from the actual promoter, inducing bending of the DNA between the operator sites. DNA-bending in the promoter region prevents RNAP binding and, thus, initiation of transcription. (F) Repression of transcription through sequestration of the activator by a secondary regulatory factor, which may or may not bind DNA itself. This mechanism requires that the activator is necessary for transcription initiation, which might be the case in some core promoters lacking –35 sequences. What accounts for the individual mechanisms of activation or repression presented here, is often also observed to result in the opposite effect, e.g. activation of transcription can also occur through sequestration of a repressor (opposite of (F)), called derepression.
Transcriptional regulators in *Lactococcus lactis*

The nucleotide sequence of the entire genome of *L. lactis* IL1403 (2,365,589 bp, ~2310 ORFs) allowed the identification of 138 genes with possible functions in gene regulation. Of these, 111 were postulated to be transcriptional regulators (24, 100). The function of only 18 of these have been investigated experimentally, including regulators involved in carbon metabolism: CcpA, LacR, SacR, BglR, XylR, MalR, and MleR (2, 12, 74, 155, 156, 205, 252), nitrogen metabolism: CodY, CmbR, GlnR, ArgR, and AhrC (77, 101, 142), this thesis), nucleotide metabolism: PurR, AldR, and PyrR (97, 128, 165), and in a range of stress responses: BusR, GadR, FlpA, FlpB, CSPs, and CtsR (210, 213) (reviewed by Guédon et al. (2002) (100)). Characteristics of the catabolite control protein CcpA and the regulators involved in nitrogen metabolism are detailed in a later section of this chapter. The functions of 34 others have been proposed on the basis of amino acid sequence homology and genome localisation, while the remaining 59 could not be classified (100). The function of six chromosomally located two-component systems (KinA/LlrA, B, C, D, E, and F) of *L. lactis* MG1363 have been analysed by integration knockout studies and phenotypic characterization (190, 191). The systems were shown to be involved in acid resistance, possibly through activation of the arginine deiminase pathway (ADI, see below), salt, osmotic and oxidative stress, and phosphatase activity (191). Additionally, a histidine kinase (*kinG*) and two response regulators (*llrG* and *llrH*) have been suggested on basis of sequence homology, but functions remain to be determined experimentally (24). *L. lactis* two-component systems are also involved in the production of bacteriocins such as nisin (*nisRK*) (134, 249) and resistance to copper (*lcoRS*) (127).

The number of transcriptional regulators in *L. lactis* is about half of that of the low-G+C ‘model’ organism *B. subtilis*, but considering that the genome size of the latter is twice as large as that of *L. lactis*, the relative number of regulators is similar for the two organisms (141). Strikingly different, however, is the number of (putative) σ factors, with 3 in *L. lactis* compared to 18 in *B. subtilis*, and 8-10 two-component systems in *L. lactis* compared to 34 in *B. subtilis* (141). It is generally recognised that the number of proteins involved in gene regulation and integration of environmental
signals correlates with the natural habitat of the organism (37). The more this niche is subject to sudden changes, the more systems bacteria apparently have to be able to sense and adapt to these changes. While *L. lactis* is naturally present in rich environments such as milk, *B. subtilis* naturally occurs in nutritionally poor soils that are subject to constantly changing conditions.

**Nitrogen regulation in Gram-positive bacteria**

Nitrogen is present in many cellular compounds, including amino acids, nucleotides, amino sugars, NAD, cofactors and vitamins, some of which are ultimately incorporated into macromolecules such as proteins, DNA, RNA and cell wall components. To deal with the fluctuation in extracellular as well as intracellular nitrogen availability, bacteria have evolved elaborate regulatory mechanisms controlling assimilation, utilisation and integration of nitrogen into central metabolism. In principle, nitrogen metabolism entails the metabolism of all nitrogen-containing compounds but most nitrogen moieties revolve around the two amino acids glutamine and glutamate, which act as direct substrates in the biosynthesis of other compounds or as donors of amino groups.

After a short introduction to nitrogen utilisation in *L. lactis*, this section will focus on the enzymes and transcriptional regulators controlling the cellular pools of glutamine and glutamate. Since an extensive overview of nitrogen regulation in Gram-positive bacteria has only been obtained in *B. subtilis*, knowledge from this organism will be used as a basis for a comparison to *L. lactis* and other Gram-positive bacteria.

**Nitrogen metabolism in *L. lactis***

An important characteristic of *L. lactis* is the ability to degrade and assimilate milk proteins (caseins) as main nitrogen source in cellular metabolism. This essential proteolytic system of *L. lactis* has been extensively studied and documented during the past two decades (for reviews see (45, 132, 137)). The proteolytic system consists of a cell wall-bound, extracellular proteinase PrtP and up to 15 intracellular peptidases, many of which are regulated on the transcriptional level by the global regulator CodY (see below). Oligopeptides produced by PrtP hydrolysis of casein are
transported across the cell wall by oligopeptide and di-/tri-peptide transport systems. In fact, the growth of *L. lactis* in milk depends mainly on utilisation of extracellular oligopeptides, and less on free amino acids (125, 174). *L. lactis* is multiple amino acid auxotrophic and, with some strain variation, leucine, valine, isoleucine, methionine, histidine, and glutamine or glutamate, are essential amino acids (59, 143, 203). The inability to produce some of these amino acids is caused by point-mutations or more severe lesions in the genes of biosynthetic pathways, which is attributed to the adaptation of this organism to nitrogen-rich environments such as milk (59, 61, 95).

Amino acids liberated by proteolytic activity can be excreted, used directly as building blocks in protein synthesis, or degraded further for the production of metabolic energy. They cannot be stored to be used as energy sources during starvation. Transcriptional regulation of amino acid metabolism in *L. lactis* has been investigated for only a few amino acids. The histidine biosynthetic operon is regulated by a T-box attenuation mechanism, while tryptophan biosynthesis is regulated through transcription antitermination. These are examples of transcription-termination regulation rather than transcription initiation (60, 202, 248). Interestingly, the expression of the *hom-thrB* threonine biosynthetic operon is not regulated in response to threonine availability (158). The *metC-cysK* operon, involved in biosynthesis of the sulfur-containing amino acids cysteine and methionine, is activated by the transcriptional regulator CmbR, possibly in response to acetylserine (77). As this thesis describes the transcriptional regulation of glutamine and arginine metabolism, an overview of both will be presented in the following.

**Glutamine and glutamate are central in nitrogen metabolism**

Glutamine is synthesised from glutamate by glutamine synthetase (GS; glutamate-ammonia ligase; encoded by *glnA*), by assimilation of ammonium in an ATP-driven reaction (Fig. 1.3). GS is the main enzyme through which ammonium, whether obtained from the environment or from intracellular catabolic reactions, is cycled back into central nitrogen metabolism. Also, the reaction performed by GS is the only means of glutamine biosynthesis (204). GS is ubiquitous in all organisms and is functional as a dodecameric enzyme in prokaryotes (195). In Gram-negative and high-G+C Gram-positive bacteria the activity of GS is regulated by a covalent adenlylation/deadenyllylation modification system. Each subunit of *E. coli* GS is
individually subject to inactivation by adenylylation (through covalent binding of ADP to a Tyr residue), causing the activity of GS to be inversely proportional to the number of adenylylated subunits. Covalent modification has not been observed for GS of Gram-positive bacteria (65, 66, 195). Additionally, GS activity is subject to allosteric regulation, by what has been termed ‘cumulative feed-back inhibition’, from end products of glutamine metabolism: serine, alanine, glycine, AMP, CTP, tryptophan, histidine, carbamoyl phosphate, and glucosamine-6-phosphate in *E. coli* (256), and glutamine, histidine, glycine, alanine, and AMP in *B. subtilis* (65). The gene encoding GS is present in an operon (glnRA) with the gene for the transcriptional regulator GlnR in a wide range of bacteria including Bacilli and LAB.

In bacteria, glutamate is mainly synthesised by glutamate synthase (GOGAT; glutamine-oxoglutarate aminotransferase; encoded by *gltAB* or *gltBD*) in a transaminase reaction between glutamine and α-ketoglutarate, yielding two molecules of glutamate (Fig. 1.3). Glutamate can also be formed during degradation of other amino acids, or from ammonium by the action of glutamate dehydrogenase (GDH). However, in many low-G+C Gram-positive bacteria including *L. lactis*, GDH is either absent (24), or almost exclusively catabolically active, as in *B. subtilis* (19). This will not be described here (for an overview see (14)). GOGAT is present in most Gram-positive bacteria, including *Bacillus* strains, *Staphylococcus aureus*, *S. mutans*, *Clostridium acetobutylicum* and *L. lactis* (14). Somewhat surprisingly, based on genome sequences, the genes encoding GOGAT are missing in some low-G+C Gram-positive bacteria closely related to *L. lactis*, namely *Enterococcus faecalis*, *S. pneumoniae* and *S. pyogenes* (14). GOGAT is made up of two subunits encoded by the *gltAB* or *gltBD* operons (the nomenclature differs between organisms). In *B. subtilis*, *B. stearothermophilus* and *Staph. aureus*, *gltAB* is divergently transcribed with *gltC*, encoding a transcriptional activator of *gltAB* (see below) (14).
Figure 1.3. Schematic representation of central nitrogen metabolism and its regulation in *B. subtilis* and *L. lactis*. Regulators and enzymes are shown as circles; those known to be present in *L. lactis* are drawn with broad black lines, whereas those present in *B. subtilis* but not identified in *L. lactis* are drawn with thin lines. Abbreviations: GS, glutamine synthetase; GOGAT, glutamate synthase; AT, general aminotransferase. Open arrows indicate the target proteins of main effector compounds, broken arrows indicate direct transcriptional regulation mediated by the regulator protein, and solid divergent arrows indicate direct protein-protein interactions. The vertical lines indicate the cell-membrane and its intracellular (In) and extracellular sides (Out). Ammonium can be transported across the membrane via AmtB or by diffusion, and is subsequently incorporated into nitrogen metabolism via GS. Question marks indicate the fact that GlnK targets have not been found, that the effectors of GltC and GltR are unknown, and that it remains unclear whether there is direct interaction between GlnR and GS.

GS and GOGAT together form a cyclic pathway of which the sum of the reactions is the formation of one molecule of glutamate from ammonium and α-ketoglutarate, at the expense of one ATP molecule (Fig. 1.3). In this process glutamine serves as the intermediate nitrogen carrier. During nitrogen starvation or when the glutamine pool is drained for biosynthetic reactions, glutamate has to be supplied by GDH by transfer of amino groups from other amino acids to α-ketoglutarate, or by the increased production of glutamine via ammonium-
assimilation (Fig. 1.3). The importance of the GS-GOGAT pathway in bacteria becomes obvious when looking at the preferred nitrogen source, the nutritional substrate that gives optimal growth and strongest repression of secondary nitrogen utilisation pathways. The preferred nitrogen source of E. coli is ammonium or glutamine, although the intracellular concentration of glutamine appears to be the main nitrogen signal (204). In B. subtilis the preferred nitrogen source is glutamine, followed by arginine and ammonium (6). Investigations into the effects of single nitrogen sources on LAB growth has been hampered by the fact that LAB are multiple amino acid auxotrophs and are unable to grow in minimal media with single amino acids or ammonium as the sole nitrogen source.

Transcriptional regulators in nitrogen metabolism

Besides the allosteric regulation of enzyme activity described above, the genes of the GS-GOGAT pathway are controlled by several transcriptional regulators in the model organism B. subtilis, including the global regulators GlnR, TnrA and CcpA, and the dedicated regulators GltC and GltR (Fig. 1.3). These and the global regulator CodY will be described in the following. The ultimate function of these regulators is what can be called nitrogen catabolite control, i.e. the repression of pathways involved in the utilisation of non-preferred nitrogen sources, in the presence of a preferred nitrogen source such as glutamine.

GlnR and repression of nitrogen utilisation pathways

B. subtilis GlnR is a DNA binding protein of 135 amino acids, and a member of the MerR family of transcriptional regulators (30). The glnR gene is located adjacent to glnA, and initial studies showed that disruption and point mutations in glnR resulted in loss of glutamine-dependent repression of GS activity (Fig. 1.3) (218, 222). Promoter deletion and mutation studies revealed that the motif TGTNA-7N-TNACA (GlnR box) at position –42 to –58 relative to the P_{glnRA} transcription start site is responsible for the GlnR-mediated repression (221). Footprinting analysis showed that a similar operator motif, located in the glnR core promoter, was also protected by GlnR. Both operators are necessary for glutamine-dependent regulation by GlnR, and stronger binding of GlnR to DNA fragments carrying both operators glnRAO1 and glnRAO2 than to fragments containing either of the single operators suggests cooperative interaction between a GlnR dimer binding at one operator and
another dimer binding at the other operator (31). Point mutation analysis has revealed that the N-terminal domain of GlnR, containing an H-T-H motif, is responsible for binding to DNA. C-terminally mutated regulators either repress glnRA expression continuously or have lost function entirely, indicating that the C-terminal domain of GlnR is necessary for nitrogen-sensing (218, 220). The 3-dimensional structure of GlnR has not been resolved in any organism, but both the *B. cereus* and *B. subtilis* GlnR proteins exist as homodimers in solution (31, 182).

*B. subtilis* GlnR is also responsible for repression of *ureABC*, encoding the urease enzyme, which hydrolyses urea to carbon dioxide and ammonium when cells are grown on poor nitrogen sources (258). The *ureABC* promoter also contains two GlnR boxes, in and upstream of the core promoter, both of which are required for efficient nitrogen-dependent repression by GlnR (258). Whereas the distance between the operators in *PglnRA* is 6 bp, the *PureABC* operators are separated by 88 bp (27, 258). It is speculated that interaction of GlnR to the *PureABC* operators might result in bending of the DNA between these sites (27, 258).

GlnR has been suggested to repress the expression of the transcriptional regulator TnrA (82), and a recent transcriptome analysis using DNA microarrays has identified 65 other genes (besides *glnR* and *glnA*) of which the expression is significantly changed in a *glnR* deletion strain, compared to the wild type strain of *B. subtilis* (122). Surprisingly, however, expression of the *ureABC* genes was not found to be changed, and direct interaction of GlnR to the promoter regions of the identified genes remain to be performed (122).

*TnrA and activation of nitrogen utilisation pathways*

The transcriptional regulator TnrA is homologous to GlnR. The two *B. subtilis* proteins show 65% similarity and 32% identity. High conservation is observed in the N-terminal H-T-H DNA binding domain (69% identity), while the C-terminal domains are far less similar and TnrA is 24 amino acids smaller than GlnR (259). TnrA (*trans-acting* nitrogen regulator) was identified as activator of the *amtBglnK* (previously called *nrgAB*) operon under nitrogen-limiting conditions (Fig. 1.3), in a random integration-knockout screening (259). Lack of transcriptional activation was also observed for the *gabP*, *nasB*, and *ureABC* genes, suggesting TnrA as a global nitrogen regulator in *B. subtilis*. Additionally, a *tnrA* mutant was unable to grow on some secondary nitrogen sources, but grew normally on primary sources such as
glutamine and arginine, demonstrating the importance of TnrA in nitrogen catabolite control (259). A range of studies have confirmed TnrA to be a global nitrogen regulator (Fig. 1.3), activating genes involved in ammonium transport and –sensing (amtBglnK) (259), asparagine degradation (86), nitrite and nitrate assimilation (180, 181), γ-aminobutyrate transport (78, 259), purine catabolism (223), oligopeptide transport (269), glutamine transport (269), and the genes encoding guanine deaminase (189) and a putative γ-glutamyltransferase (269), and its own gene tnra (207). Additionally, TnrA represses the expression of GlnR, GS and GOGAT, a putative proline transporter, a sodium-alanine symporter, and several other genes of unknown function (21, 259, 269).

The high similarity between the N-terminal DNA binding domains of B. subtilis GlnR and TnrA suggested that they might be able to recognise similar operator sites. Indeed, TnrA binds to the same operator motifs as GlnR but, in contrast to GlnR, TnrA is capable of strong binding to single operator sites (261). Single GlnR/TnrA operators have been identified in the –40 to –60 regions upstream of the promoters that are directly activated by TnrA, suggesting that TnrA functions by enhancing binding of RNAP to these promoters (261, 262, 269). This would explain how TnrA is able to repress expression of the glnRA and gltAB operons, where GlnR/TnrA boxes overlap the core promoters: binding of RNAP is prevented when TnrA is active. This type of regulation is an example of bacterial economising, where a single operator motif allows dynamic transcriptional control by two different regulators.

**Putting the pieces together with glutamine synthetase.**

In summary, GlnR and TnrA are homologous DNA binding proteins that are able to recognise similar operator sequences, with TnrA able to bind to single operators, whereas GlnR requires double operator sites. However, the two regulators operate under different conditions. In a nitrogen-rich medium (e.g. high glutamine concentration) GlnR repress the glnRA, ureABC and tnra genes, avoiding unnecessary ATP spilling via GS, maintaining the glutamate pool, and preventing expression of nitrogen utilisation pathways, including the urease and other genes of the TnrA regulon. In a nitrogen-poor medium (e.g. glutamate as sole nitrogen source) GlnR repression is lost, allowing production of TnrA, which activates a range of nitrogen
catabolic pathways. The ammonium produced in this way can be assimilated by GS and is thereby returned to central nitrogen metabolism.

Both GlnR and TnrA are unable to sense nitrogen signals on their own (56, 222). Surprisingly, all regulation by both GlnR and TnrA turned out to be dependent on GS, and it was first suggested that a metabolite, produced by GS, was a nitrogen effector molecule for GlnR and TnrA. It was ultimately proven, however, that the function of TnrA is inhibited by direct protein-protein interaction with GS (Fig. 1.3) (83, 87, 263). GS sequesters TnrA strongly under nitrogen-rich conditions, in the presence of feed-back inhibiting metabolites such as glutamine, glycine, alanine and AMP (Fig. 1.3) (83, 87, 263). GS has the opposite effect on GlnR: it enhances the binding of the regulator to DNA containing GlnR operator sites (182). It is not known whether GlnR directly interacts with GS and the exact mechanism by which GS controls GlnR-mediated regulation remains unknown (103, 182). In any event, the sequestration of transcriptional regulators by a metabolic enzyme is an unprecedented and intriguing mechanism of global transcriptional regulation.

**Single GlnR/TnrA homologues in low-G+C Gram-positives**

Genes similar to tnrA have only been discovered in *B. subtilis*, *B. stearothermophilus* and *B. halodurans* (14). In contrast, glnR is immediately adjacent to and possibly co-transcribed with glnA (encoding GS) in most other low-G+C Gram-positives, including *Staph. aureus*, *B. cereus*, *S. agalactiae*, *S. pneumoniae*, *E. faecalis*, *Lb rhamnosus* and *L. lactis* (24, 102, 182, 253). It is tempting to suggest that the glnRA syntheny is reflected in functional conservation. GlnR-mediated repression of GS has indeed been shown in *B. cereus* (182) and *Staph. aureus* (219). Interestingly, *L. lactis* GlnR is not only responsible for auto-regulatory repression of glnRA but also of the amtBglnK operon, encoding a putative ammonium transporter/sensor protein and a PII signal transduction protein (see below), and glnPQ, encoding a putative glutamine transporter (this thesis, Chapter 2). Thus, *L. lactis* GlnR seems to perform functions carried out by TnrA in *B. subtilis*. The role of GlnR in global nitrogen regulation has not been investigated in other low-G+C Gram-positives that lack TnrA homologues.
Role of GltC and GltR in regulation of gltAB (GOGAT) expression

*B. subtilis* GltC and GltR are homologous proteins (27% identity) of the LysR family of transcriptional regulators (15, 18). The *gltC* gene is transcribed divergently from the GOGAT genes *gltAB*. GltC negatively regulates its own expression and is essential for activating the expression of the GOGAT genes in the absence of glutamate (Fig. 1.3), by binding to operators in the *gltC-gltA* intergenic region (15). Mutagenesis studies have identified residues in GltC involved in nitrogen sensing, but an effector has not been determined (17). The *B. subtilis* Roc pathway (encoded by the *roc* genes) degrades arginine, ornithine and proline to glutamate, and ultimately to ammonium and α-ketoglutarate. Disruption of the individual steps of the Roc pathway suggested that the substrates or products of the last step, glutamate, ammonium and α-ketoglutarate, are effectors for GltC (20). GltR represses its own expression via operators similar to those of GltC, but no other function has been found for the wild-type regulator. In a strain carrying a mutant form of GltR *gltAB* expression was constitutive, but no nitrogen source-effect could be found, for wild-type nor mutant GltR (18).

The function of GltC and GltR has only been investigated in *B. subtilis*. The *gltC* gene is present in *B. stearothermophilus* and *Staph. aureus* and is transcribed divergently from *gltAB* in both organisms. It was noted that the presence of *gltC* correlates with the presence of a RocG-like catabolic glutamate dehydrogenase (14), suggesting that similar regulators are not functional in *L. lactis* and related organisms that lack *rocG* homologues.

The global energy state regulator CodY and nitrogen regulation

CodY of *B. subtilis* is a transcriptional repressor of several pathways involved in nitrogen metabolism (Fig. 1.3), including degradation of urea (*ureABC*), histidine (*hut*), branched-chain amino acids (BCAAs) (*bkd, ilv* and *leu* genes), arginine, ornithine and proline (*roc*), dipeptide and oligopeptide transport (*dpp* and *app* operons), γ-aminobutyrate transport (*gabP*), and several putative amino acid transporters ((177); (82) and references therein). Additionally, CodY controls the expression of genes not directly related to nitrogen metabolism, such as those involved in competence development, sporulation, chemotaxis, and synthesis of antibiotics (144, 175, 177, 200, 225). Target genes are repressed in response to high intracellular levels of GTP and BCAAs, showing that CodY is an energy state
regulator rather than a specific regulator of nitrogen metabolism (200, 226). Thus, during fast growth and a high cellular energy state, CodY prevents the expression of genes involved in utilization of secondary metabolites, and in survival. These genes are derepressed when nutrients become limiting in the early-stationary phase of growth.

An *L. lactis* CodY homologue, with 48% identity to that of *B. subtilis*, has been the subject of thorough investigation in recent years. CodY of *L. lactis* is also a negative regulator, responsible for repression of several genes of the proteolytic system, including the extracellular protease (*prtPM*), peptide transport systems (*opp* and *dtp*), and intracellular peptidases, but also of aminotransferases (*bcaT* and *araT*) involved in amino acid degradation (Fig. 1.3) (39, 62, 101). Also in this case, the intracellular concentration of BCAAs, and especially isoleucine, are effectors of CodY repression in *L. lactis* (62, 101), while the carbon source also affects CodY-mediated repression in *L. lactis* (39). In contrast to the situation in *B. subtilis*, the level of intracellular GTP has no effect on *L. lactis* CodY function (196). It has proven surprisingly difficult to determine a consensus operator motif for CodY in *B. subtilis* and *L. lactis*. However, recent studies of the *L. lactis oppD* promoter region and DNA microarray studies have pointed to a 23-bp region with weak dyad-symmetrical structure that is essential for CodY mediated repression (62).

**The carbon catabolite control protein CcpA and nitrogen regulation**

Carbon catabolite control is carried out by the transcriptional regulator CcpA in low-G+C Gram-positive bacteria. In the presence of a preferred carbon source, such as glucose or fructose, CcpA binds to catabolite-responsive elements (cre sites) upstream of regulated genes and thereby represses the expression of a range of secondary, catabolic pathways (for a review see (246)). *B. subtilis* CcpA mutants are severely inhibited in growth on minimal medium with glucose and ammonium as sole sources of carbon and nitrogen. Recently, it was shown that this growth defect could be restored by the addition of glutamate, and that the glutamate limitation was caused by lack of *gltAB* (encoding GOGAT) induction in the *ccpA* mutant (Fig. 1.3) (75, 254). The lack of *gltAB* induction was caused by the low activity of the phosphotransferase sugar transport system (PTS) in the *ccpA* strain, preventing the uptake of PTS sugars. Indeed, overexpression of PTS in *B. subtilis ccpA* restored *gltAB* induction, directly correlating GOGAT activity and PTS sugars (Fig. 1.3) (254).
Interestingly, despite utilising fundamentally different carbon and nitrogen regulatory systems, the *gltAB* genes of enterobacteria are also directly regulated by the ratio of carbon (α-ketoglutarate) and nitrogen (glutamine) (204), illustrating the importance of this pathway in connecting carbon and nitrogen metabolism, and the importance of balancing these two relative to each other.

**Ammonium sensing and regulation via AmtB and GlnK**

Recently, AmtB and GlnK have been shown to be of major importance for nitrogen sensing and regulation in bacteria. The AmtB protein belongs to the Amt superfamily and is ubiquitous in all domains of life. AmtB is a membrane-bound sensor and transporter of extracellular ammonium (241). GlnK is a cytoplasmic protein of the P II signal transduction family, responsible for regulation of a range of other cytoplasmic proteins, including enzymes and transcription factors, by direct protein-protein interaction (5). The genetic association of the *amtB* and *glnK* genes is highly conserved among Gram-positive and Gram-negative bacteria and Archaea, where their expression is increased during nitrogen starvation (5, 241). Only recently has detailed knowledge been obtained concerning the function of AmtB and GlnK in bacteria, where *E. coli* serves as the paradigm. When the extracellular ammonium concentration is low, ammonium is taken up by AmtB, and converted to glutamine via GS. Due to the low intracellular glutamine pool, the *E. coli* nitrogen regulatory system (Ntr) is active, and GlnK is reversibly inactivated by uridylylation (47, 123). When the glutamine pool rises, the Ntr system shuts down and GlnK is activated by deuridylylation, resulting in binding of GlnK to AmtB, blocking ammonium uptake (47, 123). In GlnK mutants, the Ntr system is constantly active, resulting in the expression of nitrogen utilization pathways, partial inhibition of growth, and slow recovery from nitrogen starvation (23, 47, 123). A single study in *B. subtilis* shows that, in contrast to the Gram-negative system, GlnK is constantly bound to AmtB at the membrane, and there is no indication of covalent modification of GlnK (Fig. 1.3) (64). AmtB is required for ammonium transport at low concentrations, and GlnK is required for high expression of *amtBglnK* (64). *B. subtilis* *amtBglnK* is positively regulated by TnrA during nitrogen-limitation (Fig. 1.3). In contrast, *amtBglnK* is repressed by GlnR in *L. lactis* under conditions of nitrogen-excess (this thesis, Chapter 2). Thus, the same transcriptional regulation is obtained by opposite
regulatory strategies. The exact mechanism of AmtB- and GlnK-mediated regulation in Gram-positive bacteria remains unclear.

**Nitrogen regulation in other Gram-positive bacteria**

The study of Gram-positive bacteria other than *B. subtilis* provides examples of the diversity in nitrogen regulation encountered in this branch of the prokaryotes. The best known nitrogen regulatory strategies, different from the Ntr and GlnR/TnrA systems of *E. coli* and *B. subtilis*, are found in the high-G+C Gram-positive *Corynebacterium glutamicum* and *Streptomyces coelicolor*, and the low-G+C Gram-positive *Clostridium acetobutylicum*.

In *C. glutamicum*, a soil organism primarily known for its capacity for high amino acid production and used as such in the biotechnology industry, the transcriptional regulator AmtR replaces the Ntr and GlnR/TnrA systems of *E. coli* and *B. subtilis*, respectively (119). AmtR is a global regulator responsible for repression of several nitrogen metabolic pathways, including *glnA* (GS) and *gltBD* (GOGAT) in response to nitrogen-excess. The preferred nitrogen source is ammonium, not glutamine (119, 234). Under nitrogen-limitation, the P II signal transduction protein GlnK becomes adenylylated by GlnD (in contrast to that of *E. coli*, which becomes uridylylated) (234). The adenylylated form of GlnK interacts directly with AmtR and thereby prevents DNA-binding, which ultimately leads to derepression of nitrogen-utilisation pathways (34, 234). Other peculiarities of *C. glutamicum* are the presence of at least two glutamine synthetases and three different AmtB-like ammonium transport systems (120, 169).

Nitrogen metabolism and regulation in *S. coelicolor* is less well characterised but presents interesting features, such as the presence of five homologues of the glutamine synthetase enzyme. One GS is of a class generally found in Gram-negative and high-G+C Gram-positive bacteria, regulated by covalent modification. Three others belong to the un-modified low-G+C Gram-positive class of GSs, and the last homologue belongs to a class which is typically found in eukaryotic organisms (80, 111). Additionally, despite a nitrogen signal transduction pathway employing GlnK, UTase and ATase, as seen in *E. coli*, the exact mechanism is different in *S. coelicolor* (111). Two homologous regulators involved in regulation of the GS genes of *S. coelicolor* belong, despite their names GlnR and GlnRII, to the OmpR family of
transcriptional regulators and, thus, are not similar to the GlnR/TnrA of low-G+C Gram-positive bacteria (81).

*Clostridium acetobutylicum*, a low-G+C Gram-positive bacterium, presents a mechanism of *glnA* regulation that is fundamentally different from that in the low-G+C organisms presented previously. Expression of *C. acetobutylicum* *glnA* is regulated by three promoters. P1 and P2 are located upstream of *glnA* and control the expression of the *glnA* mRNA. P3 is located downstream of *glnA* and controls the expression of a small non-coding 43-bp anti-sense RNA to the 5'-end of the *glnA* transcript (121). The ratio of anti-sense RNA over *glnA* mRNA is approximately 1.5 during conditions of nitrogen excess. Under nitrogen limitation a drastic drop in anti-sense RNA production, accompanied by a slight increase in *glnR* mRNA synthesis, results in a ratio of *glnA* mRNA over anti-sense RNA of approximately five, and consequently a similar increase in GS activity (79).

Regulation of arginine metabolism

Arginine is one of the amino acids synthesised and catabolised in close connection to both glutamine and glutamate. The *de novo* arginine biosynthesis pathway originates in glutamate, while glutamine is precursor for the high-energy compound carbamoylphosphate, which is needed for arginine synthesis (Fig. 1.4). Arginine metabolism in *E. coli* has been studied for many years and led to the first definition of a ‘regulon’. Whereas arginine biosynthesis and its regulation are similar in many bacteria, several different pathways and regulatory enzymes are involved in arginine catabolism. After an introduction to arginine biosynthesis and the arginine deiminase (ADI) catabolic pathway, this section will focus on the transcriptional regulators responsible for the control of these processes.

Arginine biosynthesis

Arginine is synthesised from glutamate in eight enzymatic steps (Fig. 1.4) (reviewed by Cunin *et al.* (52) and Glansdorff (94)). In enterobacteria and archaea of the *Sulfolobus* genus, arginine is synthesised in a linear pathway, where the first step is catalysed by acetylglutamate synthase (encoded by *argA*), and removal of the acetyl group in the fifth step is catalysed by acetylnornithinase (encoded by *argE*) (52, 247).
Other bacteria, including the Gram-positive *B. stearothermophilus* and *B. subtilis* utilise a cyclic pathway, where *argJ* encodes a bifunctional ornithine acetyltransferase that returns the acetyl group removed in the fifth step back to the first step, thereby replacing the enzymes encoded by both *argA* and *argE* (14, 265). Some organisms contain of *argJ* as well as *argA* or *argE*: in *C. glutamicum*, *P. aeruginosa* and *S. coelicolor* the ornithine acetyltransferases (*argJ*) have acetyltransferase activity, but have lost the synthase activity, which is carried out by specific acetylglutamate synthases (*argA*) (105, 112, 211).

The enzymes of the *L. lactis* biosynthetic pathway have not been characterised. However, the genes are transcribed from the three operons *argCJDBF*, *gltSargE* and *argGH* (24) (this thesis, Chapter 4). The fact that there is no homologue of *argA*, but both *argJ* and *argE* are present suggests that *L. lactis* uses a cyclic pathway with a monofunctional onithine acetyltransferase (*argJ*) only catalysing the acetylglutamate synthase function. The acetylornithinase (*argE*) as well as the ornithine acetyltransferase (*argJ*) might be able to carry out the fifth step. Enzymatic characterisation of the arginine biosynthetic pathway in *L. lactis* will have to be performed to answer these questions.

The organisation of the arginine biosynthetic pathway genes varies between bacterial species. In Gram-positive bacteria, exemplified by *L. lactis* ((24); this thesis, Chapter 4), *Lb. plantarum* (28), *Bacillus* subsp. (14), *C. glutamicum* ((211), and *S. coelicolor* (112), the arginine biosynthetic genes are clustered, although the intra-operon gene order varies slightly. In contrast, the arginine biosynthetic genes of the Gram-negative bacteria *P. aeruginosa* and *N. gonorrhoeae* are scattered over the chromosome. *E. coli* takes an intermediate position, with one *argCBH* operon, but the remainder of the biosynthetic genes transcribed as monocistronic units (52). Surprisingly, low-G+C Gram-positive bacteria related to *L. lactis* (*S. pneumoniae*, *S. pyogenes*, and *E. faecalis*) do not have any *arg* gene homologues.

Carbamoylphosphate (CP) is an essential intermediate of both arginine and pyrimidine metabolism (Fig. 1.4). It is consumed during arginine biosynthesis (by the product of *argF*, the anabolic ornithine carbamoyltransferase) and is produced during arginine degradation via the ADI pathway (by the product of *arcB*, the catabolic ornithine carbamoyltransferase; see below) or directly from glutamine via carbamoylphosphate synthetase (CPS), of which the two subunits are encoded by *carA* and *carB* (Fig. 1.4). In some cases *carA* and *carB* are co-transcribed and -
regulated with genes of the pyrimidine biosynthetic pathway (*Bacillus* subsp.). In *Lb. plantarum*, however, they are transcribed divergently to the *argCJBDF* operon (28), whereas in *L. lactis carA* is transcribed as the last gene in the *pyrRBPcarA* operon, and *carB* is transcribed monocistronically (164, 165). *L. lactis* only harbors one CPS (encoded by *carA* and *carB*), which is regulated by the availability of pyrimidines in an attenuation mechanism involving the transcriptional regulator PyrR (165), and mainly supplies CP to pyrimidine biosynthesis. *Bacillus* subsp. and *Lb. plantarum* have two CPSs: one dedicated to arginine biosynthesis and one dedicated to pyrimidine biosynthesis, which are regulated in response to arginine and pyrimidine availability, respectively (14, 185).

**Arginine catabolism**

Bacteria possess several pathways for the utilisation of arginine (reviewed by Cunin *et al.* (52)). The number of arginine degradation pathways present in a given bacterium is determined by the environmental challenges and nutritional requirements of that specific organism. The soil organism *P. aeruginosa* is known to harbor as many as four separate pathways for arginine degradation, allowing it to use arginine efficiently as sole source of carbon, nitrogen and energy, during aerobiosis as well as anaerobiosis (193). Low-G+C Gram-positive bacteria in general have one or two different pathways to break down arginine. *B. licheniformis* and *S. aureus* have two routes: the arginase pathway and the arginine deiminase pathway. *B. subtilis* and other *Bacillus* subsp. only use the arginase pathway while many LAB, including *L. lactis*, *Lb. sake*, *E. faecalis*, and *S. pneumoniae* only have the arginine deiminase pathway, or at least the required genes (84).

In the arginase pathway, which is encoded by the *rocABC*, *rocDEF* and *rocG* operons in *Bacillus* (arginine and ornithine catabolism), arginine is cleaved into ornithine and urea, both of which can be further degraded into sources of carbon and nitrogen (91). Ornithine is degraded into pyrroline-5-carboxylate, which can serve as a carbon source, or further into α-ketoglutarate and ammonium with glutamate as an intermediate (20). Urea can be degraded via urease (*ureABC*; subject to nitrogen catabolite control) into carbon dioxide and ammonium, which can be returned to nitrogen metabolism via glutamine synthetase (GS) (6, 50). The *roc* genes are activated by RocR, a transcriptional regulator of the NtrC/NifA family, in response to
ornithine or citrulline, and by the ArgR-type regulator AhrC in response to arginine (36, 92). Additionally, rocG is subject to carbon catabolite repression by CcpA and the arginase pathway is mainly active during aerobic growth conditions (16).

**Figure 1.4.** Schematic representation of the proposed arginine metabolism in *L. lactis*. The arginine biosynthetic pathway comprise the *arg* and *car* genes, the catabolic pathway consists of the *arc* genes. Genes encode enzymes as follows: *glnA*, glutamine synthetase; *gltBD*, glutamate synthase; *argJ*, ornithine acetyltransferase; *argB*, N-acetylglutamate 5-phosphotransferase; *argC*, N-acetylglutamate 5-semialdehyde dehydrogenase; *argD*, N²-acetylornithine 5-aminotransferase; *argE*, acetylornithine acetyltransferase; *argF*, anabolic ornithine carbamoyltransferase; *argG*, argininosuccinate synthetase; *argH*, argininosuccinate; *carA/carB*, carbamoylphosphate synthetase; *arcA*, arginine deiminase; *arcB*, catabolic ornithine carbamoyltransferase; *arcC*, carbamate kinase; *arcD1/2*, arginine-ornithine antiporter; *gltS*, arginine or glutamate transporter.

The arginine deiminase pathway (ADI) is commonly present in many LAB, with exceptions among strains of *L. lactis* subsp. *cremoris*, *Streptococcus* subsp., homofermentative *Lactobacillus* subsp., and *Pediococcus* subsp., and *Oenococcus oeni* subsp. (46). Although the presence of a functional ADI pathway has been used in
the past as a phenotypic characteristic of *L. lactis* subsp. *lactis*, the *L. lactis* subsp. *cremoris* strain MG1363, used throughout the work presented in this thesis, does in fact harbor a functional ADI pathway. Arginine is degraded via the ADI pathway into ammonium, carbon dioxide and ATP, with ornithine as an important by-product, by the subsequent action of arginine deiminase (encoded by *arcA*), a catabolic carbamoyltransferase (*arcB*), and carbamate kinase (*arcC*) (Fig. 1.4). As is apparent from the end-products, the ADI pathway has several physiological functions: it is involved in the production of energy, recycling of nitrogen, and resistance to acid stress (see below).

The ADI pathway is the only amino acid degradation pathway in LAB that is known to provide ATP via substrate level phosphorylation (45), and is therefore of major energetic importance in these organisms, which cannot gain energy through oxidative phosphorylation. Accordingly, glycolysis and arginine catabolism are closely connected. It was shown that *L. lactis* consumes glucose and arginine sequentially, glucose being depleted before arginine utilization is initiated. In the presence of galactose (and absence of glucose), the sugar and arginine are consumed simultaneously (49, 197). Carbon catabolite repression controls *arc* gene expression in a.o. *L. lactis* (197), *Lb. sake* (178), *S. rattus* (99), *S. gordonii* (71), and *E. faecalis* (11), most likely by the binding of CcpA to catabolite-responsive elements (*cre* sites) which have been identified in the *arc* promoter regions. Arginine can be taken up by LAB cells via an arginine/ornithine antiporter, encoded by *arcD* (197, 273). Since ornithine is produced during arginine degradation and antiport does not require ATP, arginine utilisation is executed without initial energy expenditure, in contrast to glycolysis (72).

As mentioned previously, the ammonium produced via the ADI pathway has two physiological functions, namely, in the recycling of nitrogen by ammonium assimilation via glutamine synthetase, and in protection of the cell against acid stress. Many LAB are subject to severe acid stress, whether from the production of lactic acid, or from other acids in the environment. The ammonium results in alkalinisation of the cytoplasm, thereby keeping the intracellular processes and transport systems operative (46, 133). Examples of the importance of this process are the increased acid resistance and survival of the food-related LAB, *L. lactis* (140, 235), *Lb. sanfranciscensis* (54), and *Lb. sakei* (40), the growth of the tooth-colonising *S. gordonii*, *S. sanguis*, and *S. rattus* (38, 70, 71), and the survival of the pathogenic
Introduction

organism \textit{S. pyogenes} (57, 58). The ADI proteins are active at very low pH values (38), and restoration of the ADI enzymes of \textit{L. lactis} during recovery from starvation was shown to be independent of protein synthesis, in contrast to the enzymes of the glycolytic pathway (140). However, it should be noted that despite the apparent influence of the ADI system on acid resistance, acid has not been observed to have any effect on transcriptional regulation of the \textit{arc} genes (46), and induction of the acid tolerance response in \textit{L. lactis} is dependent on protein synthesis (192).

The \textit{arcABC} genes are normally clustered in low-G+C Gram-positive bacteria, and the arginine/ornithine antiporter gene \textit{arcD} is often co-transcribed with these three genes (273). Duplication of the \textit{arc} genes is not unusual and \textit{L. lactis} probably has the most complex \textit{arc} operon known to date, with the \textit{arcABD1C1C2TD2yvaD} gene order (this thesis, chapter 4). \textit{L. lactis} even harbors an additional \textit{arcC3} gene which is not co-transcribed with the remainder of the \textit{arc} genes (24, 273).

Additionally, genes of the Crp/Fnr family of transcriptional regulators, responsible for activation of \textit{arc} expression under anaerobic conditions, are often associated with the \textit{arc} operon, as seen in \textit{P. aeruginosa} (88, 152), \textit{B. licheniformis} (160), \textit{Lb. sakei} (272), \textit{S. gordonii} (71), and \textit{E. faecalis} (11). Regulators of the Fnr-type have been identified in \textit{L. lactis}, but they are not genetically associated with the \textit{arc} operon and there is no indication that they are involved in regulation of the ADI pathway (96, 224). There is evidence that the two-component system \textit{A} of \textit{L. lactis} MG1363 is important for activation of the ADI pathway, but the exact mechanism remains to be determined (191).

Finally, arginine-dependent activation of the ADI pathway and other arginine catabolic pathways, like the arginase pathway of \textit{B. licheniformis} (see above), has been observed in a range of diverse bacteria, including \textit{E. coli} (157), \textit{Bacillus} subsp. (84), and LAB (142, 184). This control is mediated by ArgR-type regulators, which are the subject of the following section.

\textbf{Regulation of arginine metabolism by ArgR-type regulators}

The arginine-dependent repression of arginine biosynthesis and activation of arginine catabolism is mediated by DNA-binding proteins of the ArgR family of transcriptional regulators. In contrast to the diversity of mechanisms of arginine biosynthetic and catabolic pathways between bacteria, the use of ArgR-like regulators
is unusually conserved. ArgR regulators have been identified and experimentally characterised in the Gram-negative *E. coli* (148) and *S. typhimurium* (151), the Gram-positive *B. subtilis* (14), *E. faecalis* (11), *Lb. plantarum* (184), and *L. lactis* (142), the extremophilic *B. stearothermophilus* (68), *Thermotoga neapolitana* and *T. maritima* (67, 179), *Moritella profunda* (266) and the actinomycete *Streptomyces coelicolor* and *S. clavuligerus* (208, 232), while genome sequences reveal homologues in many other bacteria. An exception is *P. aeruginosa* where control of arginine metabolism is carried out by a regulator which, despite the name ArgR, is a member of the AraC/XylS family of transcriptional regulators (193, 194).

The ArgR-type regulators (i) form hexa-oligomeric complexes (Fig. 1.5), (ii) have N-terminal winged-helix-turn-helix (wH-T-H) DNA-binding domains, and (iii) act as transcriptional repressors as well as activators (Fig. 1.6). They also have a role in multimer resolution of Col1 plasmids in *E. coli*.

Arginine regulation in *E. coli* represents the paradigm of arginine regulation in prokaryotes (reviewed by Maas (157)). Chromosomal mutations that gave derepression of arginine biosynthesis and, consequently, resistance to the arginine analogues canavanine (*E. coli*) and arginine hydroxamate (*B. subtilis*), were identified as residing in the genes encoding the respective regulators (229, 230). The 156-amino-acid ArgR of *E. coli* and the 149-amino-acid AhrC of *B. subtilis* are 27 % identical with highest conservation between residue 25-50 of the N-terminal region, and residue 100-135 of the C-terminal region (188). This level of identity was sufficient for AhrC*Bs* to replace ArgR*Ec* in vivo, although ArgR*Ec* could not replace AhrC*Bs* in repression of *PargC* in *B. subtilis* (229, 233). Proteolytic cleavage of AhrC*Bs* revealed that each monomer has two domains (53), and point mutation studies of ArgR*Ec* proved that residues in the N-terminal domain are essential for DNA-binding while residues in the C-terminal domain are essential for multimerisation and arginine sensing (33, 245). Additionally, expression of part of the *E. coli argR* gene corresponding to amino acid 1-96, yielding a monomeric protein, resulted in arginine-independent repression of ornithine carbamoyltransferase (*argF*) synthesis (98). Overexpression of the C-terminal amino acids 79-156 gave no repression but, instead, a slight derepression, possibly caused by interaction of the truncated ArgR with the wild-type ArgR, which was still expressed at normal levels in the cells (157). Overproduction and isolation of ArgR*Ec* (148), AhrC*Bs* (25, 53), and ArgR of *S. typhimurium* (151), yielded protein complexes of approximately 96.000
kDa, corresponding to hexameric protein complexes, independently of the presence of arginine. More recent investigations have shown that *B. stearothermophilus* ArgR and *Thermotoga neapolitana* ArgR are maintained as trimers that assemble into hexameric structures as a function of the concentrations of arginine, DNA, and protein (67, 68, 183, 231).

![Diagram of ArgR/AhrC type regulators and ARG box operators](image)

**Figure 1.5.** Structure of the ArgR/AhrC type regulators and ARG box operators. (A) Structure of the *B. subtilis* AhrC hexamer (Adapted from Dennis et al., 2001). (B) Simplified structure of hexameric ArgR/AhrC type regulators with each monomer represented as a rounded rectangle. The upper trimer is shown by light, the lower trimer by dark shading.

Crystal structures have been obtained of ArgR from *E. coli* (236, 250), *B. stearothermophilus* (183), and *B. subtilis* (63) (Fig. 1.5). The hexameric structures are made up of dimers of trimers, where the N-terminal domains surround a central core of the C-terminal domains. Six arginine molecules are bound in the trimer-trimer interface (acting as a molecular ‘glue’), and it was shown for *B. stearothermophilus* ArgR that arginine binding induces a 15° rotation of one trimer relative to the other in addition to tighter trimer-trimer interaction, which possibly aligns the N-terminal domains for optimal DNA-binding (183). The hexameric structure of the ArgR-type regulators is unique as most other transcriptional regulators are active as dimeric or tetrameric proteins. The DNA-binding domains are so-called winged-helix-turn-helix DNA-binding domains: three α-helices, of which the two last interact with DNA, are
followed by two $\beta$-strands separated by a loop, which make up the ‘wing’ structure (116).

Arginine-dependent regulation involves binding of the regulators to so-called ARG box operators which, like the ArgR-type regulators themselves, are highly conserved in a range of diverse bacteria (163). Seventeen motifs identified in the 8 arginine biosynthetic promoters of *E. coli* allowed the definition of the ARG box as an 18-bp palindromic structure with the consensus sequence WnTGnATWWWATnCAnW (W, A or T; N, any residue; capitals, highly conserved residues) (157). ArgR operators of the arginine biosynthetic promoters in *E. coli* (157) and *M. profunda* (266) consist of two adjacent ARG boxes, separated by 3 bp, whereas they are separated by 2 bp in *B. stearothermophilus* (215), *B. licheniformis* (231), and *T. neapolitana* (231). One hexameric regulator binds simultaneously and cooperatively to both ARG boxes (Fig. 1.6), and major-groove contacts are mainly responsible for binding specificity. Binding takes place on one face of the DNA-helix, inducing $70^\circ$-$85^\circ$ DNA-bending, which is suggested to result in looping of the promoter region (53, 173, 243). ArgR regulators are able to bind single ARG boxes, such as present in the arginine catabolic *rocA* and *rocD* promoters of *B. subtilis*, and in the *argR* promoter of *E. coli* (Fig. 1.6), but the binding to such operators is generally much weaker than to double-ARG box operators (51, 173). In general, repression of gene expression is obtained by the presence of double-ARG box operators overlapping the core promoter regions of the genes, thus preventing transcription initiation. However, there are interesting exceptions: the *argC* promoters of *E. coli* as well as *B. subtilis* contains ARG boxes covering the core promoters, but footprinting analysis and electrophoretic mobility shift assays suggested that single regulators additionally bind to single ARG boxes located further downstream, within the open reading frame, resulting in looping of the DNA (53, 115). Even more intriguing is the arginine-dependent activation of the *rocA* and *rocD* promoters in *B. subtilis*. In this case, *B. subtilis* AhrC activates transcription by binding to single ARG boxes located immediately upstream of the transcriptional initiation site, thus partially overlapping the core promoters. Such an operator location would generally be expected to prevent transcription. However, the addition of the regulator increased DNA-bending in the promoter DNA, suggesting that a change in DNA-conformation might be the explanation for the increased initiation of transcription (Fig. 1.6) (173).
A unique feature of the ArgR regulators is their multi-functionality. The regulators are in general able to repress the expression of arginine biosynthetic genes and activate the expression of arginine catabolic genes (see above). Surprisingly, the *E. coli* regulator is also one of two accessory factors in the *Xer-cer* recombination
system responsible for ColE1 plasmid-multimer resolution, a mechanism completely unrelated to arginine metabolism (113, 233). The \textit{B. subtilis} AhrC regulator could replace the \textit{E. coli} regulator in this function, but no similar function has been observed for ArgR regulators of other organisms (233). Additionally, activation of the catabolic arginase (roc) pathway of \textit{B. subtilis} appears to be induced by direct protein-protein interaction between AhrC and another transcriptional activator, RocR (92).

Recent studies have shown that more than one functional ArgR-type regulator is present in the low-G+C Gram-positive bacteria \textit{E. faecalis} (11), \textit{Lb. plantarum} (184), and \textit{L. lactis} (142). The functions of the \textit{E. faecalis} regulators have not been investigated, but in \textit{Lb. plantarum} and \textit{L. lactis} the presence of two regulators is needed for repression of arginine biosynthesis (142, 184) (this thesis, chapters 3, 4, and 5). Moreover, both \textit{L. lactis} regulators are required for regulation of arginine catabolism but, in this case, the function of the regulators is clearly different, one possibly responsible for DNA-binding, and the other for sensing of arginine (this thesis, chapters 3, 4, and 5) (142). These results can only be explained by assuming that cooperation takes place between the two regulators, mediated through direct protein-protein interaction (this thesis, chapter 5).
Outline of this thesis

The goal of the work presented in this thesis was to analyse the function of three uncharacterised transcriptional regulators involved in the control of central amino acid metabolic pathways in *L. lactis*.

Chapter 2 describes the functional characterisation of the global nitrogen regulator GlnR. DNA microarray analysis, comparing the transcriptomes of a regulator deletion mutant and the wild-type strain, allowed identification of chromosomal GlnR targets. A consensus GlnR operator motif could be discerned by promoter deletion studies and was confirmed by electrophoretic mobility shift assays using a purified His-tagged derivative of GlnR. Additionally, the effect of different nitrogen effector compounds was investigated.

Chapter 3 presents the initial identification and characterisation of the two arginine regulators ArgR and AhrC of *L. lactis*. A random integration screening showed both regulators to be functional and essential for arginine-dependent repression of arginine biosynthetic genes. Chromosomal single and double regulator deletion mutants were constructed and the effect on regulation of arginine biosynthesis as well as catabolism was investigated using a novel low-copy promoter expression vector.

The regulons of the ArgR and AhrC regulators are defined in chapter 4. DNA microarray analyses, in which the transcriptome of the wild-type strain was compared to that of the single as well as to that of the double regulator mutant, showed that ArgR and AhrC are dedicated regulators of arginine metabolism in *L. lactis*. An ARG box consensus sequence could be determined for the biosynthetic promoters, but was absent in the promoter of the catabolic *arc* operon. Both regulators are essential for arginine-mediated activation of the catabolic ADI pathway, and a complex mechanism of regulation is proposed. The indirect effect of disrupted arginine-regulation on expression of the pyrimidine *de novo* biosynthetic pathway and on that of a previously uncharacterised operon is discussed.

The functional characterisation of purified ArgR and AhrC is described in chapter 5. His-tagged derivatives of both regulators were isolated to near purity after overproduction in *L. lactis*. Electrophoretic mobility shift assays confirmed that, also *in vitro*, both regulators are essential for arginine-dependent repression of arginine
biosynthesis and activation of arginine catabolism. Footprinting assays confirmed binding of the regulators to ARG boxes in a biosynthetic promoter, and suggested that regulator binding to the catabolic promoter takes place via ARG box-half sites. A regulatory model is proposed to account for the possible protein-protein interaction between ArgR and AhrC.

Chapter 6 summarises the work presented in this thesis, with an evaluation of its impact and relevance, and possible future directions for continuation of the research are proposed.