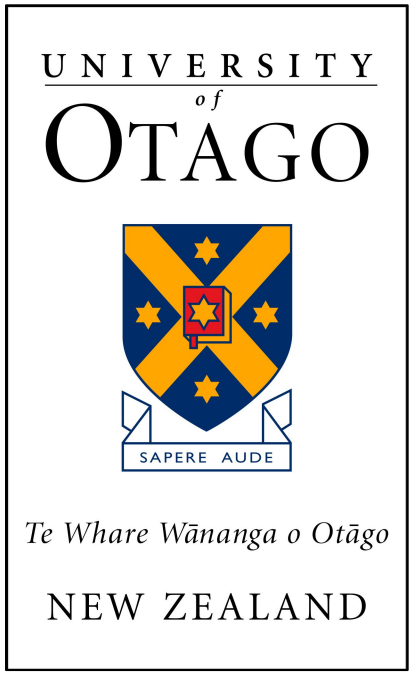
Nitrogen metabolism in *Mycobacterium smegmatis*

Michael Petridis

Dipl. Biol. (Goethe University Frankfurt)



University of Otago

June 30th 2015

This dissertation is submitted in publication format

for the degree of Doctor of Philosophy.

Declaration

This dissertation is the result of my own work and contains nothing, which is the outcome of work done in collaboration with others, except where clearly stated.

No part of this dissertation has been submitted for any other degree or qualification.

This research project was undertaken in the laboratory of Prof. Gregory M. Cook at the Department of Microbiology and Immunology, University of Otago in Dunedin, New Zealand between July 2012 and June 2015. A portion of this research was also performed in the laboratory of Prof. Stewart T. Cole at the École Polytechnique Fédérale de Lausanne in Lausanne, Switzerland in December 2013.

Thesis by Publication Format

This dissertation has been submitted by publication format. The Otago School of Medical Sciences and the Department of Microbiology & Immunology provided consultation on how to submit by this format. This thesis is presented in accordance with their guidelines.

Each of the three results chapters are primary research papers that have been submitted, or prepared for publication. The extended introduction and summary sections were written specifically for this thesis. They serve to place this work in a wider context and bring together the findings as a whole. In accordance with requirements, each chapter has been reformatted for consistency.

I took the leading role in the work presented here. For each study, I planned experiments, performed research and interpreted data. I wrote the majority of each manuscript. Each publication also contains contributions by co-authors of varying size and importance, which are clearly stated. Author contributions are detailed using the initials of each author at the start of each chapter and in the legends of figures and tables. For journal submission, further work was incorporated into the manuscript (Chapter 3; determination of X-ray structure from AmtR). These figures are omitted from this thesis.

Abstract

Nitrogen is an essential component of the bacterial cell and bacteria have developed elaborate regulatory mechanisms used to control the uptake, assimilation and metabolism of nitrogen. In this thesis, I have developed a nitrogen-limited continuous culture system to gain further insights into nitrogen metabolism in mycobacteria and their response to nitrogen limitation. I identified 357 differentially expressed genes in response to nitrogen limitation in continuous culture, including changes in amino acid metabolic pathways. I found 26 transcriptional regulators that mediated the global transcriptomic response of *Mycobacterium smegmatis* to nitrogen limitation and I identified several non-coding RNAs that might be involved in the regulation of nitrogen-regulated gene expression. Subsequently, I characterised two differentially expressed transcriptional regulators, AmtR (MSMEG_4300) and CadC (MSMEG_3297) using a combination of genome-wide expression profiling, physiological, biochemical and biophysical analyses. I identified the AmtR regulon and showed that AmtR was a transcriptional repressor of an urea degradation pathway. I identified xanthine and allantoin as ligands of mycobacterial AmtR and showed that addition of these metabolites had no effect on the release of AmtR from DNA. In further work, I demonstrated that deletion of the gene *cadC* in *M. smegmatis* resulted in a severe growth defect manifested as cell lysis during growth on rich medium. Subsequently, I identified the CadC regulon that included genes involved in the diaminopimelate (DAP) and lysine biosynthesis pathway. Supplementation with DAP or lysine could not rescue the growth defect of the Δ*cadC* mutant. Furthermore, I showed that *M. smegmatis* has a high-affinity lysine uptake system that exhibited high rates of lysine transport during growth in minimal medium, which was significantly reduced during growth in rich medium. My data suggest that a Δ*cadC* mutant is defective in the generation or replenishment of intracellular lysine and DAP levels that are essential for growth and survival in mycobacteria. I conclude that *M. smegmatis* has a broad network of regulatory systems that together enable *M. smegmatis* to adapt its nitrogen metabolism to rapidly changing environments.

Acknowledgements

First, I would like to thank my supervisor, Prof. Gregory Cook. I am extremely grateful that you offered me a PhD position in your lab and for all your effort that made it possible for me to get my student visa and start my PhD. From the very beginning I could feel your contagious passion for science and - to be honest - I never met such an enthusiastic and encouraging person before. You are an outstanding supervisor and a real inspiration in every way and I truly enjoy it to be part of your team.

I would also like to thank my co-supervisor, AsProf. Keith Ireton for his feedback and support throughout the past three years. I would like to thank Dr. Jennifer Robson and Dr. Htin Lin Aung for their encouragement, guidance, advice and support – I am extremely grateful for that. I also express gratitude to Prof. Stewart T. Cole, Andrej Benjak and Claudia Sala for their generosity hosting my stay in Lausanne and I thank AsProf. Peter Fineran, Dr. Michael Berney and Dr. Raymond Staals, three truly exceptional and passionate scientists, for their feedback and support – that is highly appreciated. I acknowledge Prof. Vic Arcus, Chelsea Vickers and Abigail Sharrock for their work on the AmtR project.

Mom, this thesis is dedicated to you! You laid the foundation for this PhD thesis almost 20 years ago and planted the idea in my mind that I should achieve something valuable in my life. You always believed in me and I thank you for your endless support. You are the best mom, I can imagine!

Εμεις μαζι, για μια ζωη

Jess, I would like to thank you for coming with me to New Zealand. Thank you for the time we spent together and your unwavering support over the past three years – thank you!

I would also like to thank my friends who helped me to maintain a balance throughout my PhD and grounded me from time to time – thanks to Christoph & Nadine, Catie, Jussi, Paddi, Rachel & Conan & Theo, Carolyn & Simon, Adrian, Anthony, Corinda, Elyse, Evelyn, Hannah, James, Kiel, Liam, Marion, Nicola, Rachel, Rita, Ron, Shaun, Todd and Yoshio. I also thank everybody else in the Department of Microbiology and Immunology, who provided support and feedback over the years.

I thank the University of Otago for awarding me a Doctoral Scholarship. I further thank The Maurice & Phyllis Paykel Trust, The Webster Centre For Infectious Diseases, The Department of Microbiology and Immunology and The Division of Health Sciences for awarding grants to support travel to the laboratory of Prof. Stewart T. Cole at the EPFL in Lausanne and the Gordon Research Conference – Tuberculosis Drug Discovery & Development in 2015. Finally, I would also like to thank the Stiftung Polytechnische Gesellschaft Frankfurt am Main who laid the foundation for my academic progress by awarding me a mainCampus academicus scholarship during my diploma thesis and their financial support that allowed me to attend the Gordon Research Conference – Bioenergetics in 2012, where I met my supervisor and a true inspiration, Prof. Gregory Cook.

Everything will be alright in the end.
If it's not alright,
it is not yet the end.

Contents

List of Figures

List of Tables

Abbreviations

Å	Ångström
ABC	ABC-transport system
AH	Allophanate hydrolase
AMP	Adenosine monophosphate
asRNA	Antisense RNA
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
°C	Degrees Celsius
cAMP	Cyclic adenosine monophosphate
CFU	Colony forming units
CO_2	Carbon dioxide
CoA	Coenzyme A
CRP	cAMP receptor protein
DAP	Diaminopimelate
DIG	Digoxigenin
DNA	Deoxyribose nucleic acid
dsDNA	Double-stranded DNA
DSF	Differential scanning fluorimetry
EMSA	Electrophoretic mobility shift assay
f.c.	Final concentration
FAD	Flavin adenine dinucleotide
FDR	False discovery rate
FMN	Flavin mononucleotide
g	Gravitational force
g	Gram
GDH	Glutamate dehydrogenase
GluS	Glutamate synthase
GO	Gene Ontology
GS	Glutamine synthetase
h	Hour

HdB	Hartmans-de-Bont minimal medium
HS	High sensitivity
K_2SO_4	Potassium sulfate
kb	Kilobases
kDa	Kilodalton
K_m	Michaelis constant
LB	Luria-Bertani
LBT	Luria-Bertani, supplemented with 0.05% (w/v) Tween-80
MAST	Motif Alignment & Search Tool
MEME	Multiple Expectation Maximization for Motif Elicitation
MFS	Major facilitator superfamily
mg	Milligram
µg	Microgram
min	Minute
MIP	Major intrinsic protein family
mL	Millilitre
µL	Microlitre
mM	Millimoles per litre
µm	Micromoles
µM	Micromoles per litre
mRNA	Messenger ribonucleic acid
NaCl	Sodium chloride
NAD(P)H	Nicotinamide adenine dinucleotide (phosphate)
$(NH_4)_2SO_4$	Ammonium sulfate
NH_4Cl	Ammonium chloride
nm	Nanometre
O/N	Over Night
OD	Optical density
p	*P*-value
PBS(T)	Phosphate buffered saline (with 0.05% (w/v) Tween-80)
PCR	Polymerase chain reaction
PE	Paired-end
poly(dI-dC)	Poly(deoxyinosinic-deoxycytidylic)
psi	Pounds per square inch

PTS	Phosphotransferase system
qPCR	Quantitative real-time PCR
RIN	RNA Integrity Number
RNA	Ribonucleic acid
rpm	Revolutions per minute
RT-PCR	Reverse transcriptase PCR
s	Second
SE	Single-end
SSS	Sodium solute superfamily
TCA	Tricarboxylic acid
t_d	Doubling time
T_m	Melting temperature
U	Unit
UA	Urea amidolyase
UC	Urea carboxylase
UTR	Untranslated region
V_{max}	Maximum velocity
v/v	Volume per volume
WT	Wild-type
w/v	Weight per volume
$Y_{glycerol}$	Glycerol yield

Chapter 1

Introduction

1.1. Overview

Mycobacteria are a diverse group of organisms comprising both environmental bacteria and opportunistic pathogens and the physiological properties of mycobacteria depend on their habitat. Understanding the mycobacterial physiology is a key element to describe ways of how mycobacteria adapt to changes in the environment. Many studies have analysed the adaptation mechanisms of mycobacteria to several stress conditions (e.g. growth rate, oxygen consumption, carbon limitation) and depicted key enzymes responsible for this adaptation (e.g. DosR). Many metabolic pathways important for mycobacterial growth, including carbon metabolism, have been thoroughly described, while nitrogen metabolism received little attention. The aim of this literature review is to present a clear view on the current knowledge of nitrogen metabolism in mycobacteria and to identify gaps that are addressed in this thesis.

1.2. Physiology of Environmental Mycobacteria

Mycobacteria are rod-shaped, acid-fast, G+C-rich organisms that are classified as Gram-positive bacteria with a complex mycolic acid-containing cell wall and belong to the actinobacteria [1]. The genus mycobacterium comprises more than one hundred species, including environmental organisms (e.g. *Mycobacterium smegmatis* and *Mycobacterium fortuitum*) and pathogens (e.g. *Mycobacterium tuberculosis*, *Mycobacterium leprae* and *Mycobacterium bovis*). Mycobacteria are categorised as obligate aerobic and heterotrophic organisms that primarily derive their energy from aerobic respiration of organic compounds [2]. However, recent studies have shown that metabolism of atmospheric hydrogen can enhance growth and survival of *M. smegmatis* under a range of conditions [3,4]. During aerobic respiration, electrons from the oxidation of either NADH or succinate flow into the menaquinone/menaquinole pool, where they are transferred to either cytochrome *c* or cytochrome *bd* under concomitant protein translocation. Protons translocated during this process drive the proton-coupled ATP synthesis via the F_1F_O ATP synthase [2,5].

Carbon metabolism in Mycobacteria

A comprehensive study by Niederweis *et al.* summarized our knowledge of nutrient acquisition in mycobacteria [6]. The nutritional requirements of mycobacteria have been intensively studied and numerous transporters for essential nutrients were identified in the genome of *M. smegmatis* [7-11]. Research on nutrient transport systems in *M. tuberculosis* is greatly understudied, despite the availability of genome sequences [6]. Both environmental and pathogenic mycobacteria are metabolically versatile and able to metabolise various carbon sources, e.g. carbohydrates, TCA cycle intermediates, fatty acids and lipids [2]. However, environmental mycobacteria have a much broader spectrum of substrates that they can metabolise as carbon and energy source, unlike *M. tuberculosis* that encodes only for a reduced number of nutrient uptake systems, a phenomenon that is related to its natural habitat in human macrophages (Figure 1.1A,B) [10]. The genome of *M. smegmatis* encodes for 28 putative carbohydrate uptake systems, including ABC-transport systems (ABC), members of major facilitator superfamily (MFS), phosphotransferase systems (PTS), members of the major intrinsic protein family (MIP) and one member of the sodium solute superfamily (SSS), however, not all of them have been functionally characterised (Figure 1.1A) [10]. In contrast, the genome of *M. tuberculosis* harbors only four ABC-transport systems and one member of the major facilitator superfamily (Figure 1.1B). *M. tuberculosis* lacks an uptake system for glycerol, a carbon source that is commonly used for its growth in *in vitro* studies [10]. Glycerol can diffuse through lipid membranes in concentrations that are sufficient for growth of *M. tuberculosis* and its metabolic route in mycobacterial species is well described [2,12]. Further studies have focused on the effect of growth rate and oxygen consumption in carbon-limited continuous cultures compiling insights into the physiology of mycobacteria [13-15].

[A]

[B]

[C]

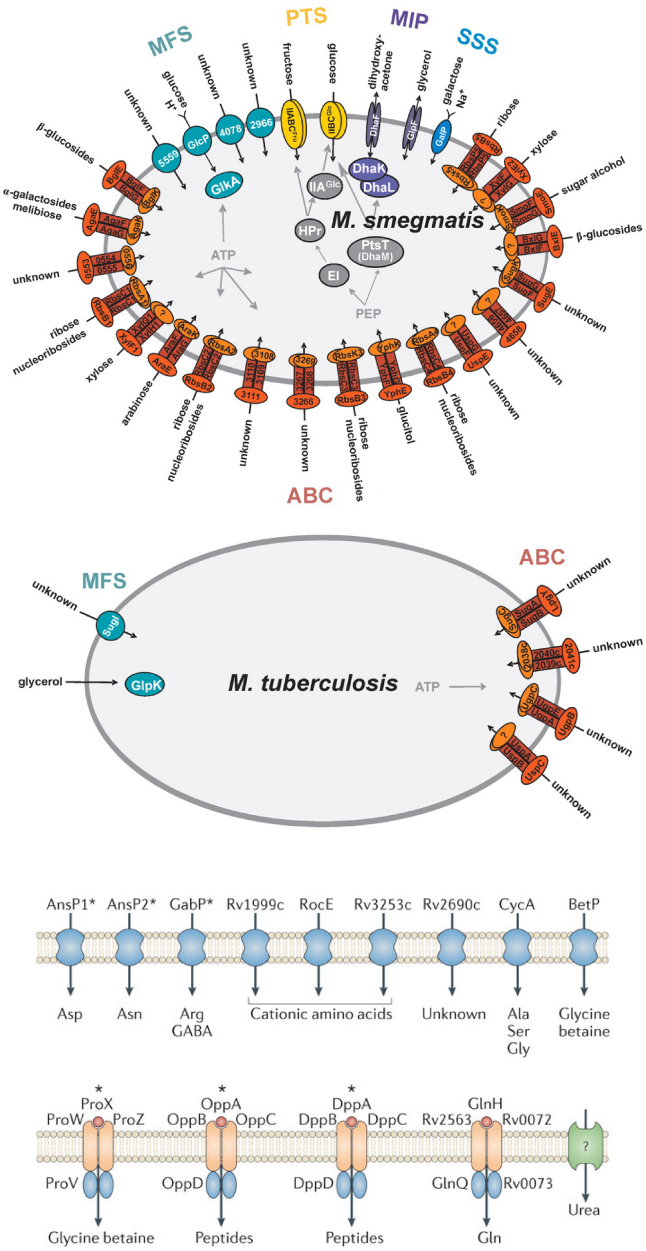


Figure 1.1. Mechanisms of carbon and nitrogen acquisition in *M. smegmatis* and *M. tuberculosis*. Transporters for uptake of carbohydrates in *M. smegmatis* **[A]** and *M. tuberculosis* **[B]**. Environmental mycobacteria can metabolise a broader spectrum of carbon sources than pathogenic mycobacteria, clearly demonstrated by the differences in nutrient acquisition mechanisms within these species. **[C]** Several types of transporters that are required for the acquisition of nitrogen sources in *M. tuberculosis*. Shown above are the nine permeases (blue), four ABC-transporters (orange) and a putative urea transporter (green), which are annotated in the genome of *M. tuberculosis*. Transporters with partial or complete functional characterisation are indicated by an asterisk. ABC: ABC-transport system; MFS: major facilitator superfaimily; MIP: major intrinsic protein; PTS: phosphotransferase system; SSS: sodium solute superfamily. For further information, see text. Figures taken from [8] and [16].

Other nutrient acquisition mechanisms in Mycobacteria

Several uptake systems for phosphorus, sulphur and nitrogen-containing solutes in mycobacteria have been discussed in the literature [7-9,11,16]. The uptake of inorganic phosphate usually occurs via the high-affinity ABC-transport system Pst. *M. smegmatis* contains one copy of the Pst phosphate uptake system, while *M. tuberculosis* harbors several copies of a Pst system in its genome, which are important for virulence *in vivo* [7,17]. Sulfate transport systems are not essential for mycobacterial growth and mutations in these yielded in methionine auxotrophy [8]. Ammonium is the preferred nitrogen source of many bacteria and mycobacterial genomes harbor genes encoding for at least one ammonium uptake system, however, no biochemical data are available for ammonium uptake in mycobacteria [6]. Mycobacterial genomes encode for various other nitrogen uptake systems, but our knowledge surrounding the characterisation of these is poor for both environmental and pathogenic mycobacteria (Figure 1.1C) [11,16]. Amino acid transport systems have turned into focus recently, since their importance in the pathogenicity (e.g. resistance to acid stress and hypoxic stress) of *M. tuberculosis* was reported [18,19].

Environmental mycobacteria can metabolise a broader spectrum of carbon and nitrogen than pathogenic mycobacteria, demonstrated by the differences in the nutrient acquisition mechanisms within different mycobacterial species (Figure 1.1). Thereby, it is uncertain how much of our knowledge about the physiology of *M. smegmatis* can contribute to our understanding about the physiology and virulence of *M. tuberculosis*. While the carbon metabolism in environmental and pathogenic mycobacteria is well studied, we lack great understanding of nitrogen metabolism in mycobacteria. Further studies need to shed light onto adaptive and regulatory mechanisms in mycobacteria in response to changing nitrogen levels. The characterisation of these mechanisms is the subject of this thesis.

1.3. Central Nitrogen Metabolism in Bacteria

Nitrogen is an essential component of almost all macromolecules (nucleic acids, proteins, cell wall, etc.) in a bacterial cell. Therefore, prokaryotes have developed comprehensive mechanisms for the uptake and assimilation of nitrogen to provide a sufficient nitrogen supply for cell growth and maintenance of intracellular metabolism. Ammonium (NH_4^+) is the preferred nitrogen source of most bacteria as it allows for a high growth rate compared to other nitrogen sources and a number of energy-dependent uptake mechanisms have been described [11,20] (for reviews see [21-23]). However, its uncharged form ammonia (NH_3) can also rapidly diffuse across the membrane and serve as nitrogen donor when extracellular ammonia concentrations are high. Both ammonium and ammonia are incorporated into glutamate and glutamine via two different assimilation mechanisms and they subsequently serve as nitrogen donors for the intracellular metabolism of bacteria. Furthermore, many bacteria can transport and metabolise alternative nitrogen sources such as nitrate, peptides, amino acids and nucleotides that are converted to ammonium. Previous work has highlighted the importance of these alternative nitrogen sources for growth and survival of bacteria, particularly when primary nitrogen sources are exhausted [18,19,24,25].

1.3.1. Uptake and assimilation of nitrogen

Transport of nitrogen sources

Ammonium is the preferred nitrogen source of most bacteria and its incorporation into glutamate and glutamine generates nitrogen donors for biosynthetic processes in bacteria. The transport of ammonium ions across the bacterial cell membrane is conducted by transmembrane proteins belonging to the Amt/Mep family and nearly all bacteria encode for at least one copy of the *amt* gene [21]. However, the exposure of bacteria to various environmental conditions (e.g. nutrient limitation) resulted in the evolution of different strategies to ensure sufficient uptake of alternate nitrogen sources and various nitrogen transport systems in actinomycetes are going to be discussed.

Some organisms belonging to the actinomycetes such as *Corynebacterium glutamicum* (Figure 1.2A [red]) and *M. smegmatis* (Figure 1.3A [red]), also possess paralogues of *amt*. Most of the knowledge about ammonium transport systems in actinomycetes is limited to *C. glutamicum*, where these transmembrane proteins have been extensively characterised [20,26-28]. The AmtA protein functions as a high-affinity uptake system for methylammonium (K_m = 44 µM), an ammonium analogue commonly used for transport measurements, while its paralogue AmtB was described as an ammonium-specific permease and showed only a very low affinity for methylammonium (K_m > 3mM) in *C. glutamicum* (Figure 1.2A [red]) [20,28].

Ammonium transporters in Mycobacteria

M. smegmatis encodes for three paralogues of the *amt* gene, *amtA* (*msmeg_4635*), *amtB* (*msmeg_2425*) and *amt1* (*msmeg_6259*) (Figure 1.3A [red]), while *M. tuberculosis* only encodes a homologue of *amtB* (*Rv2920c*, also referred to as *amt*). However, no biochemical data are available to clarify the role of the single *amt* genes in ammonium uptake in mycobacteria so far. In a considerable number of bacteria and archaea, the ammonium transporter *amtB* is clustered together with the regulatory protein P_{II}, encoded by *glnK*, and a functional interaction between AmtB and GlnK has been previously shown for several organisms [29-32]. Unlike *amtB*, the putative ammonium transporter *amtA* is not encoded within an operonic structure in *M. smegmatis*, while the gene *amt1* lies in a conserved region that has apparently originated from pseudomonads and might have been acquired via horizontal gene transfer [11]. This conserved region encodes different functional domains of a class III glutamine synthetase (*msmeg_6260*), a class II glutamine amidotransferase (*msmeg_6261*) and domains of a glutamate synthase (*msmeg_6262* and *msmeg_6263*), however, their physiological function has yet to be defined [11].

Within the actinomycetes, the gene *glnD* is in close proximity to the *amtB-glnK* operon. This gene encodes for a GlnK modifying/demodifying enzyme (uridylyl-transferase/uridylyl-removing enzyme in enterobacteria, adenylyl-transferase/adenylyl-removing enzyme in *Streptomyces coelicolor*, *C. glutamicum* and *M. tuberculosis*) [31,33,34]. Under nitrogen excess, GlnK

interacts with AmtB almost exclusively via a long surface loop (T-loop), the tip of which inserts into the cytoplasmic pore exit of AmtB and therefore blocks ammonium transport. This does not occur under nitrogen limitation, as modification of GlnK by GlnD at a conserved tyrosine residue prevents this complex formation, allowing ammonium transport through AmtB [35,36].

Acquisition of other nitrogen sources

Transcriptional regulation of the different ammonium transporters in actinomycetes is controlled by GlnR and AmtR. The ammonium transporter AmtB in *S. coelicolor* is regulated by GlnR (Figure 1.4A [red]), like all three paralogues of the *amt* gene in *M. smegmatis* (Figure 1.3A [red]) [37,38]. The two paralogues *amtA* and *amtB* in *C. glutamicum* are under control of AmtR (Figure 1.2A [red]), however, they are not transcribed under nitrogen surplus, but their transcription is induced upon nitrogen starvation [39,40]. Comparison of the GlnR and AmtR regulon in these actinomycetes reveals remarkable differences with regard to the uptake of alternate nitrogen sources. While in *S. coelicolor* the only gene reported so far to be under control of GlnR is *amtB*, the regulon of AmtR in *C. glutamicum* comprises four other nitrogen uptake mechanisms, i.e. uptake systems for glutamate, creatinine, oligo-peptides and urea [39,41]. The urea ABC transporter has been shown to have a crucial role for urea uptake in *C. glutamicum* (Figure 1.2F [cyan]) [42].

By far the most complex response with regard to the uptake of alternative nitrogen sources was reported for *M. smegmatis*, where ten nitrogen uptake systems are under transcriptional control of GlnR [43]. Transcription of a number of ABC transporters and secondary transport systems for amino acids, nucleotides and nitrate were consistently induced upon nitrogen starvation (Figure 1.3F [cyan]). However, their role and importance under these conditions has yet to be defined. One urea ABC transport system is annotated in the genome of *M. smegmatis*, but no homologue was found in pathogenic mycobacteria. A nitrate uptake system encoded by *narK* is present in both fast- and slow-growing mycobacterial species [44]. Furthermore, enzymes for the metabolism of alternative nitrogen sources are also annotated in the genomes of mycobacteria (e.g. urease, nitrate/nitrite reductase) [44,45].

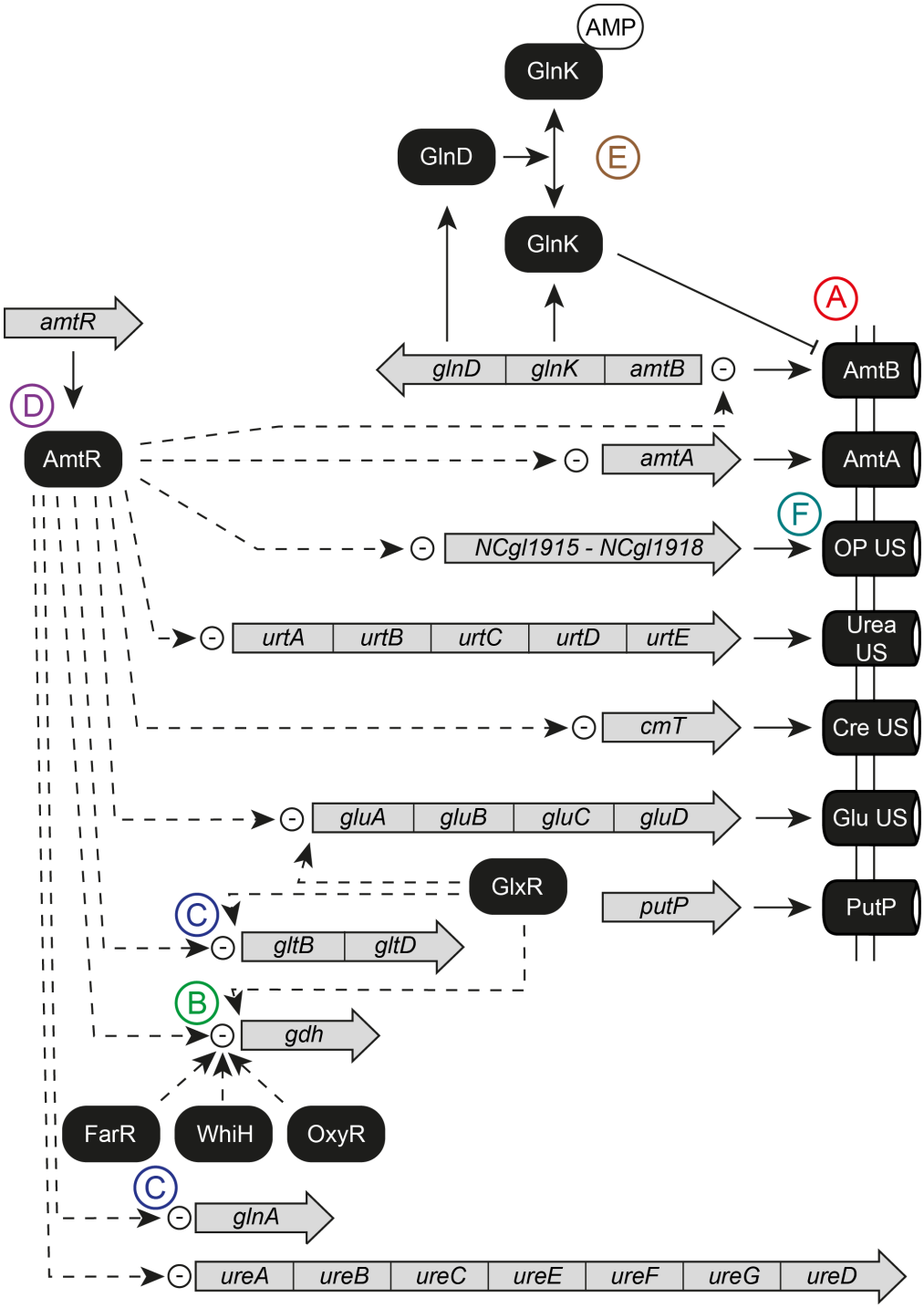


Figure 1.2. General response to nitrogen limitation and the regulatory network of AmtR in *C. glutamicum*. Grey arrows indicate genes and their organisation, solid black arrows indicate transcription into proteins (black boxes) and dashed black arrows indicate negative (minus) regulatory effect of the respective transcriptional regulator on gene expression. **[A]**: NH_4^+ transport systems; **[B]**: GDH enzyme; **[C]**: GS/GOGAT enzymes; **[D]**: AmtR regulon; **[E]**: P_{II} protein; **[F]**: other nitrogen transport systems. AMP: adenosine monophosphate; Cre: creatinine; Glu: glutamate; OP: oligopeptides; PutP: Na^+/proline symporter; US: uptake system; for further details, see text. Shown response is illustrative, but not exhaustive.

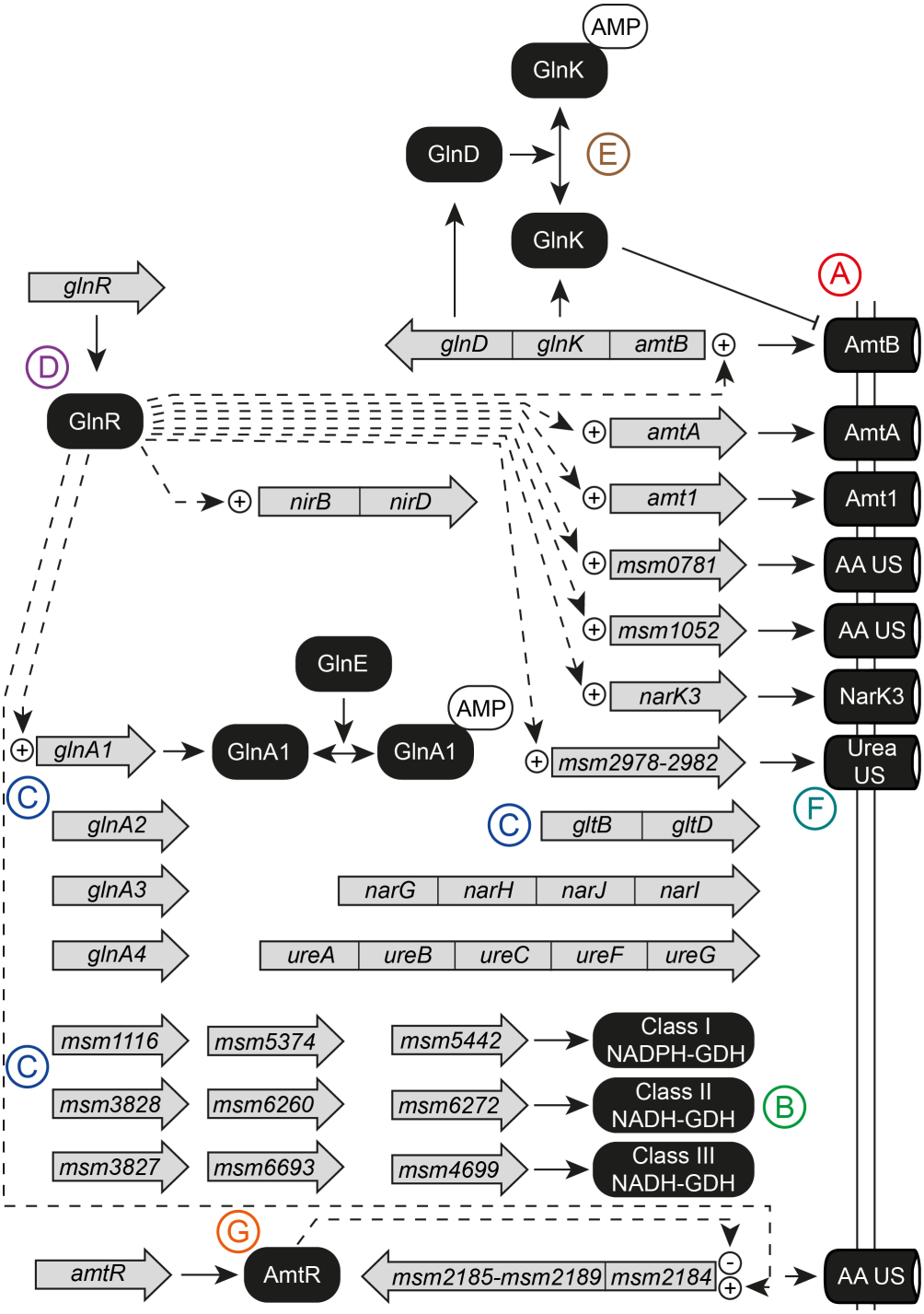


Figure 1.3. General response to nitrogen limitation and the regulatory network of GlnR and AmtR in *M. smegmatis*. Grey arrows indicate genes and their organisation, solid black arrows indicate transcription into proteins (black boxes) and dashed black arrows indicate positive (plus) or negative (minus) regulatory effect of the respective transcriptional regulator on gene expression. [A]: NH_4^+ transport systems; [B]: GDH enzymes; [C]: GS/GOGAT enzymes; [D]: GlnR regulon; [E]: P_{II} protein; [F]: other nitrogen transport systems; [G]: AmtR regulon. AA: amino acid; AMP: adenosine monophosphate; NarK3: nitrate uptake system; for further details, see text. Shown response is illustrative, but not exhaustive.

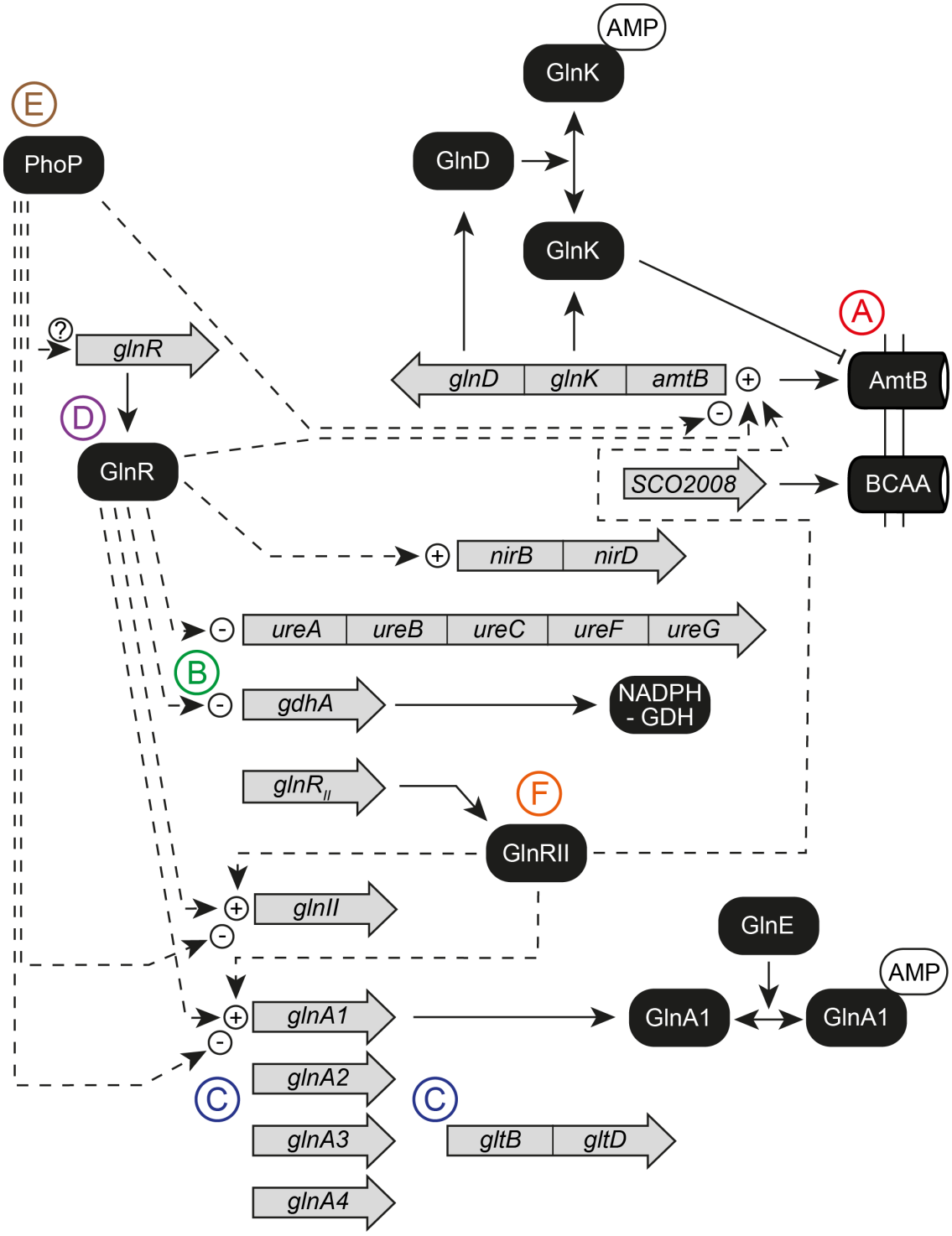


Figure 1.4. General response to nitrogen limitation and the regulatory network of GlnR in *S. coelicolor*. Grey arrows indicate genes and their organization, solid black arrows indicate transcription into proteins (black boxes) and dashed black arrows indicate positive (plus) or negative (minus) or unknown (?) regulatory effect of the respective transcriptional regulator on gene expression. **[A]**: NH_4^+ transport system; **[B]**: GDH enzyme; **[C]**: GS/GOGAT enzymes; **[D]**: GlnR regulon; **[E]**: PhoP regulon; **[F]**: $GlnR_{II}$ regulon. AMP: adenosine monophosphate; BCAA: branched chain amino acid transporter; for further details, see text. Shown response is illustrative, but not exhaustive.

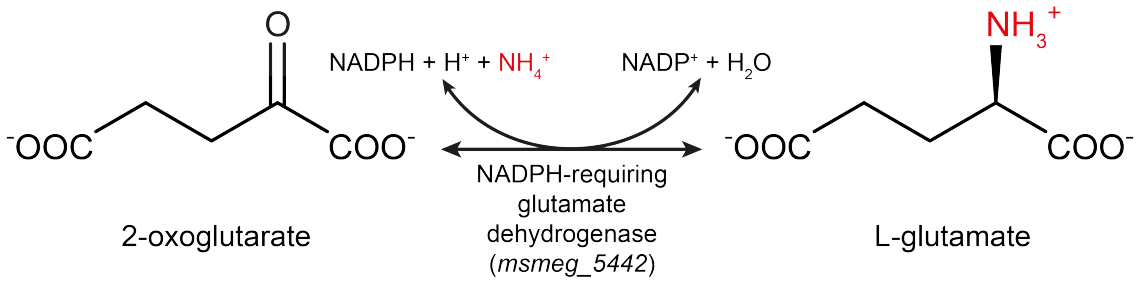
Glutamate dehydrogenase

Nitrogen can be incorporated into glutamine and glutamate via two physiologically distinct enzymes, the ATP-independent NAD(P)H-requiring glutamate dehydrogenase (GDH) and the ATP-dependent NADPH-requiring glutamine synthetase/glutamate synthase (GS/GOGAT). The GDH catalyses the reversible NAD(P)H-requiring amination of 2-oxoglutarate with NH_4^+ to form glutamate (Figure 1.5). This reaction is bioenergetically more favourable compared to the GS/GOGAT ammonium assimilation system. No utilisation of ATP is required to drive this reaction, therefore making GDH the preferred ammonium assimilation system under nitrogen rich conditions, despite its lower affinity for ammonium compared to the GS/GOGAT system (for reviews see [46,47]).

Cofactor specificity in GDH

As a reversible enzyme, GDH has the ability to either catalyse an amination (anabolism) or deamination (catabolism) and the physiological function of the GDH enzyme in an organism depends on numerous factors, such as the ecological niche, nutrient availability and the physiological state [48]. In general, GDH enzymes of prokaryotes discriminate between NADH and NADPH and this cofactor is a crucial indicator for the function of GDH enzymes. NADPH-requiring GDH enzymes (NADPH-GDH) are primarily found in organisms that rely on scavenging inorganic nitrogen sources, such as ammonium and nitrate. In these organisms, they catalyse the amination of 2-oxoglutarate to synthesise glutamate, which allows them to balance their intracellular $NADP^+$/NADPH pool for reductive biosynthesis reactions [49-52]. Organisms that live in an environment with a large supply of organic nitrogen sources often encode only for a NADH-requiring GDH (NADH-GDH) [53-55]. This enzyme has primarily a catabolic function, which bacteria use to get rid off excess ammonium and to access the carbon skeleton of 2-oxoglutarate. Interestingly, some organisms such as *M. smegmatis*, *Pseudomonas aeruginosa*, *Psychrobacter sp. TAD1* and *Thiobacillus novellus* encode for NADH- and NADPH-dependent GDH enzymes [52,56-59].

[A]



[B]

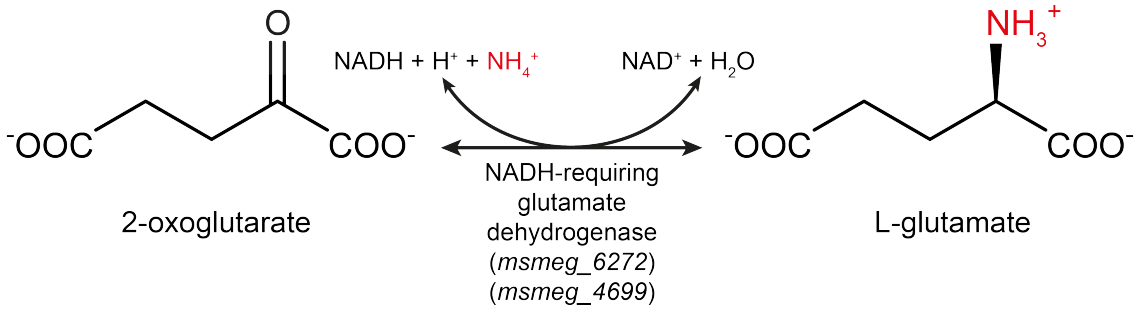


Figure 1.5. The glutamate dehydrogenase reaction. Depicted is the enzymatic reaction driven by the NADPH-requiring **[A]** and NADH-requiring **[B]** glutamate dehydrogenase (GDH). The GDH catalyses the reversible NAD(P)H-requiring amination of 2-oxoglutarate with NH_4^+ to form glutamate. GDH enzymes of prokaryotes discriminate between NADH or NADPH and this cofactor is a crucial indicator for the function of GDH enzymes *in vivo*. Genes in the genome of *M. smegmatis* that encode for NADPH-dependent and NADH-dependent GDH enzymes are indicated in brackets.

Categorisation of GDH enzymes

The GDH enzymes can be grouped into three distinct subgroups according to their coenzyme specificity, subunit composition and size. The primary class of GDH enzymes are NADPH-requiring, form hexameric structures (subunits ~48 kDa) and are generally involved in nitrogen assimilation [50]. The second class consists of NADH-requiring enzymes with a tetrameric structure (115 kDa per subunit). This class of GDH enzymes was firstly isolated from *Neurospora crassa*, a fungus belonging to the phylum *Ascomycota* [60]. The most recently identified third class of GDH enzymes is also NADH-specific and comprises a hexameric GDH (subunits ~183 kDa) and was first described in *Streptomyces clavuligerus* [55]. Mammalian GDH enzymes are less diverse and therefore more versatile in their function, that allows these proteins to interact with both NADPH and NADH [48].

GDH in Actinobacteria

C. glutamicum (Figure 1.2B [green]) and *S. coelicolor* (Figure 1.4B [green]) encode one NADPH-GDH, which plays a major role in ammonium assimilation when the ammonium concentration is high. However, a *gdh* mutant of *C. glutamicum* revealed that a functional GDH is not required for growth or glutamate production in *C. glutamicum*, although the growth rate of a *gdh*-deficient strain was dramatically reduced at ammonium concentrations higher than 10 mM [61]. In contrast to *C. glutamicum* and *S. coelicolor*, *M. smegmatis* encodes for three copies of GDH enzymes in its genome (Figure 1.3B [green]), a class I NADPH-GDH (*msmeg_5442*), a class II NADH-GDH (*msmeg_6272*) and a class III NADH-GDH (*msmeg_4699*) (Figure 1.5). The class I and III enzymes encode for functional GDH enzymes that have been biochemically characterised, however, the presence of a functional class II GDH has to be shown in the future [52,59]. In a comprehensive study, Harper *et al.* showed that expression of the class II and class III enzymes were elevated two hours after exposure to nitrogen starvation [52]. In contrast, transcription of the class I GDH gene *msmeg_5442* was downregulated within the first 30 minutes of nitrogen limitation, whereas its activity increased concurrently. This implies an

additional post-translational regulatory mechanism of the class I NADPH-GDH in *M. smegmatis*, as previously reported for the GDH enzymes of *C. glutamicum* and *M. tuberculosis* [51,52,59,62]. Enzyme activity studies with NADPH-GDH and NADH-GDH of *M. smegmatis* revealed a high aminating activity for both enzymes, but only the NADPH-GDH was functional in the deaminating reaction. The activity towards the deaminating reaction of the NADPH-dependent GDH altered in response to nitrogen availability, while the aminating reaction was not affected. However, amination activity of the NADH-dependent GDH enzymes dropped under nitrogen excess [52].

In *M. smegmatis*, none of the GDH-encoding genes (*gdh*) are under control of the global nitrogen regulator GlnR (Figure 1.3B [green] and Figure 1.3D [purple]). This is in contrast to *S. coelicolor* (Figure 1.4B [green] and Figure 1.4D [purple]) and *C. glutamicum* (Figure 1.2B [green] and Figure 1.2D [purple]), where *gdhA* is regulated by GlnR or AmtR, respectively [41,63]. Besides the regulatory effect of the transcriptional repressor AmtR on the regulation of *gdh* in *C. glutamicum*, expression of the *gdh* gene is also regulated by FarR (regulator of arginine metabolism), GlxR (central carbon metabolism), WhiH and OxyR (both unknown function) (Figure 1.2B [green]) [63-65]. Although the GDH enzyme is known to play a key role in ammonium assimilation under nitrogen surplus, in *C. glutamicum* both the expression of *gdh* and the GDH activity are upregulated in response to nitrogen limitation [63]. This was previously reported only for the GDH enzyme of *Ruminococcus flavefaciens* so far and the authors proposed a mechanism to maintain a sufficient glutamate pool through GDH activity in order to keep the glutamate:glutamine ratio balanced [66].

Regulation of GDH

Determination of enzymatic structures of different GDH enzymes has allowed elucidation of post-translational regulatory mechanisms through protein-protein interaction and allosteric regulation (for review see [67]). Adenosine monophosphate (AMP) is an essential activator of GDH activity in *Streptomyces clavuligerus* [55]. Furthermore, the amino acid L-arginine has a positive allosteric effect on the GDH enzyme in *Pseudomonas aeruginosa*,

whereas increasing concentrations of citrate cause allosteric inhibition [68]. The single NADH-GDH (*Rv2476c*) of *M. tuberculosis* shares a 71% amino acid identity with the class III enzyme of *M. smegmatis* and a post-translational regulatory cascade initiated by the serine/threonine protein kinase G (PknG; *Rv0410c*) has been reported to modulate function of NADH-GDH in *M. tuberculosis*, as previously described for *C. glutamicum* [59]. There, PknA phosphorylates PknG, which then modulates glutamate metabolism via phosphorylation of OdhI, a 2-oxoglutarate dehydrogenase complex [62]. In *M. tuberculosis*, autophosphorylation of PknG in response to an unknown signal results in phosphorylation and consequent inactivation of the glycogen accumulation regulator GarA (*Rv1827*), the homologue of OdhI in mycobacterial genomes [59]. GarA contains a forkhead-associated domain, suggesting involvement in a serine/threonine kinase signal transduction pathway, but its function in *M. tuberculosis* has not yet been defined, although it was reported to affect glycogen metabolism in the closely related organism *M. smegmatis* [69]. Unphosphorylated GarA inhibits activity of the NADH-GDH in *M. tuberculosis* by direct protein-protein interaction shown by affinity purification. A protein-protein interaction of unphosphorylated GarA was also reported for the NADH-GDH homologue MSMEG_4699 of *M. smegmatis*, suggesting a similar mechanism in fast-growing mycobacteria [59].

Ammonium assimilation in *M. tuberculosis* usually occurs via the GS/GOGAT pathway, while GDH is mostly catabolic and results in glutamate degradation into ammonium and 2-oxoglutarate [16,70]. The glycogen accumulation regulator A (GarA) plays a key role in regulation of GDH and GOGAT activity in *M. tuberculosis*. Unphosphorylated GarA inhibits GDH and stimulates activity of GOGAT, resulting in glutamate synthesis [16]. In *M. bovis*, the GDH enzyme is important for the utilisation of glutamate as the sole nitrogen source to ensure nitrogen supply of the cell via transamination reactions and an important role for GDH in survival of the *M. tuberculosis* complex during latency has also been proposed [52,70]. The low energy requirements of GDH may allow a sufficient nitrogen supply when energy preservation is necessary, however, further work is required to fully ascertain the function of GDH enzymes in mycobacteria [52].

Glutamine synthetase/Glutamate synthase (GS/GOGAT)

The GS/GOGAT system comprises two separate enzymes, the glutamine synthetase (GS) and the glutamate synthase (GOGAT), which form a cyclic machinery to incorporate ammonium into L-glutamate (Figure 1.6). This system is the preferred pathway of ammonium assimilation under nitrogen limitation due to its much higher affinity for ammonium than the GDH. The GS is encoded by the gene *glnA* and is a pivotal enzyme in nitrogen metabolism that catalyses the ATP-dependent incorporation of ammonium into glutamate to form glutamine (Figure 1.6). This assimilatory pathway utilises around 15% of the intracellular ATP pool, thus, the activity of GS is subject to regulation at both the transcriptional and enzyme activity levels to prevent energy wastage under nitrogen surplus [47,71]. The GOGAT enzyme consists of two subunits, a small subunit encoded by *gltB* and a large subunit encoded by *gltD*, that drive the reductant-dependent reaction of L-glutamine and 2-oxoglutarate to form two molecules of L-glutamate (Figure 1.6) [72].

Categorisation of GS enzymes

The glutamine synthetase is a ubiquitous enzyme that has been described in all three domains of life (for review see [73,74]). GS enzymes occur in three distinct types, GS-I, GS-II and GS-III. These show variation in subunit size, composition and the type of covalent modification, depending on the organism and its ecological niche [46].

The genomes of bacteria and archaea commonly harbor genes that encode for GS-I, an enzyme that consists of twelve identical subunits with an approximate molecular mass of around 55 kDa each, that are arranged in two planar hexamers and complex two Mn^{2+} ions in their active site [75]. GS-I enzymes can be subdivided into GS-I-α and GS-I-β enzymes, depending on whether they undergo post-translational modification that regulates their enzyme activity [46]. The GS-I-α enzyme, which lacks a post-translational regulatory mechanism, is characteristic for archaea and low G+C Gram-positive bacteria, while GS-I-β enzymes occur in the genomes of all other bacteria and apart from cyanobacteria their activity is subject to post-translational regulation

through adenylylation via GlnE [46,73]. Recent studies also reported the remarkable finding of GS-I-like genes in higher plants and the vertebrate eye lens, that also form a dodecameric structure similar to prokaryotic GS-I enzymes, but did not show enzymatic activity [76,77].

The GS-II enzymes are encoded by the gene *glnII* and their structure differs significantly from GS-I enzymes. Structural analysis of GS-II enzymes revealed a decameric structure that is formed by two pentameric rings with a molecular mass of approximately 46 kDa per subunit [78,79]. They are mainly present in eukaryotes, but genes encoding these enzymes have also been identified in some bacterial species that also encode for GS-I enzymes, such as *Frankia*, *Rhizobium*, *Agrobacterium* and *Streptomyces* [80-82]. This co-occurence of GS-I and GS-II enzymes is proposed to be due to a gene duplication that preceeded the divergence of prokaryotes and eukaryotes. The finding of GS-I-like enzymes in eukaryotes supports the hypothesis of a paralogous evolution of GS-I and GS-II genes [46,73,76,77,83].

The gene *glnN* encodes for GS-III enzymes that have been found in a few bacterial organisms and that was also identified in the genome of the amoeba *Dictyostelium discoideum* [84-88]. GS-III enzymes consist of two hexameric rings that form a dodecameric structure with a molecular mass of around 75 kDa per subunit and are very distinct to GS-I and GS-II enzymes, mainly due to differences in their quaternary structure [88]. Surprisingly, the two hexameric rings comprising the GS-III dodecamer have been found to associate across opposite interfaces relative to GS-I and GS-II enzymes [88].

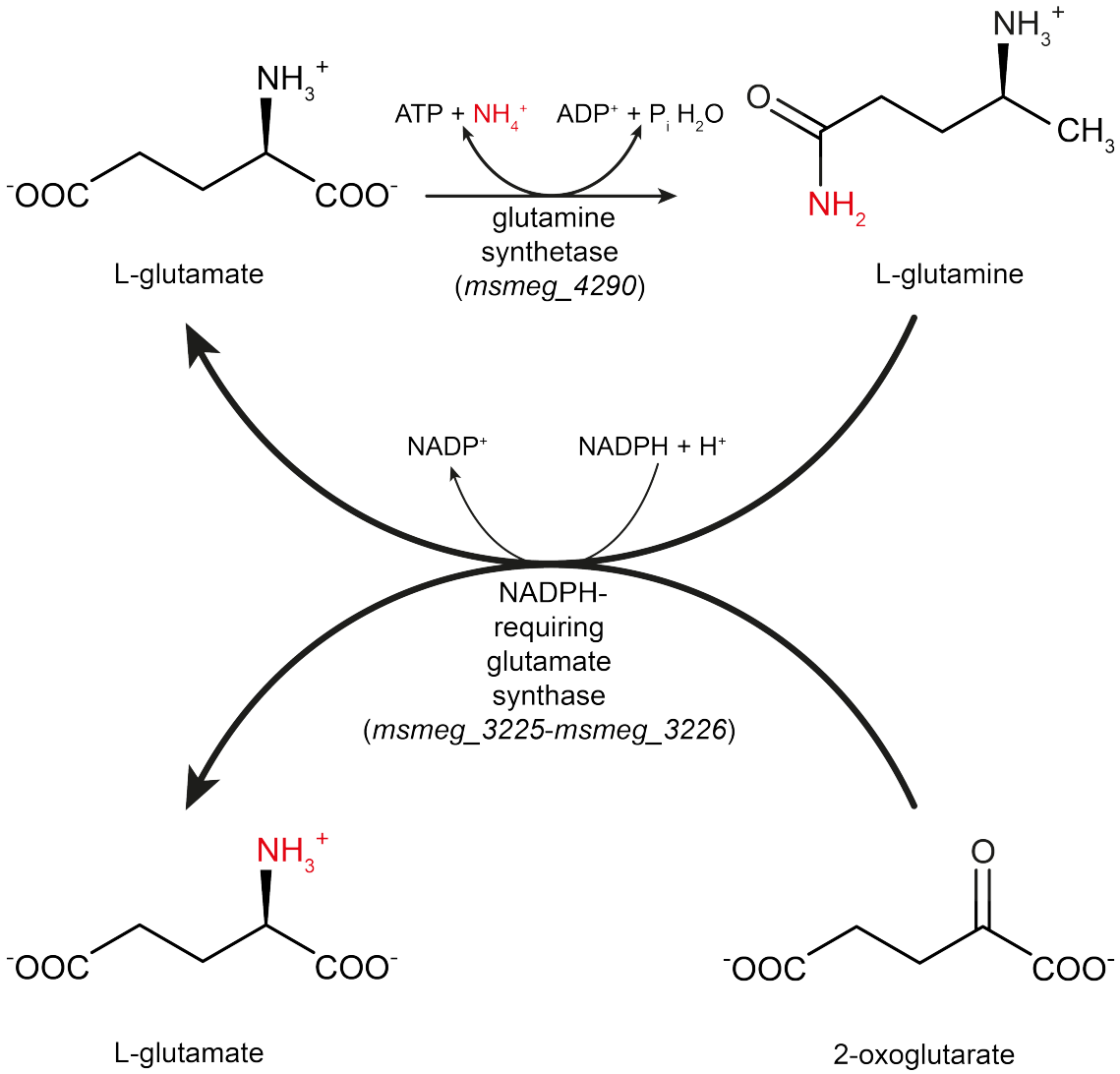


Figure 1.6. The glutamine synthetase/glutamate synthase (GS/GOGAT) reaction. Depicted is the enzymatic reaction driven by the glutamine synthetase (GS) and the NADPH-requiring glutamate synthase (GOGAT). The GS/GOGAT system is the preferred pathway of ammonium assimilation under nitrogen limitation due to its high affinity for ammonium. The GS catalyses the ATP-dependent incorporation of ammonium into glutamate to form glutamine and GOGAT drives the reductant-dependent reaction of L-glutamine and 2-oxoglutarate to form two molecules of L-glutamate. Genes of *M. smegmatis* that encode for the primary GS and GOGAT enzymes are indicated in brackets. For a full list of homologous enzymes of GS and GOGAT, see Table 1.1.

GS in Mycobacteria

Some organisms, like *Escherichia coli* only encode for a single copy of a GS-I enzyme, however, the genomes of many organisms comprise a number of genes that encode for multiple GS enzymes [47]. The genome of *S. coelicolor* harbors four GS-I enzymes and one GS-II enzyme, but only the GS-I and GS-II enzymes encoded by *glnA1* and *glnII*, respectively, have been reported to be involved in nitrogen metabolism, while the other homologs of GS-I (GlnA2, GlnA3, GlnA4) do not contribute significantly to glutamine synthesis (Figure 1.4C [blue]) [89].

Mycobacterial genomes also harbor multiple copies of GS enzymes (Table 1.1) [11]. The genes *glnA1* (*msmeg_4290* / *Rv2220*) and *glnA2* (*msmeg_4294* / *Rv2222c*) are found in all mycobacterial genomes (Figure 1.3C [blue]), where they are clustered in a conserved region together with *glnE*, an adenylyltransferase that regulates glutamine synthetase activity. The genes *glnA3* (*msmeg_3561* / *Rv1878*) and *glnA4* (*msmeg_2595* / *Rv2860c*) are present in all mycobacterial genomes (Figure 1.3C [blue]), except *M. leprae* [11,90]. *M. smegmatis* contains further two enzymes that are homologous to GlnA4 (*msmeg_1116*, *msmeg_3828*), two GS-like proteins (*msmeg_3827*, *msmeg_5374*), as well as two GS enzymes (*msmeg_6260*, *msmeg_6693*) that show a sequence homology between 35%-45% to the GS-III enzymes of *Pseudomonas*, *Synechococcus* and *Agrobacterium* species (Table 1.1), (Figure 1.3C [blue]) [11]. Physiological and enzymatic studies with the four GS enzymes of *M. tuberculosis* revealed that only GlnA1 is essential for *in vitro* growth, which is supported by the observation that GlnA1 is the only GS enzyme that displays enzymatic activity [91]. The physiological role and function of the various other GS-encoding genes, especially in the genome of *M. smegmatis*, remains to be investigated. Interestingly, in pathogenic mycobacteria GS is highly abundant in the extracellular milieu, while the release of GS is very little in fast-growing mycobacteria [92,93]. Previous studies have shown that accumulation of extracellular proteins in mycobacteria strongly depends on nutrient availability, however, it is also suggested that extracellular GS modulates the ammonium level in the phagosome, which might influence the phagosomal pH [52,93].

Regulation of GS

The glutamine synthetase is a strictly regulated enzyme on both transcriptional and post-translational levels to avoid unnecessary enzyme activity of GS. Transcription of the primary GS-encoding gene is under control of the major nitrogen regulatory system that mediates transcription in response to changing nitrogen levels (e.g. NtrBC in *Enterobacteria*, GlnR in *S. coelicolor* and *M. smegmatis*, AmtR in *C. glutamicum*) [37,39,43]. Interestingly, the *glnA* gene of *E. coli* is transcribed from two different promoters, one of which is under transcriptional control of the cAMP receptor protein providing a novel regulatory linkage between carbon metabolism and nitrogen assimilation [94].

Generally, bacterial GS is also subject to a post-translational regulatory mechanism through adenylylation via the adenylyl transferase GlnE, which allows a more precise adjustment to the cellular requirements. GlnE modulates GS activity through adenylylating and deadenylylating domains in response to multiple extracellular and intracellular stimuli with each GS monomer in the homododecamer subject to modification by GlnE at a conserved tyrosine residue [95]. The number of adenylylated subunits directly reflects the GS activity and its ability to assimilate ammonium into glutamate to form glutamine. Stepwise adenylylation of GS through GlnE results in a gradual decrease of the ammonium assimilation activity of GS [96,97].

The glutamate synthase

The second enzyme in this cyclic machinery is the glutamate synthase (GluS) that, like GS, is also ubiquitous in all three domains of life. The GluS of bacteria consists of two subunits, a small subunit encoded by *gltB* and a large subunit *gltD*, that drive the reductant-dependent reaction of L-glutamine and 2-oxoglutarate to form two molecules of L-glutamate (Figure 1.6). Based on their biochemical properties, GluS enzymes can be categorised into three distinct classes, NADPH-requiring GluS enzymes (NADPH-GluS), NADH-requiring GluS enzymes (NADH-GluS) and ferredoxin-requiring GluS enzymes (Fd-GluS) (for review see [72]).

The NADPH-GluS enzyme is the most abundant GluS in bacteria and is composed of two separately encoded subunits, GltB and GltD, that contain several cofactors and iron-sulphur clusters [72]. The large subunit GltB (160 kDa) harbors an FMN cofactor and catalyses the reaction of L-glutamine and 2-oxoglutarate to L-glutamate. The required reducing equivalents are provided by the small subunit GltD (~52 kDa), an FAD-dependent NADPH oxidoreductase [72]. GltD transfers reducing equivalents from NADPH through an FAD cofactor and three iron-sulphur clusters to the FMN cofactor in the large subunit, where they are utilised to drive the aforementioned reaction (Figure 1.6). Some archaea encode for homologs of the small or large subunits of GluS, e.g. *Pyrococcus* species only harbor the small subunit, which can function as GluS after assembly to a homotetramer and the genomes of *Methanococcus jannaschii* and *Archaeoglobus fulgidus* revealed the presence of truncated versions of GltB that show homology to the central part of bacterial GltB [98-101].

The NADH-GluS and the Fd-GluS enzymes are mostly found in eukaryotes, e.g. higher plants (Fd-GluS), nonphotosynthetic tissue of plants (NADH-GluS) and fungi (NADH-GluS), however, they have also been detected in cyanobacteria (Fd-GluS). Both enzymes are composed of a single polypeptide with an approximate molecular mass of 165 kDa (Fd-GluS) and 200 kDa (NADH-GluS), respectively and it is currently hypothesized that these enzymes are derived from a *gltB-gltD* gene fusion [72].

The glutamate synthase enzyme in Mycobacteria

The GluS-encoding genes *gltB* and *gltD* are highly conserved in mycobacterial genomes and their transcription is induced upon nitrogen starvation [11,71]. However, this is not regulated by GlnR in *M. smegmatis* (Figure 1.3C [blue] and Figure 1.3D [purple]), unlike in *C. glutamicum*, where expression of *gltBD* is under control of AmtR (Figure 1.2C [blue] and Figure 1.2D [purple]) and is repressed under nitrogen surplus [43,102]. Furthermore, no regulation of GluS activity via post-translational modification in response to changing nitrogen availability has been reported so far, in either *C. glutamicum*, *M. smegmatis* or *M. tuberculosis*, thus leaving it uncertain how

mycobacteria regulate GluS transcription and activity [103]. Interestingly, *M. smegmatis* harbors additional copies of *gltB* (*msmeg_5594*, *msmeg_6263*, *msmeg_6459*) and *gltD* (*msmeg_6262*, *msmeg_6458*) in its genome (Table 1.1). However, they are also not regulated by GlnR and no further regulatory mechanism has been identified yet and their physiological role under nitrogen limitation has to be shown (Figure 1.3C [blue]) [11].

Table 1.1. Homologs of the GS- and GOGAT-encoding genes in the genome of *M. smegmatis*. Listed are all annotated homologs of *gln*, *gltB* and *gltD* with their gene name, function, locus tags for *M. smegmatis* mc^{2}155 and *M. tuberculosis* H37Rv if present and references with studies for representative well-studied homologs. Table adapted from [11].

Gene	Function	*M. smegmatis* mc^{2}155 locus	*M. tuberculosis* H37Rv locus	Reference
glnA1	NH_4^+ assimilation	msmeg_4290	Rv2220	[91]
glnA2	unknown; similar to GS-I	msmeg_4294	Rv2222c	[91]
glnA3	unknown; similar to GS-I	msmeg_3561	Rv1878	[91]
glnA4	unknown; similar to GS-I	msmeg_2595	Rv2860c	[91]
glnA	unknown; homolog to *glnA4*	msmeg_1116	n/a[a]	[91]
glnA	unknown; homolog to *glnA4*	msmeg_3828	n/a	[91]
glnA	unknown	msmeg_3827	n/a	*E. coli* K12 putative glutamine synthetase (*b1297*)
glnA	unknown	msmeg_5374	n/a	*R. palustris glnA* (RPA0984)
glxA	unknown; similar to GS-III	msmeg_6260	n/a	[11]
glnA	unknown; similar to GS-III	msmeg_6693	n/a	*A. tumefaciens* glnA *(Atu4230*)
gltB	glutamate synthase; small subunit	msmeg_3225	Rv3859	[102]
gltD	glutamate synthase; large subunit	msmeg_3226	Rv3858	[102]
gltB	unknown; similar to *gltB*	msmeg_5594	n/a	[102]
glxD	unknown; similar to *gltB*	msmeg_6263	n/a	[11]
gltB	unknown; similar to *gltB*	msmeg_6459	n/a	[102]
glxC	unknown; similar to *gltD*	msmeg_6262	n/a	[11]
gltD	unknown; similar to *gltB*	msmeg_6458	n/a	[102]

[a] n/a: not available

1.3.2. Regulation of nitrogen metabolism

The P_{II} protein

The P_{II} protein is a central player in regulation of bacterial nitrogen metabolism. It generally consists of a functionally versatile P_{II} signal transduction protein and a P_{II} protein-modifying enzyme that together modulate a wide range of regulatory pathways related to the uptake, assimilation and metabolism of nitrogen. The regulatory function of the P_{II} protein is finely adjusted by two distinct mechanisms: the reversible covalent modification of a conserved residue on the T-loop and allosteric binding of effector molecules. The P_{II} protein consists of three monomers and all of them can be subject to uridylylation by GlnD, which depends on the intracellular glutamine:2-oxoglutarate ratio. A ratio in favor of glutamine signals nitrogen excess, whereas a ratio in favor of 2-oxoglutarate signals nitrogen deficiency. Subsequently, an increasing uridylylation of GlnB has a continuously decreasing effect on the NtrB-dependent dephosphorylation of NtrC in *E. coli* [104]. Uridylylated GlnB also stimulates deadenylylation of the GS enzyme under nitrogen depletion, whereas GlnB in its unmodified form has the opposite effect [22]. The intracellular stimulus of both P_{II} homologs is 2-oxoglutarate, a well-described indicator of nitrogen deficiency that acts directly on GlnB. The homotrimeric P_{II} protein contains three binding sites for 2-oxoglutarate and three binding sites for ATP, respectively, that can bind to both its unmodified and uridylylated form in a synergistic manner. However, binding of one molecule of 2-oxoglutarate exerts a strong negative cooperativity on the binding of further molecules of 2-oxoglutarate to the P_{II} protein, an effect that is not observed for ATP binding [22,105].

The P_{II} protein is an extensively studied enzyme and its physiological, biochemical and structural properties are largely understood (for review see [22,106,107]). A phylogenetic analysis revealed that P_{II} proteins cluster into seven major groups, depending on the organisms in which they occur [32]. The molecular mass of a P_{II} monomer is around 13 kDa. It was firstly described as a regulatory mechanism of the GS-I enzyme in *E. coli* and since then, different intracellular stimuli, further interaction partners and several high-resolution structures have been reported [30,108-112].

An X-ray structure revealed that the P_{II} protein is a homotrimer that consists of a disk-like core with threefold symmetry. Each subunit is composed of two α-helices, six β-strands and three loop-structures (B-loop, C-loop, T-loop) that are important for the regulatory function of the P_{II} protein [106,113]. The T-loop harbors a tyrosine residue that is essential for the function of the P_{II} protein and which is subject to post-translational regulation via GlnD, while the B- and C-loops together with the T-loop form clefts wherein effector molecules can bind into [22].

E. coli contains two homologs of the P_{II} protein, GlnB and GlnK and both are post-translationally modified by the uridylyl transferase GlnD at a conserved tyrosine residue (serine residue in cyanobacteria) [108,114,115]. GlnB and GlnK are structurally very similar and can compensate for each other in several cases. However, a distinct expression profile and slightly different conformations of the three loops that are important for effector molecule binding and protein-protein interactions have been reported, thus resulting in distinct specificities for certain functions [106]. GlnB is constitutively expressed and is the primary nitrogen sensor under nitrogen surplus, whereas expression of GlnK is induced upon nitrogen starvation and regulates the expression of several genes in response thereto [116]. The gene encoding GlnB is often in close proximity to *glnA*, while *glnK* is encoded in a highly conserved operon together with the ammonium channel *amtB* [29]. Under nitrogen excess GlnB interacts with the trimeric AmtB protein via blocking of the cytoplasmic pore exit by the T-loop and thus disrupting ammonium transport, which has also been reported for *C. glutamicum* [31,35].

The P_{II} protein in Actinobacteria

Mycobacteria contain only one copy of a highly conserved P_{II} protein, GlnK, that is subject to post-translational modification via the adenylyl-transferase/adenylyl-removing enzyme, GlnD (Figure 1.3E [brown]) [34,117]. The genetic organization of *glnK*, *glnD* and *amtB* is the same as in *E. coli* and transcription of this operon in actinobacteria is under control of the global

nitrogen regulatory proteins, GlnR or AmtR [37,39,43]. Structural comparison of *M. tuberculosis* GlnK with the *E. coli* GlnK:AmtB complex also suggests a similar function of mycobacterial GlnK [117]. Actinobacteria do not comprise an NtrBC two-component system that is subject to post-translational regulation by the P_{II} protein. However, the P_{II} protein GlnK regulates the activity of the global nitrogen regulator AmtR and is essential for nitrogen control in *C. glutamicum* [39,118]. In other actinobacteria, such as *S. coelicolor*, *M. smegmatis* and *M. tuberculosis*, no interaction at a protein level between GlnK and the nitrogen regulatory system GlnR has been reported so far [34,37]. Furthermore, in *S. coelicolor*, *C. glutamicum*, *M. smegmatis* and *M. tuberculosis*, the activity of the adenylyltransferase/adenylyl-removing enzyme GlnE is not regulated by the P_{II} protein, GlnK [31,33,34]. There are still many open questions about the function of the P_{II} protein in mycobacteria and recent findings propose a more general role in stress protection of the cell [34]. Comparing gene expression of *M. tuberculosis* wild-type and a *glnD* deletion strain revealed differential expression of only 23 genes, including fumarate reductase, NADH dehydrogenase and an alkyl hydroperoxide reductase, however, none of the differentially expressed genes are predicted to be involved in nitrogen metabolism [34].

The nitrogen regulatory protein GlnR

The genomes of actinomycetes harbor two distinct global nitrogen regulatory proteins, AmtR and GlnR, that control gene expression in response to changing nitrogen levels. These two orphan regulatory proteins usually occur as sole global regulatory systems of nitrogen metabolism, however, some organisms encode for both AmtR and GlnR, but the physiological role of this co-occurence is not fully understood [11].

GlnR belongs to the OmpR-family of transcriptional regulators that harbor a C-terminal helix-turn-helix DNA-binding domain, which is highly conserved amongst other actinomycetes [41,119]. The OmpR protein was first described as part of the EnvZ-OmpR two-component system in *E. coli* that regulates expression of *ompF* and *ompC*. Both these genes have several OmpR-binding sites that consist of two 10-bp binding boxes each. These are located in their respective promoter regions and differ in their binding affinity for

OmpR. Upon phosphorylation by EnvZ, phosphorylated OmpR binds to these tandem repeats in a discontinuous manner, which is also referred to as “galloping model“, allowing a stepwise regulation of *ompF* and *ompC* transcription [120].

Transcriptional regulatory proteins that belong to the OmpR-family are usually subject to phosphorylation by a histidine sensor kinase [120]. Activation of the transcriptional regulator results from the phosphorylation of a highly conserved aspartate residue, as previously shown for related members of the OmpR-family [121,122]. This phosphorylation site is present in the GlnR protein of actinomycetes and was shown to be essential for the GlnR-mediated transcriptional response of *M. smegmatis* to nitrogen limitation, suggesting that GlnR might be subject to phosphorylation by a yet unidentified histidine kinase [43,71,123].

The function of GlnR in Actinobacteria

In actinomycetes, GlnR-dependent transcriptional regulation has been studied in *S. coelicolor* (Figure 1.4D [purple]) and *M. smegmatis* (Figure 1.3D [purple]) and initial studies showed that key enzymes of nitrogen metabolism are under control of GlnR. The GlnR regulon of *S. coelicolor* includes genes involved in uptake and assimilation of ammonium (*glnA*, *glnII*, *gdhA*, *amtB*), nitrite reduction (*nirBD*), urea degradation (*ureA*) and regulatory proteins (*glnK*, *glnD*) and their transcription is regulated in response to the intracellular nitrogen levels (Figure 1.4D [purple]) [37,41]. It was also reported, that GlnR is not subject to autoregulation, but transcription of *glnR* depends strongly on the availability of nitrogen [37,41]. Further work showed that GlnR acts as a bifunctional transcriptional regulator, activating transcription of the majority of its regulon under nitrogen starvation (e.g. *glnA*, *glnII*), but for a small subset of genes (e.g. *gdhA*, *ureA*) repression of gene expression occur under the same conditions (Figure 1.4D [purple]) [41]. This GlnR-mediated increase in transcription of *glnA* and *glnII* (both glutamine synthetases) as well as the repression of *gdhA* (glutamate dehydrogenase) in response to nitrogen limitation is physiologically reasonable. These three gene products are all

involved in ammonium assimilation and the synthesis of the nitrogen donors glutamine and glutamate. Under nitrogen-limiting conditions only the GS enzymes are able to assimilate ammonium, while the GDH with its higher K_M is ineffective in ammonium assimilation in nitrogen-starved cells. Interestingly, *S. coelicolor* encodes for another OmpR-like transcriptional regulator with high homology to GlnR, named $GlnR_{II}$, but is absent in the genome of *M. smegmatis*. $GlnR_{II}$ recognizes a similar promoter subset to GlnR, however, a *glnR$_{II}$* mutant was not auxotrophic for glutamine like a *glnR* deletion strain (Figure 1.4F [orange]) [37]. The GlnR regulon of *M. smegmatis* further comprises the glutamine synthetase *glnA2*, two additional ammonium transporters, *amtA* and *amt1* and several amino acid and nucleotide transporters, but not *gdhA* the gene encoding glutamate dehydrogenase (Figure 1.3D [purple]) [43,124]. Whether GlnR also regulates transcription of these genes in *M. tuberculosis* in response to nitrogen limitation is yet to be investigated.

Under nitrogen limitation GlnR regulates gene expression through binding at two subsequent GlnR-binding sites with the consensus sequence gTnAc-n_6-GaAAc-n_6 - GtnAC-n_6-GAAAc-n_6 (capitalization = high conservation) [41]. This consensus sequence was also identified in the majority of other actinobacteria, e.g. *S. coelicolor*, *M. smegmatis*, *Nocardia farcinica*, *Streptomyces avermitilis* and *Rhodococcus sp.*, but not in *C. glutamicum* [41]. A further consensus sequence of GlnR with the sequence TTCAC-n_6-GTAAC (capitalization ≠ conservation) was characterised recently in the promoter region of *amtB* that overlaps with the binding region of the phosphate response regulator, PhoP [125]. The same study revealed a competitive binding mechanism of PhoP and GlnR for transcriptional regulation of *amtB* [125]. Interestingly, it was previously reported that expression of *glnR*, *glnA*, *glnII* and *amtB* are under negative control of PhoP in *S. coelicolor*, however, this link between nitrogen and phosphate metabolism requires further investigation (Figure 1.4E [brown]) [126].

The nitrogen regulatory protein AmtR

The second global nitrogen regulatory protein in actinobacteria is the AmtR protein. This global repressor protein is the sole transcriptional regulator governing the nitrogen starvation-induced gene expression in corynebacteria, but can also be found in some genomes of *Streptomyces* (e.g. *S. avermitilis*) and *Mycobacterium* (e.g. *M. smegmatis*) that also encode for GlnR.

The AmtR protein belongs to the TetR-family of transcriptional regulators that contain a N-terminal helix-turn-helix motif responsible for interaction with the DNA. Through site-directed mutagenesis studies, several amino acids have been identified, which are crucial for DNA-binding and protein folding of AmtR in *C. glutamicum* [127]. TetR-type transcriptional regulators usually bind small effector molecules in order to fulfil their function. In *C. glutamicum*, AmtR interacts only with the adenylylated P_{II} protein GlnK, that occurs under nitrogen limitation, resulting in the release of AmtR from its target DNA so that expression of the AmtR regulated genes is elevated [27,39,128].

AmtR-dependent transcriptional regulation was most intensively studied in *C. glutamicum*, where transcriptome and proteome analysis revealed that at least 35 genes are directly controlled by AmtR (Figure 1.2D [purple]) [11,27,39,40]. These include genes involved in ammonium uptake (*amtA*, *amtB*), ammonium assimilation (*glnA*, *gltBD*), urea metabolism (*ureABCEFGD*) and regulatory proteins (*glnK*, *glnD*) that show elevated gene expression under nitrogen starvation (Figure 1.2D [purple]) [39]. Further transcriptome and metabolome studies also revealed that AmtR regulates *dapD*, a succinylase involved in diaminopimelate biosynthesis, establishing a link between nitrogen control and L-lysine biosynthesis [40].

During nitrogen surplus, AmtR of *C. glutamicum* acts as transcriptional repressor blocking transcription of the aforementioned genes through binding at an AmtR-binding site with the palindromic sequence tttCTAT-N_6-AtAGat (capitalization = high conservation) [27,39,127]. While all AmtR-regulated genes are upregulated in response to nitrogen limitation, their basal expression level under nitrogen excess conditions depends on the numbers and alterations

of the AmtR-binding site [39,129]. A recent study also highlighted the importance of the number of spacing nucleotides between the binding motif half sites, which is crucial for binding of AmtR and consequent transcriptional repression [127]. The AmtR protein is lacking an autoregulatory mechanism due to a shortened spacing region, although a putative AmtR-binding site was identified upstream of the *amtR* gene itself [27,127].

AmtR in Mycobacteria

The presence of the AmtR protein is not restricted to corynebacteria, but it is also found in genomes of closely related species such as *M. smegmatis*, *N. farcinica* and *S. avermitilis*, all of which also encode for homologs of GlnR (Figure 1.3G [orange]) [11,71]. Interestingly, *M. smegmatis* is one of a few environmental mycobacteria that encodes for a homolog of AmtR, while AmtR is absent in the genomes of mycobacteria that belong to the *M. tuberculosis* complex (Figure 1.3G [orange]) [11,23]. No AmtR-binding sites could be identified in any of the mycobacterial genomes using the AmtR consensus sequence of *C. glutamicum* [11]. Jessberger *et al.* used a bioinformatic approach and looked at the genomes of different actinomycetes, where they identified five genes (*msmeg_2184*, *msmeg_2185*, *msmeg_2186*, *msmeg_2187* and *msmeg_2189*) that showed a high co-occurence and co-localisation with the *amtR* gene (e.g. in *S. avermitilis* and *Kineococcus radiotolerans*) and also verified transcriptional regulation of these genes by AmtR in *M. smegmatis* (Figure 1.3G [orange]) [38]. Recent work in *S. avermitilis* has further shown that *amtRsav* (*amtR* gene in *S. avermitilis*) regulates the same set of genes as reported for *M. smegmatis* [130]. However, in *M. smegmatis* at least four of these genes are also under the control of GlnR (Figure 1.3D [purple] and Figure 1.3G [orange]), implying a fine-tuning regulatory mechanism for AmtR in *M. smegmatis*, rather than a function as a global nitrogen regulator like in *C. glutamicum* [38]. It is still unclear, whether mycobacterial AmtR binds a small effector molecule, inducing the release of the repressor protein from its binding site or if it is interacting with another protein like in *C. glutamicum*. A number of putative AmtR ligands have been tested so far (e.g. urea, various amino acids), but all metabolites had no effect

on AmtR:DNA binding in *M. smegmatis* [38]. More research is now required to determine the genome-wide effect of AmtR in *M. smegmatis*. Further studies on the characterisation of the structure, biochemistry and biophysical properties of AmtR are required to understand the reason for the co-occurence of homologs of the two global nitrogen regulatory proteins, GlnR and AmtR, in some actinomycetes like *M. smegmatis*.

Links to other regulators of cell metabolism

Nitrogen metabolism is tightly coupled to carbon metabolism through the rapid interconversion of several metabolites in the TCA cycle, especially 2-oxoglutarate into glutamate and vice versa (Figure 1.5 and Figure 1.6). Key enzymes of nitrogen metabolism that catalyse these reactions are usually under control of the global nitrogen regulatory proteins, however, a regulatory linkage between nitrogen assimilation and carbon metabolism has been reported for a number of organisms [25,94,103,131-134].

Links between nitrogen and carbon metabolism

In *E. coli*, transcription of the genes *glnA*, encoding the glutamine synthetase and *glnH*, encoding a high-affinity glutamine uptake system, can occur from two different promoters *glnAp1* and *glnAp2* [94]. Transcription of these genes is usually regulated by NtrBC, however, a regulatory mechanism through the cAMP receptor protein (CRP) was also reported, providing a novel regulatory linkage between carbon metabolism and nitrogen assimilation in *E. coli* [94,135,136]. Similar observations have been made in *C. glutamicum*, where both transcription and activity of the glutamine synthetase and glutamate synthase are affected by the intracellular carbon availability [103]. Furthermore, the amino acid L-leucine has several effects on gene expression in *E. coli* that are summarised as the Lrp response. This Lrp response is controlled by the leucine-responsive regulatory protein Lrp that governs expression of genes involved in amino acid metabolism, carbon metabolism and pili synthesis (for review see [137,138]). In total, 138 LRP-binding regions were identified in the genome of *E. coli*, however, 17 of them are located in intragenic regions, suggesting putative regulatory functions for these genomic regions. LRP further

controls expression of 11 other transcriptional regulatory proteins that altogether regulate another 34 genes. Interestingly, different patterns of regulation for these regions have been identified, where L-leucine can have a potentiating, antagonising, or no effect on LRP [133]. Recent work from Cho *et al.* further showed that the transcription factors ArgR (arginine repressor) and TrpR (tryptophan repressor) regulate transcription in response to intracellular arginine and tryptophan levels in *E. coli* [134]. Functional classification of genes regulated by Lrp, ArgR and TrpR show a distinct regulation of genes involved in amino acid metabolism, but more than half of the regulated genes were categorised into another functional class (e.g. carbohydrate metabolism, energy metabolism, lipid metabolism, cofactor metabolism and vitamin metabolism) [134]. Interestingly, both Lrp and ArgR were also shown to bind to the promoter region of *gltBD*, the genes encoding glutamate synthase [133,134].

The P_{II} protein in *Synechocystis* mediates the uptake of inorganic carbon and nitrogen sources and the CbrA-CbrB two-component system of *Pseudomonas aeruginosa* controls the utilisation of multiple organic carbon (e.g. glucose, pyruvate and mannitol) and nitrogen sources (e.g. L-arginine, L-histidine and L-proline) [25,131]. A further link between carbon and nitrogen metabolism was reported, when Buchinger *et al.* identified AmtR as transcriptional regulator of *mez*, encoding for malic enzyme, connecting AmtR to carbon metabolism in *C. glutamicum* [40].

Links between nitrogen and phosphate metabolism

A genome-wide transcriptomic analysis in *S. coelicolor* suggested a link between nitrogen and phosphate metabolism. These are linked through complex regulatory networks that are regulated by GlnR and the PhoRP two-component system that coordinate nitrogen and phosphate metabolism [37,126,139]. Direct interaction of PhoP with the promoter regions of *glnA*, *glnII* and the *amtB-glnK-glnD* operon was shown, indicating a competitive binding mechanism of PhoP and GlnR (Figure 1.4D [purple] and Figure 1.4E [brown]) [125]. Interestingly, PhoP also binds to the promoter region of *glnR*, thus further strengthening the link between nitrogen and phosphate metabolism [140].

1.3.3. Response to nitrogen limitation in Actinobacteria

Nitrogen metabolism and its regulation have been extensively studied in different bacterial species and canonical mechanisms such as ammonium uptake and its assimilation via GDH or GS/GOGAT have been described. All bacteria harbor genes that encode for these mechanisms to ensure a sufficient nitrogen supply for various cellular processes (e.g. protein, nucleotide and peptidoglycan biosynthesis). However, due to fluctuating nitrogen availability in the environment, bacteria can face conditions where available nitrogen sources are limited. In order to ensure their survival, bacteria must have evolved mechanisms to cope with changes in nitrogen availability and to overcome these periods of nitrogen depletion. This response to nitrogen limitation is mostly subject to complex regulatory systems that differ between various species (e.g. NtrBC, GlnR and AmtR). While the knowledge about the response of enteric bacteria is profound, understanding the intracellular response of actinobacteria to nitrogen depletion has turned into focus just recently, shedding light into regulatory mechanisms and metabolic pathways that are important under these conditions [46,47,141,142].

Nitrogen limitation in C. glutamicum

The response to nitrogen depletion in *C. glutamicum* has been the subject of many studies and was analysed in both batch culture and continuous culture [141,143,144]. When grown in batch culture, differential gene expression of 248 genes (102 upregulated, 146 downregulated) was observed, including the AmtR regulon. Categorisation of these genes into functional groups revealed changes of various transport mechanisms and metabolic pathways involved in nitrogen metabolism and amino acid biosynthesis. Furthermore, a massive decrease in protein biosynthesis (34 ribosomal proteins downregulated) was reported, most likely the result of growth rate changes [141]. *C. glutamicum* induces the expression of genes that encode transport and assimilation systems for ammonium, urea and creatinine, while transcription of a glutamate uptake system is decreased under nitrogen limitation (Figure 1.2F [cyan]) [141]. Furthermore, transcription of the

compatible solute uptake systems *betP* and *putP* are downregulated with a concomitant decrease in proline uptake. Genes encoding for proteins that are involved in amino acid biosynthesis were also downregulated [141].

Analysis of the transcriptional response to nitrogen limitation in a continuous culture showed differential expression of 285 genes. Again, the AmtR regulon was fully represented, but major changes in gene expression compared to batch culture grown cells were observed (e.g. glutamate uptake system, ribosomal proteins, cell division and cell wall synthesis) [145]. Transcription of genes encoding proteins of the L-aspartate, L-arginine, L-leucine, L-threonine and aromatic amino acid biosynthesis pathways were decreased [144]. Furthermore, the gene *lysA*, encoding the diaminopimelate decarboxylase, is repressed under nitrogen limitation. It has been suggested that this is to avoid a drain of ammonium into L-lysine biosynthesis [144]. Comparison of the transcriptional response of cells grown in continuous culture to cells grown in batch culture revealed major advantages of a continuous culture that allows establishment of well-defined conditions and enables discrimination between specific and non-specific cellular responses to nutrient limitation [145].

Nitrogen limitation in S. coelicolor

The GlnR regulon in *S. coelicolor* is comparatively small and comprises only 15 genes. Eight of these genes are directly involved in nitrogen metabolism and their function in ammonium transport and assimilation was previously described, whereas the remaining genes are assigned to other functional categories (Figure 1.4D [purple] [41]. However, a proteomic analysis revealed indirect modulation of 54 proteins by GlnR under nitrogen limitation. These proteins were assigned to amino acid metabolism, carbon metabolism, antibiotic biosynthesis and the general stress response, implying a much more global function of GlnR in *S. coelicolor* and a close interaction of nitrogen metabolism with other intracellular functions [146]. A genome-wide analysis of the transcriptomic response to nitrogen limitation indicated overlaps to the GlnR regulon and revealed further genes with elevated expression upon nitrogen limitation [147]. However, almost no overlap to the genes identified by Tiffert *et*

al. was detectable [146,147]. Both analyses showed the upregulation of merely one amino acid transporter, but besides that, no further nitrogen compound uptake systems and only a few genes involved in amino acid and nucleotide metabolism that showed differential gene expression in response to nitrogen austerity were reported [146,147].

Nitrogen limitation in M. smegmatis

M. smegmatis is one of the few organisms that encode for both global nitrogen regulatory protein, GlnR and AmtR, but to date no role has been reported for AmtR in mycobacteria (Figure 1.3D [purple] and Figure 1.3G [orange]) [11,71]. GlnR has been the subject of several studies aiming to unravel its crucial function in regulation of the transcriptomic response to nitrogen limitation in mycobacteria (Figure 1.3D [purple]) [38,43,123,124]. In a genome-wide analysis, Jenkins *et al.* identified GlnR as transcriptional regulator that mediates the expression of more than one hundred genes once external nitrogen sources are exhausted with the emphasis of these in regard to their functional characterisation on nitrogen scavenging [43]. Amongst the differentially expressed genes are several ammonium transport systems, uptake systems for nucleotides, amino acids, peptides and inorganic nitrogen sources and primary nitrogen metabolism pathways. Additionally, genes that encode for pathways important for the utilisation of nitrate, urea and cellular components as alternative nitrogen sources, which was consistent with observations reported by Jessberger *et al.* (Figure 1.3D [purple] [38,43]. The vast majority of GlnR-regulated genes were also identified in a microarray study using batch culture nitrogen run out experiments. In this study, Williams *et al.* reported more than 1000 genes that are differentially expressed in *M. smegmatis* in response to nitrogen limitation [142]. Surprisingly, only a remarkably low proportion of these genes is assigned to nitrogen metabolism, whereas central carbon metabolism and cellular processes, such as DNA and RNA synthesis, exhibit major differences in their transcription level in response to nitrogen limitation [142]. Amino acid metabolism is only slightly affected, which is in contrast to a recent study published by Elharar *et al.*, which provided evidence that nitrogen limitation in mycobacteria induces the Pup proteasome,

which ultimately leads to protein degradation into amino acids [148]. Amino acids are further suggested to play a key role in *M. tuberculosis* pathogenesis [18,19]. Nitrogen metabolism in *M. tuberculosis* was recently reviewed by Gouzy *et al*., discussing the central nitrogen metabolism and its relevance for the physiology and virulence in this opportunistic pathogen [16].

Metabolomic studies on nitrogen-limited cells of *M. smegmatis* also showed an accumulation of the sugar-derived compound, glucosylglycerate, suggesting it as novel marker of nitrogen stress [149]. Glucosylglycerate accumulation in *M. smegmatis* was only observed in response to nitrogen limitation, but not to oxidative or osmotic stress. Glucosylglycerate is suggested to act as nitrogen capacitor under nitrogen limitation, however, glucosylglycerate biosynthesis is not essential in *M. smegmatis* [149]. This metabolite was initially identified in *Mycobacterium phlei* as part of the methylglucose lipopolysacharides and since then has also been identified in other mycobacterial species [150-152]. However, in actinobacteria, accumulation of glucosylglycerate is not only linked to nitrogen depletion, but also to salt stress, where it acts as compatible solute [149,153].

In summary, nitrogen is an essential component of all living cells and prokaryotes have developed elaborate mechanisms for the uptake, assimilation and metabolism of nitrogen. However, despite the importance of nitrogen for cell growth and maintenance of intracellular metabolism, there have been only a few studies on nitrogen metabolism of free-living saprophytes and pathogenic species of mycobacteria. We lack general understanding of how mycobacteria adapt to a nitrogen-limited environment as well as the regulatory mechanisms that mediate this adaptation. The work presented in this thesis will contribute new knowledge and insight into addressing these questions in the fast-growing saprophytic species *M. smegmatis* and will provide valuable tools to study nitrogen metabolism in pathogenic mycobacteria such as *M. tuberculosis*.

1.4. Aims of This Study

The overall aim of this study was to elucidate the molecular mechanisms used by *M. smegmatis* to sense and respond to nitrogen limitation. To address this aim I have performed the following objectives:

1. **Development of a nitrogen-limited continuous culture of *M. smegmatis*.** This objective will set the platform for this study and be used to define changes in the transcriptome of *M. smegmatis* in response to nitrogen limitation.

2. **Identification of the nitrogen-regulated transcriptome of *M. smegmatis*.** This objective will determine the genome-wide expression profile of *M. smegmatis* in response to nitrogen limitation using RNA-sequencing analysis.

3. **Characterisation of the physiological role of the global nitrogen regulator AmtR in *M. smegmatis.*** AmtR is a TetR-like transcriptional regulator with unknown function in environmental mycobacteria. A combination of physiological, molecular and biochemical analyses will be performed to identify the molecular function of AmtR in *M. smegmatis*.

4. **Characterisation of adaptation mechanisms to nitrogen limitation in *M. smegmatis*.** Expression of many transcriptional regulatory proteins is expected to be affected in response to nitrogen limitation. After these transcriptional regulators are identified using RNA-sequencing, I will determine the role of a selected transcriptional regulator on the adaptation of *M. smegmatis* to nitrogen limitation.

Chapter 2

Defining the nitrogen regulated transcriptome of *Mycobacterium smegmatis* using continuous culture

Michael Petridis, Andrej Benjak and Gregory M. Cook †

Article Submitted: BMC Genomics

Author contributions: Authors are listed in order of magnitude of their contribution in each role. Corresponding author is indicated by a cross (†). MP and GMC planned research. MP performed the research with assistance from AB for the bioinformatics analysis of the RNA-sequencing data. MP and GMC interpreted data and wrote the manuscript. MP produced all Figures and Tables for the manuscript.

2.1. Abstract

Nitrogen is essential for microbial growth and its importance is demonstrated by the complex regulatory systems used to control the transport, assimilation and metabolism of nitrogen. Recent studies are beginning to shed light on how mycobacteria cope with nitrogen limitation and several regulators (e.g. GlnR, P_{II}) have been characterised at a molecular level. To gain further insight into the response of mycobacteria to nitrogen limitation, we developed a nitrogen-limited chemostat. We compared the transcriptional response of nitrogen-limited cells to carbon-limited cells using RNA-Seq analysis in a continuous culture model at a constant growth rate. Our findings revealed significant changes in the expression of 357 genes (208 upregulated, 149 downregulated; >2-fold change, FDR<5%) in response to nitrogen limitation in continuous culture. The vast majority (68%) of the GlnR regulon was differentially expressed under nitrogen limitation in continuous culture and ~52% of the 357 genes overlapped with a previously published study investigating the response of *M. smegmatis* to nitrogen limitation in batch culture, while only 17% of the genes identified in the batch culture were differentially expressed in our chemostat model. Moreover, we identified a unique set of 45 genes involved in the uptake and metabolism of nitrogen that were exclusive to our chemostat model. We observed strong downregulation of pathways for amino acid catabolism (i.e. alanine, aspartate, valine, proline and lysine), suggesting preservation of these amino acids for critical cellular function. We found 16 novel transcriptional regulators that were directly or indirectly involved in the global transcriptomic response of *M. smegmatis* to nitrogen limitation and identified several non-coding RNAs that might be involved in the transcriptional or post-transcriptional regulation of nitrogen-regulated gene expression. Our data highlight the versatile metabolic capability of *M. smegmatis* and provide a molecular framework for understanding how environmental mycobacteria respond to nitrogen-depleted environments.

2.2. Introduction

Carbon and nitrogen are key components of organic material and their availability in the environment is necessary for growth and survival of microorganisms. The major mechanisms of carbon source utilisation and its regulation are well defined for many groups of microorganisms, but our knowledge of nitrogen metabolism and its regulation is largely confined to enteric bacteria (for review see [22,46,47]). The importance of nitrogen for microbial growth is demonstrated by the complex regulatory systems used to control nitrogen assimilation in response to internal and external nitrogen levels. In some bacteria, nitrogen availability is sensed by the intracellular ratio of 2-oxoglutarate:glutamine [109]. This signal ensures that uptake of nitrogen sources commensurate with the metabolic requirements of the organism i.e. carbon or nitrogen. High 2-oxoglutarate levels signal nitrogen limitation, while high glutamine levels signal nitrogen excess in bacteria [46,109,154].

In enteric bacteria, nitrogen metabolism is regulated by the two-component regulatory system NtrBC and the signal transduction protein P_{II} [104]. These regulatory systems coordinate nitrogen metabolism by regulating genes involved in ammonium assimilation, amino acid transport and nitrate metabolism through protein-protein interactions [32,46,104,109]. An integral component of this regulation is GlnD, an uridylyl transferase/uridylyl-removing enzyme that senses and transmits the nitrogen status of the cell to the P_{II} protein [155]. The P_{II} protein in its uridylylated state stimulates phosphorylation of the response regulator NtrC by the NtrB kinase, which results in expression of the NtrBC-regulated genes under nitrogen limitation [156]. Homologues of the P_{II} protein are found in many bacterial genomes, where they are mainly involved in the regulation of nitrogen metabolism. However, the P_{II} protein may have additional roles, as recently described in cyanobacteria, where it regulates the metabolism of inorganic carbon [131].

Actinomycetes are a group of Gram-positive bacteria that inhabit a wide range of aquatic, terrestrial and human habitats. In actinobacteria, two different regulatory systems have been identified that regulate the

transcriptional response to nitrogen limitation, AmtR, a TetR-type transcriptional regulator and GlnR, an OmpR-type transcriptional regulator [39,41,157]. AmtR was identified in *C. glutamicum* to regulate more than 30 genes involved in nitrogen metabolism, including the P_{II} protein, glutamine synthetase, glutamate synthase and urease [39]. In *S. coelicolor*, GlnR mediates the transcriptional response to nitrogen limitation by elevating expression of the *amtB-glnK-glnD* operon and the glutamine synthetase *glnA*, while repressing transcription of the glutamate dehydrogenase *gdhA* [37,41,158].

Slow-growing mycobacteria such as *M. tuberculosis* appear to harbor only GlnR. A conserved GlnR-binding site has been identified upstream of genes involved in nitrogen metabolism in *M. tuberculosis* [41], however, there is no direct evidence for involvement of GlnR in the regulation of nitrogen metabolism and its role as global nitrogen regulator in *M. tuberculosis* has not been defined. *M. tuberculosis* can exploit different nitrogen sources, however, organic nitrogen sources (e.g. asparagine and aspartate) are preferred for extracellular and intracellular growth [18,19]. Amino acids can function as both carbon and nitrogen sources, but although *M. tuberculosis* can metabolise a variety of amino acids *in vitro*, only a small number (aspartate, glutamate, asparagine and glutamine) can support growth of *M. tuberculosis* at the acidic pH that prevails in macrophages and their availability is essential for virulence *in vivo* [16]. Inorganic nitrogen sources such as urea, ammonium and nitrate are less efficient for growth than amino acids [16], but the reasons for this remain unknown.

The genome of the fast growing saprophytic actinobacterium *M. smegmatis* contains copies of both global nitrogen regulators, GlnR and AmtR [11]. It has recently been demonstrated that GlnR regulates the expression of more than one hundred genes in response to nitrogen limitation in *M. smegmatis* [43], while the AmtR regulon remains unknown. Previous microarray studies using batch culture nitrogen run out experiments have identified 1090 genes in *M. smegmatis* that are differentially expressed in response to nitrogen limitation. However, only a small subset of these genes is under control of GlnR, while the regulatory mechanisms for the majority have yet to be identified [43,142]. During nitrogen run out studies, the growth rate

changes significantly as nitrogen becomes depleted and therefore these types of experiments may not unequivocally uncover genes that respond solely to nitrogen limitation.

The aim of the current study was to deliver a molecular framework for how a mycobacterial cell responds to nitrogen-depleted and -replete conditions at a defined growth rate. To address this aim, we developed a nitrogen-limited chemostat for *M. smegmatis* at a constant growth rate of 0.12 h^{-1} (t_d = 5.7 h) followed by RNA-sequencing analysis to identify genes responding to nitrogen limitation.

2.3. Results and Discussion

2.3.1. Development and validation of a nitrogen-limited continuous culture to understand the global transcriptomic response to nitrogen limitation

The first step of this study was to establish a defined minimal medium limited for either nitrogen or carbon to identify the molecular response of *M. smegmatis* to nitrogen limitation. We decided to use HdB medium, because it was previously used in our lab for carbon limitation studies in continuous culture. We modified HdB medium [1] by replacing 0.05% (w/v) Tween-80 with 0.05% (w/v) Tyloxapol, which cannot be metabolised as a carbon source, and $(NH_4)_2SO_4$ with NH_4Cl and K_2SO_4 (11.4 mM) to avoid simultaneous limitation for sulphur and nitrogen. Inorganic nitrogen sources (e.g. ammonium) are less efficient for growth than amino acids, however, we chose an inorganic nitrogen source to allow tighter control of the nitrogen to carbon ratio in the medium. Using batch culture, we identified glycerol as a suitable carbon source, with a high cell yield and no change in external pH during the course of the experiment compared to other carbon sources such as glucose, acetate or succinate (Figure 2.1A). When cells were grown in batch culture on glycerol (25 mM) under nitrogen-replete conditions, growth ceased at an optical density (OD_{600}) of 3.5 (Figure 2.1B). When the culture was limited for nitrogen (carbon excess), the final OD_{600} was 2.1 (Figure 2.1B). Based on these experiments, we used 30 mM glycerol and 10 mM NH_4Cl (nitrogen-replete) / 1.25 mM NH_4Cl (nitrogen-depleted) for continuous culture studies.

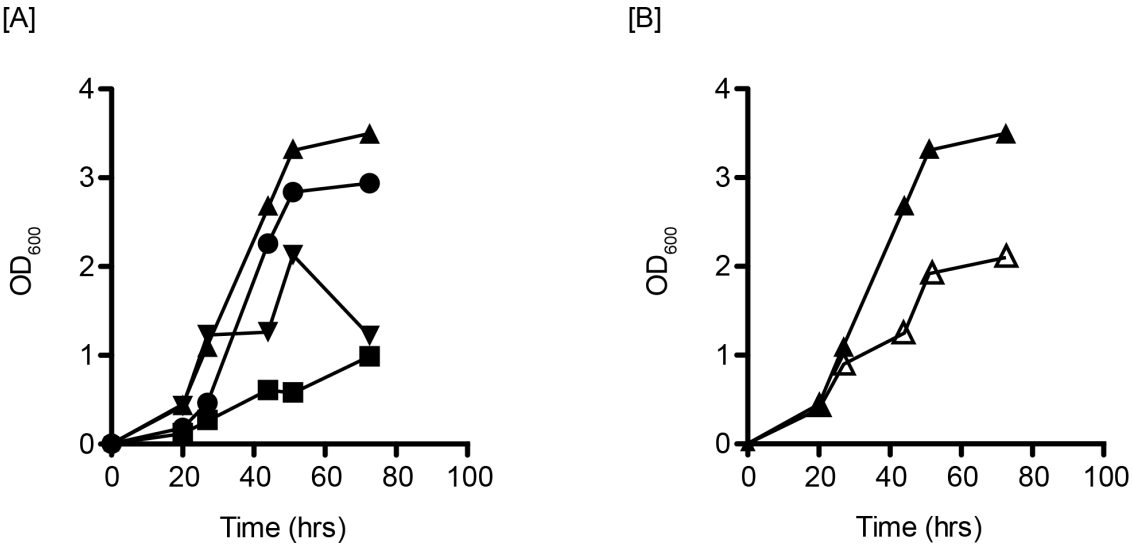


Figure 2.1. Growth of *M. smegmatis* in batch culture with different carbon sources under nitrogen-depleted and nitrogen-replete conditions. **[A]** Growth curves of *M. smegmatis* in nitrogen-replete (10 mM NH_4Cl) HdB minimal medium supplemented with different carbon sources. 12.5 mM glucose (closed circles), 18.75 mM succinate (closed triangles down), 25 mM glycerol (closed triangles) and 37.5 mM acetate (closed squares) were added as carbon source and growth was monitored by optical density measurements at 600 nm. **[B]** Comparison of growth in nitrogen-depleted (1 mM NH_4Cl) (open triangles) and nitrogen-replete (10 mM NH_4Cl) (closed triangles) HdB minimal medium with 25 mM glycerol as sole carbon source. These experiments were performed in single biological replicate.

In continuous culture, *M. smegmatis* was grown at 50% O_2 saturation and at a dilution rate of 0.12 h^{-1} until steady-state conditions were achieved (typically after 3-4 resident times; the resident time is defined as the time to entirely exchange the volume of the bioreactor and is dependent on the reactor size, reactor volume and the flow rate of fresh medium). During steady-state growth all intrinsic parameters of the cells remain constant (e.g. growth rate) and the dilution rate, pH and temperature of the medium remain constant. Nitrogen-limited cultures (1.25 mM NH_4Cl / 30 mM glycerol) reached steady-state at an OD_{600} of 0.8 and no residual ammonium could be detected (Figure 2.2A). The residual concentration of glycerol in the medium was approximately 18 mM indicating that the cells were not carbon-limited (Figure 2.2B). At this stage, addition of 25 mM NH_4Cl into the culture vessel led to an immediate increase in optical density, demonstrating that the cultures were indeed limited for nitrogen (Figure 2.2A). Addition of 10 mM glycerol to a nitrogen-limited culture had no effect (Figure 2.2C). In the carbon-limited culture (10 mM NH_4Cl / 30 mM glycerol), the steady-state-phase was reached after 56 h at an OD_{600} of 4.3 with the residual ammonium concentration in the medium remaining constant at 0.5-1 mM (Figure 2.2E). A pulse of 1 mM NH_4Cl into the culture vessel had no effect on growth (Figure 2.2F), while the addition of glycerol led to a significant increase in OD_{600}, demonstrating nitrogen excess (replete) under these conditions (Figure 2.2D).

For each condition, three independent biological replicates were grown until steady-state-phase and cells were harvested after at least 3 volume changes in the chemostat vessel. The analysis of the transcriptome using RNA-sequencing revealed a total of 357 differentially expressed genes under nitrogen limitation (>2-fold change, FDR<5%) i.e. approximately 5% of the *M. smegmatis* genome (Table 2.5, supplementary data). In total, 208 genes were upregulated and 149 were downregulated. Classification of these genes into functional categories showed major changes in transport proteins, genes that are associated with nitrogen and amino acid metabolism and genes that are functionally assigned as regulatory proteins. A large number of genes encoding for hypothetical proteins were differentially expressed (36 upregulated; 38 downregulated), however, their function in response to nitrogen limitation needs to be investigated (Figure 2.3).

Previous work published by Berney *et al.* characterised the response of *M. smegmatis* to carbon limitation in continuous culture and a comparison with this dataset allowed us to identify genes in our analysis of the nitrogen-regulated transcriptome that rather show differential expression in response to carbon limitation than to nitrogen limitation [15]. In total, 58 genes were identified that were mostly associated with central carbon metabolism (e.g. succinate dehydrogenase) and fatty acid metabolism (e.g. polyketide synthase, fatty acid synthase, acyl-CoA dehydrogenase, long chain fatty acid-ACP ligase) and these genes were excluded from further analysis.

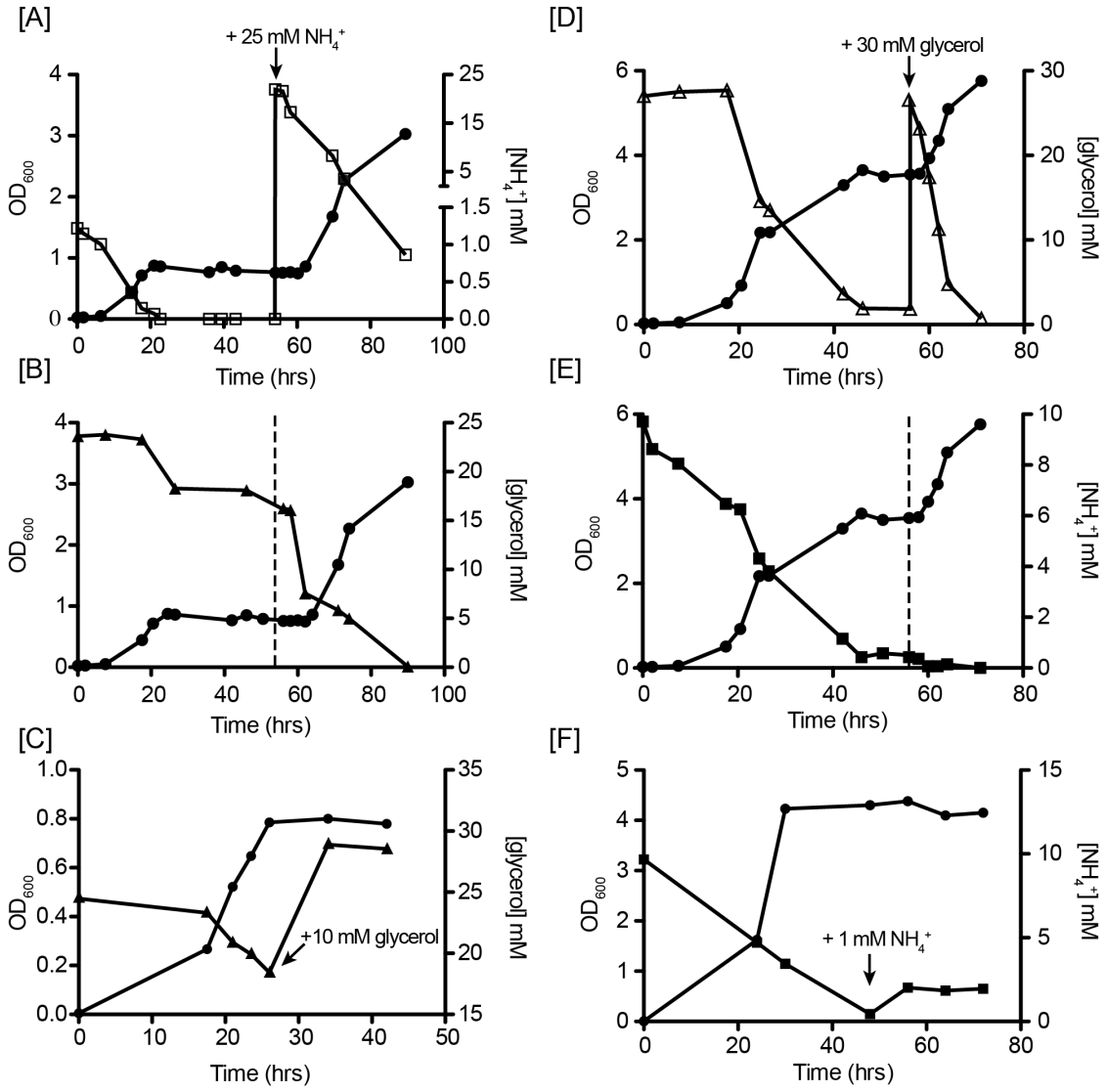


Figure 2.2. Growth of *M. smegmatis* in continuous culture under nitrogen-depleted and nitrogen-replete conditions. Optical density (OD_{600}), residual glycerol and ammonium concentrations in the medium were monitored during the experiment. The effect of a pulse with 25 mM NH_4Cl on residual ammonium (open squares) **[A]** and glycerol (closed triangles) **[B]** concentrations in the medium and OD_{600} (closed circles) in a nitrogen-depleted culture. **[C]** A pulse with 10 mM glycerol had no effect on a nitrogen-depleted culture. The effect of a pulse with 30 mM glycerol on residual glycerol (open triangles) **[D]** and ammonium (closed squares) **[E]** concentrations in the medium and OD_{600} (closed circles) in a nitrogen-replete culture. **[F]** A pulse with 1 mM NH_4Cl had no effect on the nitrogen-replete culture, respectively.

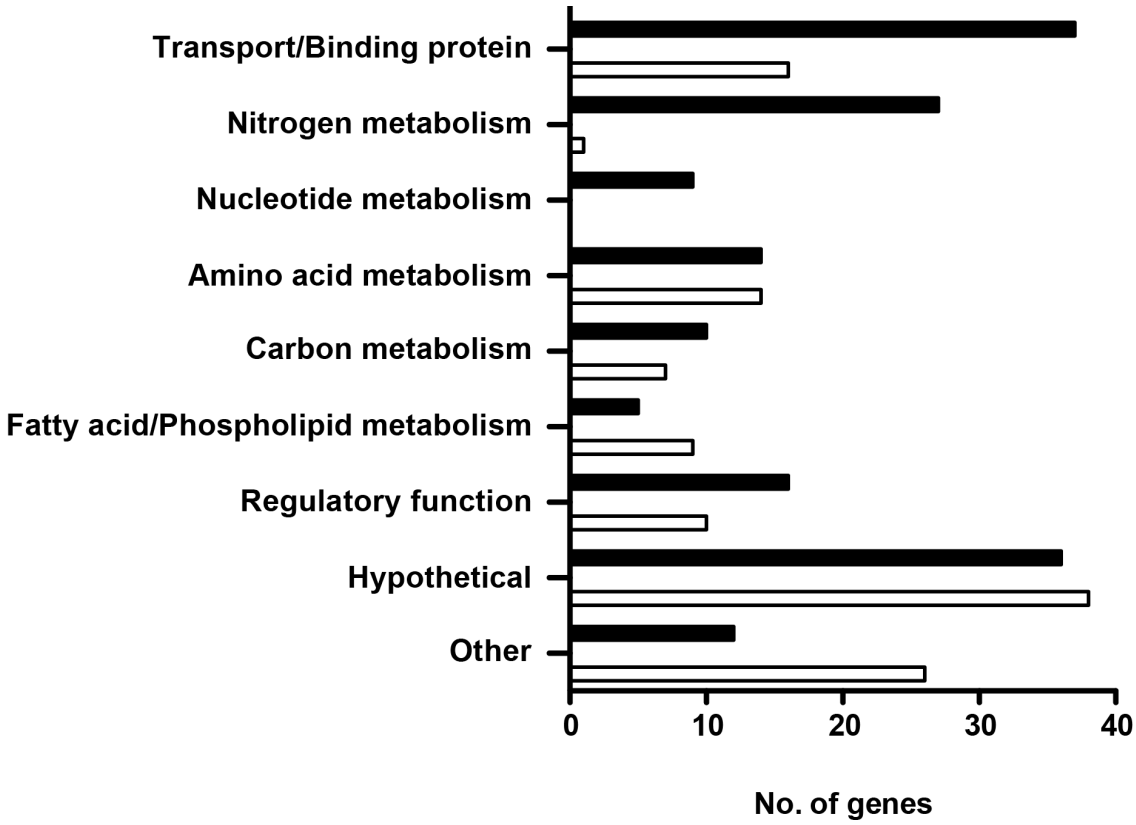


Figure 2.3. Functional categorisation of differentially expressed genes. Genes were assigned to functional categories using the Biocyc Database Collection. Bars indicate number of genes that were upregulated (black) and downregulated (white) under nitrogen limitation.

2.3.2. Nitrogen limitation studies in continuous culture versus batch culture

Previous work published by Williams *et al.* focused on the response of *M. smegmatis* to nitrogen stress in batch culture nitrogen run out experiments and showed differential expression of 1090 genes (574 upregulated, 516 downregulated) (Figure 2.4A) [142]. Surprisingly, 903 of these genes were not differentially expressed in our continuous culture. In fact, only 17% of the genes reported by Williams *et al.* responded to nitrogen limitation in continuous culture, including 70 genes that were predicted to be under control of GlnR (Figure 2.4A) [43,142]. Expression of a similar set of ammonium and nucleotide uptake systems and metabolic pathways were elevated in both nitrogen limitation studies (Table 2.1 and Table 2.2), however, we identified a unique set of genes involved in the metabolism of amino acids, nucleotides and urea that were nitrogen-responsive in continuous culture (Table 2.1 and Table 2.2). We further identified 36 genes, including regulatory enzymes, which showed an inversed expression profile in continuous culture compared to batch culture. In batch culture it is often difficult to assign transcriptional changes to a single stimulus due to changes in growth rate, nutrient depletion and end product buildup. It is striking that there was very little overlap in the downregulated genes in response to nitrogen limitation between the batch culture and continuous culture (Figure 2.4C). A major portion of the 497 downregulated genes in batch culture is associated with the general reduction in cellular metabolism due to a reduced growth rate. Development of a nitrogen-limited continuous culture enabled us to define transcriptional changes purely in response to nitrogen limitation, excluding growth rate effects. We showed that using continuous culture we can identify the set of genes involved in nitrogen uptake and metabolism that were reported by Williams *et al*., however, this system allowed us to further extend our knowledge towards the transcriptional response to nitrogen limitation and reduce the number of genes responding to other environmental factors [142].

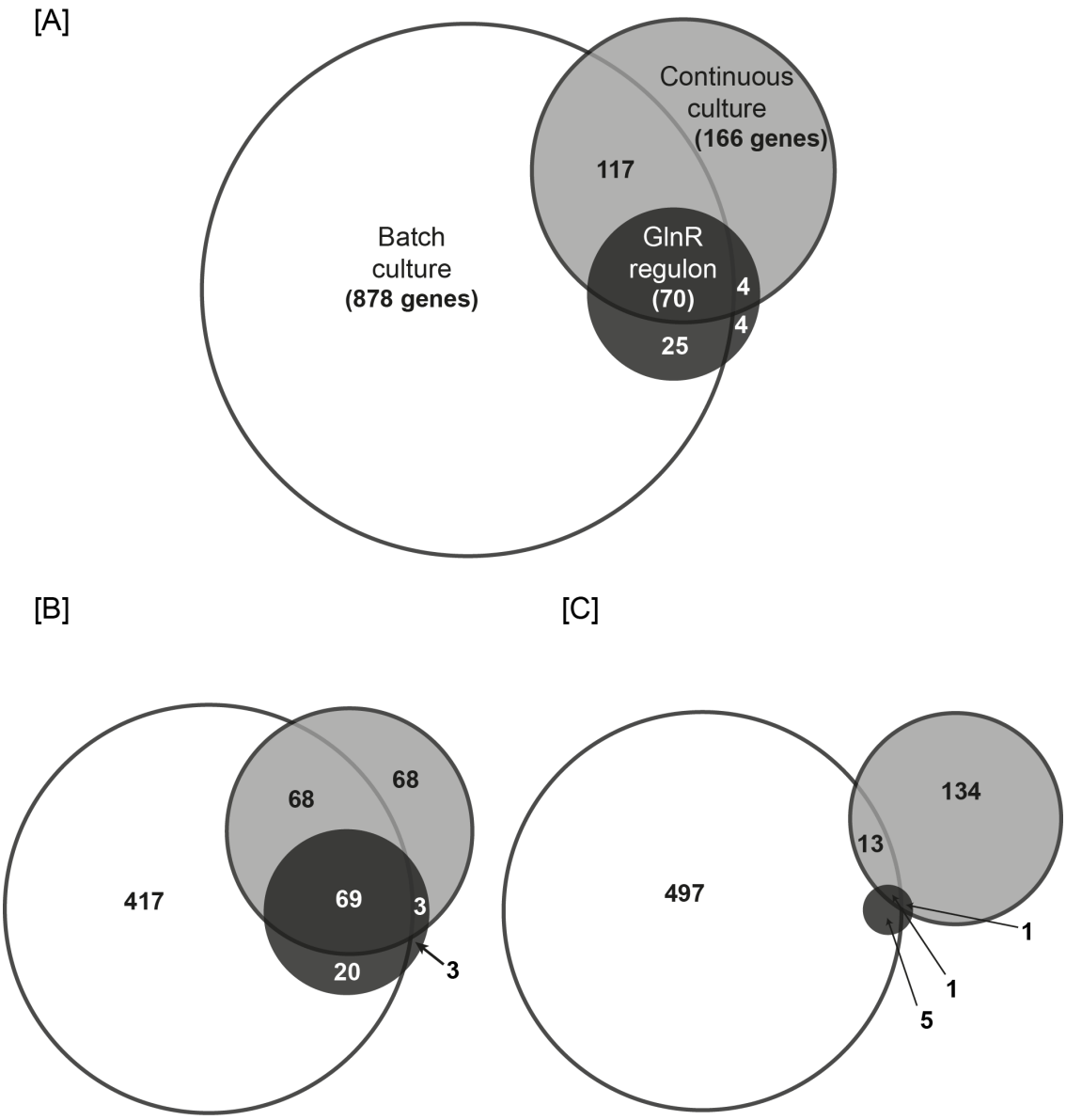


Figure 2.4. Schematic distribution of overlapping genes that were differentially expressed comparing nitrogen-depleted continuous culture versus nitrogen-depleted batch culture [142]. In this comparison, we included genes that were reported to be under control of GlnR [43]. **[A]** Included are all genes that were reported to be differentially expressed in continuous culture and batch culture. **[B]** Upregulated genes that are upregulated in continuous culture and batch culture and activated by GlnR. **[C]** Downregulated genes that are downregulated in continuous culture and batch culture and are repressed by GlnR. Numbers indicate the total number of genes that fall into the respective category.

2.3.3. Nitrogen limitation activates the expression of genes involved in scavenging nitrogen sources in the environment

In this study, a total of 53 genes encoding for transporters or corresponding binding proteins were differentially expressed in response to nitrogen limitation (Table 2.5, supplementary data). Of the upregulated genes, 18 encoded for transporters that are generally involved in uptake of both organic and inorganic nitrogen containing compounds like amino acids (e.g. *msmeg_2525*), nucleotides (e.g. *msmeg_5730*) and ammonium (e.g. *msmeg_4635*) (Table 2.1). Previously published data by Berney and Cook (2010) showed downregulation of nine of these transporters in a carbon-limited continuous culture comparing slow growth (t_d 69 h) versus fast growth (t_d 4.6 h) [15]. All three ammonium transporters and several amino acid (3 of 8) and nucleotide transporters (3 of 7) showed a significant decrease in transcript level, however, transcription of all of these genes was enhanced upon nitrogen limitation. The GlnR regulon comprises 15 of these nitrogen compound transporters including the three ammonium transporters and the majority of nucleotide and amino acid transporters. The most prominent amino acid permease (*msmeg_2525*) was not under direct control of GlnR. In continuous culture, we identified 4.8-fold upregulation of a polar amino acid ABC transport system that has not been reported previously to be affected by nitrogen limitation. Further work is required to identify the regulatory mechanisms that mediate the uptake of amino acids under nitrogen depletion. We observed a decrease in gene expression for the catabolism of five amino acids suggesting *M. smegmatis* attempts to balance the availability of ammonium and particular amino acids under nitrogen depletion. This finding is supported by a recent report of a proteasome-mediated amino acid recycling mechanism in *M. smegmatis* under nutrient limitation [146]. The transcriptomic response of *M. smegmatis* to increase its ability for the uptake (scavenging) of nitrogen compounds is remarkably different from other actinobacteria like *C. glutamicum*, where no amino acid transport systems were upregulated under nitrogen limitation [141].

Peptides appear to play an important role in replenishing the intracellular ammonium pool necessary for the different anabolic pathways in *M. smegmatis* (Table 2.2 and Figure 2.5). Amongst the genes induced in response to nitrogen limitation we identified a peptidase, an aminopeptidase and an aliphatic amidase catalysing the successive degradation of peptides to amino acids and the subsequent recovery of ammonium (Table 2.2). Furthermore, expression of a gene encoding the dipeptide transporter DppB (*msmeg_1085*) was elevated 7.7-fold (FDR<0.1%). Dipeptide transporters facilitate the import of dipeptides into the cytoplasm, where they can be hydrolyzed in order to replenish the intracellular amino acid pool or serve for incorporation into proteins [159]. The importance of (di-)peptides as nitrogen donors in mycobacteria remain to be explored, whereas previous work suggested that peptides cannot be used as carbon source [160].

Table 2.1. List of genes involved in uptake of nitrogen compounds that are differentially expressed in response to nitrogen limitation in *M. smegmatis* mc^{2}155

Predicted transported substrate	mc^{2}155 locus[a]	Expression ratio[b]	FDR[c]	Description
Amino acids	msmeg_2525	13.18	8.41E-26	amino acid permease
	msmeg_2184	7.09	4.86E-11	amino acid permease
	msmeg_6735	6.11	1.56E-13	amino acid permease
	msmeg_2981	5.81	5.53E-29	branched-chain amino acid ABC transporter permease
	msmeg_1613[d]	4.88	3.38E-07	polar amino acid ABC transporter inner membrane protein
	msmeg_3231	2.16	1.80E-02	cysteine ABC transporter permease/ATP-binding protein
	msmeg_6876	2.14	9.32E-03	branched-chain amino acid transport ATP-binding protein
	msmeg_3203[d]	0.33	4.71E-05	transporter LysE family protein
Ammonium	msmeg_4635	7.49	9.82E-37	ammonium transporter
	msmeg_6259	2.70	2.06E-04	ammonium transporter
	msmeg_2425	2.70	3.95E-11	ammonium transporter
Nucleotides	msmeg_2570	6.79	4.33E-08	xanthine / uracil permease
	msmeg_5730	6.55	1.50E-26	permease for cytosine / purines / uracil / thiamine / allantoin
	msmeg_4011	6.30	2.56E-12	pyrimidine permease RutG
	msmeg_6660	4.22	1.88E-17	permease for cytosine / purines / uracil / thiamine / allantoin
	msmeg_1177	4.22	4.90E-08	cytosine / purines / uracil / thiamine / allantoin permease
	msmeg_1293	3.98	1.09E-17	xanthine / uracil permease

	msmeg_3402	2.23	3.57E-06	cytosine permease
Nitrate	msmeg_0433	3.67	1.73E-04	nitrite extrusion protein

[a] Locus number of gene in *M. smegmatis* mc^2155

[b] Mean gene expression ratio of three biological replicates

[c] *P*-values of gene expression ratio from three biological replicates were corrected for multiple testing using the Benjamini and Hochberg False Discovery Rate

[d] Genes in the nitrogen regulated transcriptome that were discovered in this study

2.3.4. Strong induction of nucleotide catabolism and recovery of ammonium

In our analysis, we identified a novel set of genes that were significantly upregulated upon nitrogen limitation and appeared to be involved in the breakdown of nucleotides and inorganic nitrogen sources (Table 2.2). Nucleotides contain a high level of nitrogen and their degradation plays an important role in nitrogen metabolism in a large number of different microorganisms (e.g. *Bacillus subtilis*), where purine catabolic pathways are described as an alternative pathway of nitrogen utilisation once primary nitrogen sources are exhausted [24]. Genes encoding for proteins involved in the degradation of purine nucleotides were previously reported to be upregulated in *M. smegmatis* during nitrogen limitation [142]. The enzyme guanine deaminase is upregulated 3.7-fold (FDR<0.1%) catalysing the reaction from guanine to xanthine and releasing 1 mol of ammonium per 1 mol guanine. The three enzymes, uricase (up 2.8-fold, FDR<0.1%), transthyretin (up 3.5-fold, FDR<0.1%) and 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase (up 3.9-fold, FDR<0.1%) catalyse the subsequent steps of urate degradation resulting in the production of allantoin. This metabolite is then further degraded with the concomitant production of ureidoglycolate and urea, which is broken down in either a one-step reaction via urease or a two-step reaction via urea carboxylase/allophanate hydrolase enzyme complex to ammonium (Figure 2.5). Transcription of the urease-encoding genes was elevated also in batch culture, while genes encoding for an allophanate hydrolase (*msmeg_0435-msmeg_0436*) were upregulated only under nitrogen depletion in continuous culture [142].

We observed an upregulation of the pyrimidine nucleotide degradation gene cluster and detected differential expression of a third copy of a putative phenylhydantoinase (*msmeg_3553*; up 3-fold, FDR<0.1%) that plays an important role in the degradation of pyrimidine nucleotides. Catabolism of pyrimidine nucleotides can occur through three different pathways; a reductive pathway in which uracil is reduced to two molecules of CO_2 and one molecule of ammonium and β-alanine, respectively, an oxidative pathway and a recently identified Rut pathway [161]. *M. smegmatis* harbors a reductive uracil

degradation pathway [162], which is induced upon nitrogen limitation. However, not all of the genes in this pathway have been identified in *M. smegmatis*. After the import of uracil by several nucleotide permeases, the first step of this reductive pathway is catalysed by a dihydropyrimidine dehydrogenase that is able to reduce uracil and thymine to 5,6-dihydrouracil and 5,6-dihydrothymine (Figure 2.5), followed by a reaction catalysed by a phenylhydantoinase. This enzyme hydrolyzes the opening of the heterocyclic ring and β-alanine synthase (*msmeg_3555*; up 1.9-fold, FDR<1%) is catalysing the last step of this pathway (Figure 2.5).

Table 2.2. List of genes involved in nitrogen metabolism that are differentially expressed in response to nitrogen limitation in *M. smegmatis* mc^{2}155

Metabolic pathway	mc^{2}155 locus[a]	Expression ratio[b]	FDR[c]	Description
Amino acids	msmeg_2526	8.68	7.15E-35	tyramine oxidase
	msmeg_3993	5.19	3.03E-14	Asp/Glu racemase
	msmeg_0567	4.89	2.27E-12	selenophosphate synthetase
	msmeg_3973	3.85	3.71E-05	N-methylhydantoinase
	msmeg_4459[d]	3.22	1.59E-03	agmatinase
	msmeg_3317	3.06	1.59E-08	dihydrodipicolinate reductase N-terminal domain-containing protein
	msmeg_6261	2.73	4.86E-04	glutamine amidotransferase
	msmeg_6260	2.63	1.54E-04	glutamine synthetase
	msmeg_2493	2.42	3.78E-03	aminotransferase class I and class II family protein
	msmeg_6197[d]	2.40	1.58E-02	diaminopimelate decarboxylase
	msmeg_2494	2.20	8.20E-03	Xaa-Pro aminopeptidase
	msmeg_5374	2.18	4.69E-05	glutamate-ammonia ligase
	msmeg_6263	2.07	1.21E-02	glutamate synthase
	msmeg_2100[d]	2.02	9.01E-03	peptidase M20/M25/M40
	msmeg_6256[d]	0.49	7.47E-06	aspartate-semialdehyde dehydrogenase
	msmeg_5612[d]	0.40	1.25E-07	amino acid acetyltransferase
	msmeg_2691[d]	0.40	6.92E-03	N-acetylglutamate synthase
	msmeg_1762[d]	0.32	4.60E-07	piperideine-6-carboxylic acid dehydrogenase
	msmeg_5454[d]	0.28	7.84E-08	choloylglycine hydrolase
	msmeg_5119[d]	0.25	5.76E-09	1-pyrroline-5-carboxylate dehydrogenase
	msmeg_0019[d]	0.21	5.06E-23	amino acid adenylation protein

	msmeg_5117[d]	0.19	5.23E-07	proline dehydrogenase
	msmeg_1414[d]	0.19	5.97E-03	amidinotransferase
	msmeg_1413[d]	0.18	1.92E-05	ornithine-oxo-acid transaminase
	msmeg_0022[d]	0.18	3.49E-28	L-ornithine 5-monooxygenase
	msmeg_0021[d]	0.18	1.83E-21	aspartate alpha-decarboxylase
	msmeg_1764[d]	0.13	1.45E-03	L-lysine aminotransferase
Nucleotides	msmeg_4012	9.42	1.31E-24	phenylhydantoinase
	msmeg_5729	7.24	5.08E-19	hydantoin racemase
	msmeg_1294	3.89	2.39E-17	OHCU decarboxylase
	msmeg_1296	3.78	2.28E-07	uricase
	msmeg_2748	3.69	1.44E-05	soluble pyridine nucleotide transhydrogenase
	msmeg_1298	3.67	6.88E-12	guanine deaminase
	msmeg_3996	3.53	3.56E-13	phenylhydantoinase
	msmeg_1295	3.52	2.55E-12	transthyretin
	msmeg_3553[d]	2.95	1.98E-04	phenylhydantoinase
	msmeg_6116	2.64	1.09E-08	OHCU decarboxylase
	msmeg_3473	2.55	6.53E-03	uracil phosphoribosyltransferase
	msmeg_5727	2.21	4.92E-06	allantoicase
Urea	msmeg_2187	3.33	5.04E-06	urea amidolyase
	msmeg_1425[d]	2.60	2.80E-08	creatininase subfamily protein
	msmeg_3623	2.45	1.22E-06	urease accesory protein UreG
	msmeg_0435[d]	2.40	6.48E-06	allophanate hydrolase subunit 2
	msmeg_0436[d]	2.17	1.19E-04	allophanate hydrolase subunit 1
	msmeg_2189	2.16	9.92E-03	allophanate hydrolase

	msmeg_3624	2.11	1.40E-03	urease accesory protein UreF
	msmeg_3627	2.06	3.27E-04	urease subunit γ
C-N bond	msmeg_5358	7.23	9.62E-35	acetamidase/formamidase
	msmeg_5359	4.64	6.37E-14	cyanate hydratase
	msmeg_0571	4.06	1.08E-17	carbon-nitrogen hydrolase
	msmeg_3995	3.99	1.09E-13	N-carbamoyl-L-amino acid amidohydrolase
	msmeg_6266	3.68	5.45E-18	thiocyanate hydrolase subunit β
	msmeg_4367	3.67	2.87E-17	formamidase
	msmeg_6267	3.42	2.79E-15	thiocyanate hydrolase subunit γ
	msmeg_0566	2.96	3.15E-06	aliphatic amidase
	msmeg_4381	2.58	1.56E-08	amidase
	msmeg_3403	2.38	2.83E-06	formamidase
	msmeg_5373	2.35	1.43E-04	nitrilase 2
	msmeg_0443[d]	2.28	1.82E-02	carbon-nitrogen hydrolase
	msmeg_3397	2.26	3.75E-04	acetylornithine deacetylase
	msmeg_2491	2.25	3.45E-02	acetylornithine deacetylase
Nitrate/Nitrite	msmeg_0427	4.42	1.52E-03	nitrite reductase, large subunit
	msmeg_0332[d]	0.16	7.18E-10	2-nitropropane dioxygenase
CoA biosynthesis	msmeg_6097[d]	0.49	9.14E-05	pantoate-β-alanine ligase

[a] Locus number of gene in *M. smegmatis* mc^2155

[b] Mean gene expression ratio of three biological replicates

[c] *P*-values of gene expression ratio from three biological replicates were corrected for multiple testing using the Benjamini and Hochberg False Discovery Rate

[d] Genes in the nitrogen regulated transcriptome that were discovered in this study

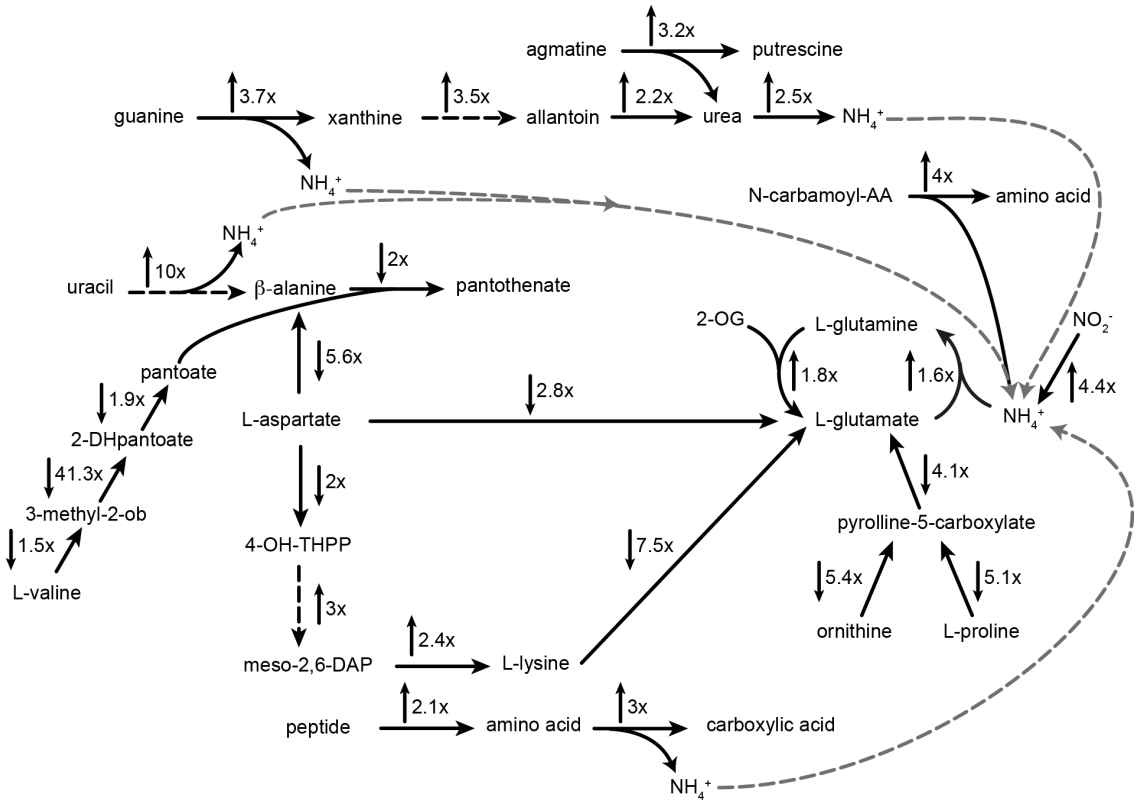


Figure 2.5. Diagram of differentially expressed pathways involved in nitrogen metabolism and ammonium assimilation in response to nitrogen limitation in *M. smegmatis*. Shown are selected metabolic pathways that are discussed. Fold change in gene expression and predicted directionality of reaction are indicated by numbers and arrows. Solid black arrows indicate one-step reactions and dotted black arrows indicate multi-step reactions. This overview is not exhaustive. (2-DHpantoate: 2-dehydropantoate; 2-OG: 2-oxoglutarate; 3-methyl-ob: 3-methyl-2-oxobutanoate; 4-OH-THPP: 4-hydroxy-2,3,4,5-tetrahydrodipicolinate; AA: amino acid; meso-2,6-DAP: meso-2,6-diaminopimelate; NH_4^+: ammonium; NO_2^-: nitrite).

2.3.5. Amino acid catabolism is strongly downregulated under nitrogen limitation

A large cohort of genes, particularly in the catabolism of the amino acids alanine, aspartate, valine, proline and lysine were downregulated 5-40-fold in response to nitrogen depletion (Table 2.2), (Figure 2.5). Interestingly, most of these genes showed an elevated transcription in batch culture. The amino acid alanine has a substantial importance as a central metabolite in bacterial metabolism and also its role in the synthesis of peptidoglycan as D-/L-alanine is essential for bacteria. Several pathways have been described contributing to the biosynthesis of L-alanine and β-alanine, e.g. catabolism of amino acids (i.e. valine and cysteine) and transamination of pyruvate. In *M. smegmatis*, genes encoding for all of the aforementioned pathways were differentially expressed in response to nitrogen limitation (Table 2.2). The conformation of the amino acid β-alanine does not allow its incorporation into proteins, but it serves together with pantoate as precursor of coenzyme A (CoA) biosynthesis, which is essential for a functional TCA cycle, as well as fatty acid and cholesterol biosynthesis. Degradation of purine nucleotides via an L-aspartate-alpha-decarboxylase (PanD) results in production of β-alanine and PanD was identified as the predominant pathway of β-alanine synthesis in the closely related *C. glutamicum*, where a *panD* mutant exhibited β-alanine auxotrophy [163]. However, expression of *panD* was strongly downregulated in *M. smegmatis* under nitrogen limitation. Valine degradation via the intermediates 3-methyl-2-oxobutanoate and 2-dehydropantoate to (R)-pantoate was repressed at the same time, suggesting the demand to prevent unnecessary consumption of amino acids for CoA biosynthesis (Figure 2.5).

A second pathway of L-aspartate (Asp) catabolism was differentially expressed in *M. smegmatis* under nitrogen limitation, which is linked to the concomitant biosynthesis of lysine. This pathway is a nine-step reaction including important metabolites such as L-aspartate semialdehyde (homoserine biosynthesis) and meso-2,6-diaminopimelate (constituent of bacterial cell walls). Interestingly, the initial steps of aspartate catabolism were repressed, while the degradation of meso-2,6-diaminopimelate to lysine was upregulated (Figure 2.5). Lysine can act as donor of an amino group by transferring an

ammonium group to 2-oxoglutarate to form glutamate under nitrogen excess, however, this pathway of lysine catabolism via a lysine aminotransferase was downregulated 7.5-fold (FDR<1%), indicating an intracellular accumulation of aspartate and lysine and suggesting a secondary function of these amino acids [164,165]. Previous studies discussed the importance of intracellular lysine to control growth rate in mycobacteria and suggested a link between lysine accumulation and fatty acid metabolism [166]. Proline has been shown to serve as a mechanism for methylglyoxal detoxification, when anabolic and catabolic processes were imbalanced [165] and we showed that proline degradation was repressed under nitrogen depletion.

2.3.6. A large number of transcriptional regulatory systems are differentially expressed in response to nitrogen limitation

We identified 26 differentially expressed transcriptional regulators that are either directly or indirectly responding to nitrogen limitation in *M. smegmatis* (Table 2.3). Only a handful of these regulators have been characterised, including the nitrogen regulatory protein P_{II} (*msmeg_2426*) and the P_{II} adenylyl transferase (*msmeg_2427*). The mycobacterial copy of the P_{II} protein is not required for the regulation of the glutamine synthetase activity and does not act as regulator of the transcriptional response to nitrogen limitation [34]. This is in contrast to *C. glutamicum*, where the P_{II} protein was identified as the sole signal transduction protein, binding to AmtR and releasing this repressor from its target DNA, in order to allow transcription of genes involved in nitrogen uptake, assimilation and metabolism [39].

Another well-described transcriptional regulator is the OmpR-type response regulator GlnR, which has been identified as a mediator of the transcriptomic response to nitrogen limitation in *M. smegmatis* [43]. Determination of the GlnR regulon, by combining expression profiling of *M. smegmatis* wild-type and a Δ*glnR* deletion mutant under nitrogen-starvation conditions in batch culture revealed a total of 103 genes directly controlled by GlnR [43]. A large portion of these genes (72%) were also differentially expressed under nitrogen limitation in our continuous culture model, however,

GlnR itself was not among them, indicating a posttranslational regulatory mechanism for GlnR like in *S. coelicolor* [41]. Comparison of the differentially expressed transcriptional regulatory proteins in batch culture versus continuous culture showed an overlap of six genes with four of these under control of GlnR (Table 2.3). We further observed 16 transcriptional regulatory proteins that were differentially expressed in continuous culture (Table 2.3). The TetR-like transcriptional regulator AmtR showed a 1.7-fold upregulation under nitrogen-replete conditions and was therefore outside the selected cut-off. However, in our RNA-seq analysis we identified 127-fold downregulation of an antisense transcript (asRNA) of almost the entire *msmeg_4300* gene under nitrogen-depleted conditions (Table 2.4). These data suggest a mechanism of post-transcriptional regulation by an asRNA where transcription of *msmeg_4300* is enhanced upon nitrogen excess and modulates the translation efficiency of the AmtR encoding gene.

We identified several other cases of potential non-coding small RNAs that might regulate expression of genes involved in fatty acid and central carbon metabolism under nitrogen limitation (Table 2.4). In our RNA-seq analysis, we observed a 102-fold downregulation of an asRNA of *msmeg_4299* gene under nitrogen depletion, while *msmeg_4299* showed a 2.7-fold upregulation under nitrogen limitation (Table 2.4). The intergenic region *msmeg_3131-msmeg_3132* comprises the promotor regions of the genes *msmeg_3131* (long-chain acyl-CoA synthetase) and *msmeg_3132* (DNA-binding protein) and revealed a 4.7-fold increase in transcription under nitrogen excess (Table 2.4). Interactions between sRNAs and the 5'UTR of mRNAs affect the translation efficiency, while an interaction with the 3'UTR of mRNAs usually do not affect translation, but mRNA stability [167]. These findings are supported by a strong upregulation of genes involved in the degradation of fatty acids as well as the downregulation of genes involved in the TCA cycle as both metabolic pathways are CoA-dependent (Table 2.5, supplementary data).

Table 2.3. List of genes involved in regulatory mechanisms in response to nitrogen limitation in *M. smegmatis* mc^{2}155

mc^{2}155 locus[a]	Expression ratio[b]	FDR[c]	Description
msmeg_1082	7.04	1.06E-08	two-component regulator HTH luxR-type DNA binding domain
msmeg_3997	3.46	4.55E-15	regulatory protein
msmeg_5731	3.44	1.87E-07	GntR family transcriptional regulator
msmeg_1597	3.29	1.59E-14	transcription factor WhiB
msmeg_6236	3.27	1.32E-12	two-component system - regulatory protein
msmeg_6238	2.80	2.81E-10	two-component system - sensor kinase
msmeg_2427	3.19	2.01E-14	P_{II} uridylyl-transferase
msmeg_2426	2.18	5.48E-07	nitrogen regulatory protein P_{II}
msmeg_4006	3.11	1.56E-08	CdaR family transcriptional regulator
msmeg_4368	3.11	1.49E-13	regulatory protein - FmdB family
msmeg_3297	2.78	1.21E-05	CadC family transcriptional regulator
msmeg_3298	2.36	4.75E-03	response regulator receiver domain-containing protein
msmeg_6198	2.41	1.85E-05	DNA-binding protein
msmeg_6824	2.16	1.03E-02	MarR family transcriptional regulator
msmeg_0778	2.11	1.33E-02	transcriptional regulator
msmeg_1420	2.07	1.46E-02	transcriptional regulatory protein
msmeg_5673	0.49	9.39E-05	transcriptional regulator
msmeg_3177	0.49	4.19E-02	transcriptional regulatory protein
msmeg_4394	0.48	5.57E-04	LysR family transcriptional regulator
msmeg_6789	0.45	2.14E-02	GntR family transcriptional regulator
msmeg_0473	0.38	2.21E-04	LuxR family transcriptional regulator

msmeg_5987	0.35	9.99E-03	two-component regulator
msmeg_6555	0.33	6.11E-07	TetR family transcriptional regulator
msmeg_0051	0.31	1.59E-05	transcription factor WhiB family protein
msmeg_1953	0.27	1.77E-06	transcription factor WhiB
msmeg_0622	0.19	4.57E-08	DNA-binding protein

[a] Locus number of gene in *M. smegmatis* mc^2155

[b] Mean gene expression ratio of three biological replicates

[c] *P*-values of gene expression ratio from three biological replicates were corrected for multiple testing using the Benjamini and Hochberg False Discovery Rate

Bold: within GlnR regulon; *Italic*: differentially expressed in the same direction in batch culture versus continuous culture; Underlined: inversed expression in batch culture versus continuous culture

Table 2.4. List of differentially expressed intergenic and antisense-strand regions (Expression ratio >4.0, FDR <0.05)

mc^2 155 locus[a]	Expression ratio[b]	FDR[c]	Description	Note
rev_MSMEG_6817	17.42	4.94E-38	sigma-70 family RNA polymerase sigma factor	antisense RNA?
MSMEG_2525-MSMEG_2526	8.95	1.57E-27	between amino acid permease AND copper methylamine oxidase	
rev_MSMEG_6242-MSMEG_6243	8.71	9.26E-49	between alcohol dehydroge-nase AND protein of unknown function	
rev_MSMEG_5358-MSMEG_5359	6.74	1.18E-24	between acetamidase/formamidase AND cyanate hydratase	
MSMEG_1424-MSMEG_1425	5.76	3.50E-28	between FMN-dependent dehydrogenase AND creatininase	
rev_MSMEG_4634	5.33	2.14E-02	hypothetical protein	
MSMEG_3131-MSMEG_3132	4.74	1.25E-21	between long-chain acyl-CoA synthetase AND DNA binding protein	small RNA?
rev_MSMEG_1596-MSMEG_1597	4.55	1.99E-13	between transcriptional regulator AND transcription factor WhiB	
rev_MSMEG_6239-MSMEG_6240	4.46	5.80E-15	between 1,3-propanediol dehydrogenase AND protein containing vWa	
rev_MSMEG_2982-MSMEG_2983	4.20	6.14E-09	between amino acid transporter AND hypothetical protein	
rev_MSMEG_3811-MSMEG_3812	0.24	3.94E-02	between universal stress protein AND acyl-CoA-thioesterase	
MSMEG_0911-MSMEG_0912	0.22	2.14E-19	between isocitrate lyase AND 3-hydroxybutyryl-CoA dehydrogenase	
rev_MSMEG_1999-MSMEG_2000	0.19	1.50E-12	between hypothetical AND ketosteroid-isomerase related protein	
rev_MSMEG_0020-MSMEG_0021	0.19	5.75E-02	between periplasmic binding protein AND aspartate 1-decarboxylase	
rev_MSMEG_1737	0.13	3.57E-35	RNA pseudouridine synthase	antisense RNA?
rev_MSMEG_0910	0.12	2.20E-07	hypothetical protein	antisense RNA?

rev_MSMEG_1738-MSMEG_1740	0.08	4.06E-04	between conserved transmembrane protein AND dehydrogenase/reductase	
MSMEG_4297-MSMEG_4298	0.03	4.93E-53	between acyltransferase AND 3-methyl-2-oxobutanoate hydroxymethyltransferase	
rev_MSMEG_4299	0.01	6.23E-50	enoyl-CoA hydratase/isomerase	antisense RNA?

[a] Locus number of gene in *M. smegmatis* mc^2155

[b] Mean gene expression ratio of three biological replicates

[c] *P*-values of gene expression ratio from three biological replicates were corrected for multiple testing using the Benjamini and Hochberg False Discovery Rate

Interestingly, three regulatory proteins (one putative two-component system and two putative orphan response regulators) showed a different expression profile in continuous culture compared to batch culture. The operon *msmeg_6236-msmeg_6238* encodes for a two-component system with unknown function and the gene *msmeg_3297* encodes for a CadC-like transcriptional regulator that has been linked to lysine-dependent acid adaptation in response to an acidic external pH in *E. coli* [168]. No change in external pH was observed in our experiments, suggesting the CadC-like protein in *M. smegmatis* was performing a different function.

2.3.7. Conclusions

Herein we report the transcriptomic response of *M. smegmatis* to nitrogen limitation in a continuous culture model at a defined growth rate (Figure 2.6). We show that amino acid metabolism plays an important role in the adaptation of *M. smegmatis* to nitrogen depletion (Figure 2.6). We identified 16 novel transcriptional regulators that were either directly or indirectly involved in the global transcriptomic response of *M. smegmatis* to nitrogen limitation. We identified several non-coding RNAs that might be involved in transcriptional or post-transcriptional regulation of gene expression and we propose a regulatory mechanism through a trans-acting asRNA for the AmtR protein in *M. smegmatis*. Comparison of our chemostat data to the global transcriptomic response in batch culture nitrogen run out experiments revealed that only 17% of the previously described nitrogen-regulated genes overlapped between batch and continuous culture. Our data highlight the versatile metabolic capability of *M. smegmatis* and provide a molecular framework for understanding how environmental mycobacteria respond to nitrogen-depleted environments.

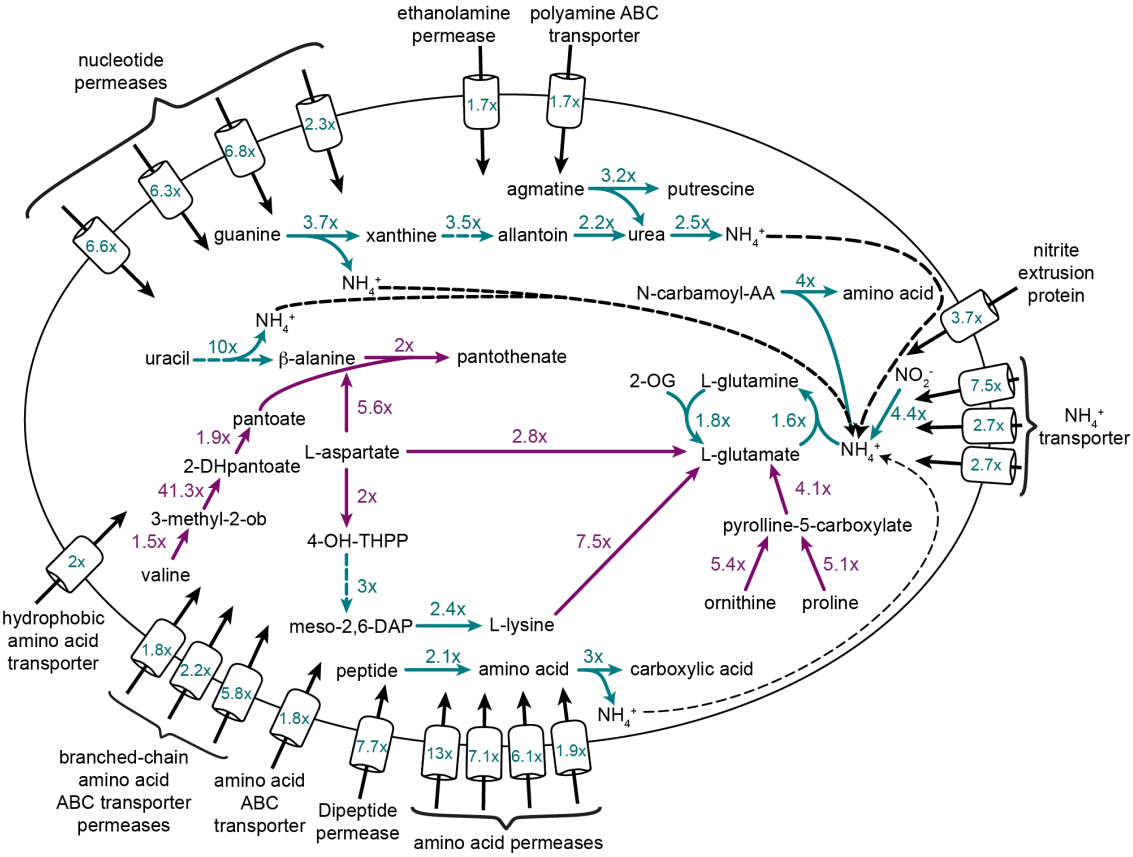


Figure 2.6. Schematic overview of differentially expressed nitrogen uptake systems and pathways involved in nitrogen metabolism in response to nitrogen limitation in *M. smegmatis*. Shown are selected upregulated (cyan) and downregulated (purple) genes that are directly involved in nitrogen uptake, metabolism and ammonium assimilation. Fold change in gene expression and predicted directionality of reaction are indicated by numbers and arrows. This overview is not exhaustive. (2-DHpantoate: 2-dehydropantoate; 2-OG: 2-oxoglutarate; 3-methyl-ob: 3-methyl-2-oxo-butanoate; 4-OH-THPP: 4-hydroxy-2,3,4,5-tetrahydrodipicolinate; AA: amino acid; meso-2,6-DAP: meso-2,6-diaminopimelate; NH_4^+: ammonium; NO_2^-: nitrite).

2.4. Materials and Methods

2.4.1. Bacterial strains, media and growth conditions

M. smegmatis mc^2155 was grown in LBT for standard growth and culture maintenance or modified HdB minimal medium. Briefly, Tween-80 that can be metabolised as carbon source was replaced by Tyloxapol and $(NH_4)_2SO_4$ was replaced by NH_4Cl and K_2SO_4 to exclude a simultaneous limitation of sulphur. Modified HdB minimal medium was prepared with 0.2% (w/v) glycerol (unless otherwise stated) as sole carbon source, NH_4Cl in various concentrations as sole nitrogen source, 11.74 mM K_2SO_4 as sulphur source, 26.5 mM Na_2HPO_4 and 11 mM KH_2PO_4 as phosphate buffer, 10 mL per liter of a 100x trace metal solution and 0.05% (w/v) Tyloxapol at 37°C with agitation (200 rpm). Aerobic starter cultures were inoculated to an initial optical density (OD_{600}) of 0.05 and grown with agitation (200 rpm) at 37°C. At an OD_{600} between 0.5 – 1.0, a sample was withdrawn to inoculate the bioreactor to a starting OD_{600} of 0.005. Before inoculation, the culture medium in the bioreactor was stirred at 400 rpm at 37°C for 30 minutes to fully saturate the medium with oxygen. The oxygen probe was then set to 100%. After inoculation the culture was left in batch mode maintaining 50% oxygen saturation until the OD_{600} reached around 80% of the steady-state OD_{600} and was then switched to chemostat mode at a dilution rate of 0.12 h^{-1} corresponding to a doubling time t_d of 5.7 hrs. After entering steady-state, the chemostat was left running for at least three volume change before cell harvest. Samples to measure residual glycerol concentration and ammonium concentration were taken in intervals of four to six hours. Glycerol concentration was measured according to Garland and Randle [169] and ammonium concentration was measured according to Weatherburn [170]. After entering steady-state, the chemostat was left running for at least three volume changes before cell harvest.

2.4.2 RNA extraction and reverse transcriptase PCR

Cells were lysed by four cycles at 4800 rpm for 30 s in a mini-Beadbeater (Biospec Products), with 30 s on ice between each of the cycles.

Total RNA was extracted using TRIzol® reagent (Ambion) according to the manufacturers instructions. DNA was removed by treatment with 3 U RNase-free DNase using the TURBO DNA-free kit (Ambion) according to the manufacturers instructions. The quality of the RNA was checked with the Bioanalyzer (RIN >9) and the concentration was determined using a NanoDrop ND-100 spectrophotometer.

Next, depletion of ribosomal RNA and strand-specific library preparation were performed, using the EpiCentre ScriptSEq™ Complete Kit for Bacteria according to manufacturers instructions. Quantification of nucleic acids was performed using a Qubit® Fluorometer to ensure DNA contamination of the samples was less than 10% and libraries were created. Prepared libraries have undergone a quality control using an Agilent 2100 Bioanalyzer: DNA 1000 Labchip, Quant-iT dSDNA HS Assay for quantification and Quant-iT RNA Assay and Quant-iT Protein Assay for percentage contamination check, using an Invitrogen Qubit® Fluorometer. Libraries were run on a Illumina MiSeq 300 cycle Kit_v2 with a paired-end (PE) read length of 2x150 and a PhiX control library was also loaded and used as control for the run.

2.4.3. Analysis of RNA-sequencing data

Adapter sequences were removed from raw fastq files using Flexbar [171] and reads shorter than 50 bp were discarded. Sequences were then mapped against the *M. smegmatis* genome (GenBank NC_008596.1) using Bowtie2 with the ”very sensitive” option. Counts for each gene, intergenic regions and counts on reverse strand were calculated with featureCounts [172], taking into account the reads’ strand-specificity. Statistical analysis was performed using DESeq (Figure 2.7) [173]. Genes with a base mean expression value >20 were excluded from further analysis. Venn diagrams were generated using the BioVenn program [174]. Gene functions were assigned using public databases (NCBI [175], UniProt [176], KEGG [177], Biocyc Database Collection [178]). The Biocyc Database Collection was used for assignment of gene functions and the functional assignment was confirmed using other databases [175-178]. The RNA sequencing run report is shown in Table 2.6.

2.5. Supplementary Data

Table 2.5. Complete list of genes that are differentially expressed in response to nitrogen limitation in continuous culture in *M. smegmatis* mc^{2}155

mc^{2}155 locus[a]	Expression ratio[b]	FDR[c]
MSMEG_0014	0.261	1.52E-14
MSMEG_0017	0.188	2.29E-09
MSMEG_0018	0.133	1.09E-15
MSMEG_0019	0.212	5.06E-23
MSMEG_0020	0.261	1.12E-18
MSMEG_0021	0.178	1.83E-21
MSMEG_0022	0.180	3.49E-28
MSMEG_0051	0.305	1.59E-05
MSMEG_0055	0.316	2.89E-13
MSMEG_0063	0.465	2.01E-05
MSMEG_0064	0.489	1.21E-05
MSMEG_0077	0.400	4.44E-06
MSMEG_0078	0.379	1.91E-10
MSMEG_0082	0.500	3.57E-05
MSMEG_0114	0.352	3.39E-02
MSMEG_0116	0.301	3.02E-06
MSMEG_0157	0.495	1.07E-03
MSMEG_0168	0.386	1.68E-02
MSMEG_0172	0.307	2.77E-11
MSMEG_0194	0.461	9.91E-07
MSMEG_0282	2.021	1.58E-03
MSMEG_0312	4.852	4.71E-07
MSMEG_0313	6.150	4.81E-20
MSMEG_0314	6.093	1.59E-27
MSMEG_0332	0.164	7.18E-10
MSMEG_0333	0.235	1.30E-11
MSMEG_0334	0.289	1.38E-03
MSMEG_0335	0.186	9.20E-09
MSMEG_0336	0.190	1.19E-10
MSMEG_0337	0.210	2.02E-14
MSMEG_0338	0.208	2.28E-15
MSMEG_0339	0.351	8.14E-03
MSMEG_0360	0.497	1.05E-04
MSMEG_0385	2.139	4.42E-05
MSMEG_0406	4.696	4.60E-17
MSMEG_0427	4.424	1.52E-03
MSMEG_0431	3.372	2.06E-04
MSMEG_0432	2.599	3.43E-05
MSMEG_0433	3.672	1.73E-04
MSMEG_0435	2.400	6.48E-06
MSMEG_0436	2.166	1.19E-04
MSMEG_0443	2.277	1.82E-02
MSMEG_0473	0.374	2.21E-04
MSMEG_0546	0.453	2.12E-05
MSMEG_0565	2.673	5.87E-03
MSMEG_0566	2.959	3.15E-06
MSMEG_0567	4.890	2.27E-12
MSMEG_0568	4.973	1.25E-18
MSMEG_0569	3.955	7.80E-18
MSMEG_0570	3.575	9.13E-09
MSMEG_0571	4.056	1.08E-17
MSMEG_0572	4.428	4.20E-20
MSMEG_0615	0.217	1.40E-12
MSMEG_0616	0.194	1.73E-04
MSMEG_0617	0.165	3.62E-08
MSMEG_0618	0.159	2.57E-04
MSMEG_0619	0.142	6.03E-22
MSMEG_0620	0.185	3.87E-25
MSMEG_0621	0.185	2.16E-24
MSMEG_0622	0.185	4.57E-08
MSMEG_0623	0.183	6.72E-18
MSMEG_0624	0.203	1.16E-14
MSMEG_0626	0.201	9.07E-11
MSMEG_0643	0.484	9.45E-04
MSMEG_0657	2.331	2.38E-03
MSMEG_0687	0.497	3.82E-02
MSMEG_0688	0.354	2.78E-10
MSMEG_0726	0.467	1.30E-05
MSMEG_0728	0.460	4.59E-06
MSMEG_0752	2.328	3.33E-07
MSMEG_0778	2.110	1.33E-02
MSMEG_0779	2.188	3.01E-04
MSMEG_0780	2.013	5.59E-03
MSMEG_0911	0.120	5.16E-19
MSMEG_0912	0.452	1.83E-06
MSMEG_1082	7.042	1.06E-08
MSMEG_1145	0.492	6.86E-03
MSMEG_1153	2.057	6.59E-03
MSMEG_1167	0.418	2.68E-03
MSMEG_1169	0.262	1.57E-08
MSMEG_1177	4.219	4.90E-08
MSMEG_1184	2.880	6.24E-07
MSMEG_1193	0.260	2.32E-08
MSMEG_1212	0.447	1.01E-02
MSMEG_1219	2.801	1.25E-03
MSMEG_1220	3.092	7.02E-04
MSMEG_1226	2.140	8.87E-03
MSMEG_1232	7.135	1.06E-04
MSMEG_1292	2.367	8.75E-03
MSMEG_1293	3.978	1.09E-17
MSMEG_1294	3.888	2.39E-17
MSMEG_1295	3.520	2.55E-12
MSMEG_1296	2.784	2.28E-07
MSMEG_1297	3.546	2.57E-12
MSMEG_1298	3.665	6.88E-12
MSMEG_1299	4.452	4.06E-13

MSMEG_1413	0.186	1.92E-05
MSMEG_1414	0.189	5.97E-03
MSMEG_1419	0.314	1.71E-06
MSMEG_1420	2.068	1.46E-02
MSMEG_1421	7.662	6.27E-07
MSMEG_1422	7.043	4.61E-15
MSMEG_1423	8.774	1.60E-47
MSMEG_1424	8.968	1.41E-48
MSMEG_1425	2.598	2.80E-08
MSMEG_1426	2.203	4.44E-06
MSMEG_1530	0.155	2.05E-08
MSMEG_1597	3.288	1.59E-14
MSMEG_1612	4.755	6.14E-08
MSMEG_1613	4.877	3.38E-07
MSMEG_1627	2.277	4.93E-03
MSMEG_1669	0.456	8.34E-07
MSMEG_1670	0.427	1.82E-08
MSMEG_1671	0.459	1.26E-06
MSMEG_1672	0.371	4.45E-09
MSMEG_1738	0.095	2.34E-23
MSMEG_1762	0.320	4.60E-07
MSMEG_1764	0.133	1.45E-03
MSMEG_1821	2.787	3.15E-05
MSMEG_1872	0.497	4.80E-05
MSMEG_1885	0.465	4.88E-04
MSMEG_1887	0.395	2.26E-02
MSMEG_1953	0.272	1.77E-06
MSMEG_1987	3.411	1.03E-06
MSMEG_1988	5.614	8.14E-07
MSMEG_1989	5.380	2.08E-09
MSMEG_1990	6.613	1.99E-10
MSMEG_1999	0.170	4.04E-26
MSMEG_2010	0.475	1.53E-06
MSMEG_2027	0.498	5.97E-03
MSMEG_2097	2.037	4.54E-03
MSMEG_2100	2.024	9.01E-03
MSMEG_2113	4.268	1.59E-06
MSMEG_2116	2.744	5.38E-10
MSMEG_2121	6.964	3.71E-30
MSMEG_2122	7.826	2.85E-32
MSMEG_2123	6.788	2.00E-32
MSMEG_2124	8.548	1.15E-25
MSMEG_2130	0.166	2.72E-25
MSMEG_2131	0.149	1.06E-12
MSMEG_2132	0.307	3.87E-02
MSMEG_2184	7.086	4.86E-11
MSMEG_2187	3.327	5.04E-06
MSMEG_2188	0.284	2.33E-05
MSMEG_2189	2.159	9.92E-03
MSMEG_2346	2.708	3.29E-02
MSMEG_2425	2.695	3.95E-11
MSMEG_2426	2.180	5.48E-07
MSMEG_2427	3.190	2.01E-14
MSMEG_2451	2.297	2.31E-02
MSMEG_2456	2.091	1.48E-04
MSMEG_2457	2.198	1.44E-03
MSMEG_2462	2.196	1.03E-04
MSMEG_2491	2.253	3.45E-02
MSMEG_2493	2.423	3.78E-03
MSMEG_2494	2.198	8.20E-03
MSMEG_2522	6.074	1.77E-18
MSMEG_2523	10.737	3.18E-10
MSMEG_2524	9.873	9.33E-07
MSMEG_2525	13.177	8.41E-26
MSMEG_2526	8.676	7.15E-35
MSMEG_2527	0.241	2.06E-02
MSMEG_2542	0.112	8.45E-14
MSMEG_2569	6.809	3.10E-05
MSMEG_2570	6.785	4.33E-08
MSMEG_2601	0.458	2.69E-05
MSMEG_2691	0.400	6.92E-03
MSMEG_2748	3.694	1.44E-05
MSMEG_2757	0.445	1.99E-03
MSMEG_2798	2.209	7.63E-06
MSMEG_2926	2.522	3.33E-02
MSMEG_2977	0.412	1.17E-03
MSMEG_2978	4.997	2.55E-19
MSMEG_2979	5.102	2.72E-25
MSMEG_2980	5.357	1.84E-25
MSMEG_2981	5.814	5.53E-29
MSMEG_2982	4.664	2.72E-25
MSMEG_3117	2.157	7.69E-05
MSMEG_3163	2.302	4.30E-07
MSMEG_3164	2.203	8.55E-07
MSMEG_3177	0.493	4.19E-02
MSMEG_3203	0.332	4.71E-05
MSMEG_3231	2.160	1.80E-02
MSMEG_3232	2.498	4.66E-02
MSMEG_3233	2.284	1.39E-06
MSMEG_3297	2.783	1.21E-05
MSMEG_3298	2.362	4.75E-03
MSMEG_3300	2.393	4.33E-05
MSMEG_3301	2.529	2.62E-06
MSMEG_3302	2.701	1.55E-04
MSMEG_3317	3.056	1.59E-08
MSMEG_3357	2.177	2.61E-06
MSMEG_3397	2.256	3.75E-04
MSMEG_3401	2.158	2.73E-05
MSMEG_3402	2.225	3.57E-06
MSMEG_3403	2.375	2.83E-06
MSMEG_3465	0.381	1.41E-07
MSMEG_3472	2.144	9.01E-03
MSMEG_3473	2.548	6.53E-03
MSMEG_3553	2.951	1.98E-04
MSMEG_3554	2.538	3.55E-03
MSMEG_3560	2.479	2.00E-03
MSMEG_3566	0.306	9.28E-12
MSMEG_3611	2.097	1.65E-04
MSMEG_3616	0.449	1.16E-04
MSMEG_3619	2.540	1.50E-07
MSMEG_3622	2.417	1.65E-04
MSMEG_3623	2.454	1.22E-06
MSMEG_3624	2.105	1.40E-03
MSMEG_3627	2.064	3.27E-04
MSMEG_3629	0.297	4.29E-10
MSMEG_3683	0.484	1.40E-03
MSMEG_3811	0.317	5.78E-11
MSMEG_3973	3.846	3.71E-05
MSMEG_3992	4.216	6.31E-14
MSMEG_3993	5.190	3.03E-14
MSMEG_3994	4.625	1.33E-16
MSMEG_3995	3.985	1.09E-13

MSMEG_3996	3.529	3.56E-13	MSMEG_5230	0.288	8.95E-11
MSMEG_3997	3.468	4.55E-15	MSMEG_5231	0.307	9.65E-04
MSMEG_4006	3.112	1.56E-08	MSMEG_5245	0.492	1.84E-02
MSMEG_4008	4.081	6.27E-03	MSMEG_5358	7.232	9.62E-35
MSMEG_4010	4.785	1.50E-02	MSMEG_5359	4.638	6.37E-14
MSMEG_4011	6.297	2.56E-12	MSMEG_5373	2.348	1.43E-04
MSMEG_4012	9.422	1.31E-24	MSMEG_5374	2.184	4.69E-05
MSMEG_4013	3.505	4.60E-07	MSMEG_5403	3.379	1.21E-02
MSMEG_4298	0.024	9.41E-50	MSMEG_5418	0.343	4.05E-12
MSMEG_4299	2.397	7.52E-06	MSMEG_5420	0.422	1.05E-07
MSMEG_4301	0.498	2.01E-05	MSMEG_5454	0.281	7.84E-08
MSMEG_4303	0.415	1.40E-03	MSMEG_5478	2.073	1.94E-02
MSMEG_4365	0.474	7.09E-05	MSMEG_5480	2.462	4.66E-03
MSMEG_4367	3.673	2.87E-17	MSMEG_5505	0.333	5.97E-12
MSMEG_4368	3.110	1.49E-13	MSMEG_5519	0.385	7.17E-04
MSMEG_4381	2.577	1.56E-08	MSMEG_5557	2.638	9.17E-07
MSMEG_4382	2.568	4.07E-07	MSMEG_5612	0.404	1.25E-07
MSMEG_4394	0.482	5.57E-04	MSMEG_5613	0.464	1.58E-05
MSMEG_4396	0.444	2.06E-06	MSMEG_5634	0.489	2.21E-05
MSMEG_4456	2.094	1.02E-03	MSMEG_5659	0.494	1.05E-02
MSMEG_4459	3.222	1.59E-03	MSMEG_5660	0.493	2.96E-03
MSMEG_4501	7.407	1.94E-22	MSMEG_5673	0.493	9.39E-05
MSMEG_4509	0.138	2.17E-07	MSMEG_5726	2.033	3.58E-03
MSMEG_4510	0.134	1.61E-26	MSMEG_5727	2.213	4.92E-06
MSMEG_4511	0.104	9.63E-14	MSMEG_5728	2.341	1.34E-06
MSMEG_4512	0.102	4.71E-05	MSMEG_5729	7.242	5.08E-19
MSMEG_4513	0.088	1.25E-12	MSMEG_5730	6.550	1.50E-26
MSMEG_4514	0.112	2.96E-07	MSMEG_5731	3.438	1.87E-07
MSMEG_4515	0.117	1.07E-18	MSMEG_5884	2.004	1.22E-02
MSMEG_4516	0.239	3.91E-05	MSMEG_5912	2.716	2.84E-08
MSMEG_4524	0.238	3.49E-13	MSMEG_5921	0.462	3.40E-02
MSMEG_4536	0.418	1.42E-04	MSMEG_5987	0.349	9.99E-03
MSMEG_4567	2.032	9.01E-04	MSMEG_5988	0.470	2.83E-06
MSMEG_4570	2.455	1.39E-06	MSMEG_5989	0.450	3.12E-06
MSMEG_4635	7.493	9.82E-37	MSMEG_6031	0.399	2.50E-03
MSMEG_4636	5.518	1.85E-21	MSMEG_6033	0.104	2.21E-35
MSMEG_4637	5.662	1.30E-12	MSMEG_6096	0.476	9.12E-04
MSMEG_4638	7.357	1.28E-15	MSMEG_6097	0.490	9.14E-05
MSMEG_4639	4.194	9.81E-03	MSMEG_6098	0.471	1.02E-03
MSMEG_4643	0.064	5.35E-06	MSMEG_6116	2.637	1.09E-08
MSMEG_4685	3.107	3.69E-11	MSMEG_6193	0.487	7.59E-06
MSMEG_4702	0.463	4.42E-03	MSMEG_6196	2.080	6.02E-04
MSMEG_4727	4.847	1.98E-23	MSMEG_6197	2.399	1.58E-02
MSMEG_4728	6.164	1.20E-20	MSMEG_6198	2.413	1.85E-05
MSMEG_4729	5.394	7.49E-15	MSMEG_6211	2.195	2.83E-02
MSMEG_4730	4.644	8.79E-14	MSMEG_6232	2.446	1.54E-02
MSMEG_4731	6.597	1.16E-18	MSMEG_6236	3.268	1.32E-12
MSMEG_4732	4.849	4.05E-15	MSMEG_6237	3.958	5.05E-10
MSMEG_4756	0.304	4.50E-14	MSMEG_6238	2.802	2.81E-10
MSMEG_4757	0.268	9.23E-20	MSMEG_6239	3.134	9.63E-14
MSMEG_4920	0.274	1.33E-18	MSMEG_6240	9.952	3.24E-45
MSMEG_4921	0.386	3.78E-10	MSMEG_6241	11.098	1.47E-55
MSMEG_4965	2.354	1.80E-05	MSMEG_6242	9.345	9.11E-52
MSMEG_5029	2.120	8.84E-03	MSMEG_6256	0.486	7.47E-06
MSMEG_5047	0.451	2.84E-03	MSMEG_6259	2.699	2.06E-04
MSMEG_5084	3.110	1.00E-09	MSMEG_6260	2.634	1.54E-04
MSMEG_5102	0.409	1.85E-02	MSMEG_6261	2.728	4.86E-04
MSMEG_5117	0.192	5.23E-07	MSMEG_6262	2.144	2.84E-03
MSMEG_5119	0.249	5.76E-09	MSMEG_6263	2.065	1.21E-02
MSMEG_5184	0.437	3.03E-06	MSMEG_6266	3.675	5.45E-18
MSMEG_5187	0.167	3.59E-12	MSMEG_6267	3.420	2.79E-15

MSMEG_6268	3.827	2.95E-14
MSMEG_6346	2.105	7.32E-04
MSMEG_6419	0.147	6.69E-09
MSMEG_6446	0.358	4.86E-05
MSMEG_6447	0.414	4.31E-03
MSMEG_6554	0.423	3.04E-04
MSMEG_6555	0.328	6.11E-07
MSMEG_6602	2.796	6.24E-07
MSMEG_6660	4.222	1.88E-17
MSMEG_6734	6.021	7.57E-20
MSMEG_6735	6.105	1.56E-13
MSMEG_6789	0.449	2.14E-02
MSMEG_6816	13.107	4.60E-33
MSMEG_6824	2.160	1.03E-02
MSMEG_6876	2.144	9.31E-03
MSMEG_6904	2.033	4.67E-06
MSMEG_6914	2.519	2.52E-06

[a] Locus number of gene in *M. smegmatis* mc^2155

[b] Mean gene expression ratio of three biological replicates

[c] *P*-values of gene expression ratio from three biological replicates were corrected for multiple testing using the Benjamini and Hochberg False Discovery Rate

Table 2.6. RNA-sequencing run summary report

Level	Yield Total (G)[a]	Aligned (%)[b]	% Perfect [Num Cycles][c]	% <=3 errors [Num cycles][d]	Error rate (%)[e]	Intensity Cycle 1[f]	% Intensity Cycle 20[g]	% >= Q30[h]
Read 1[i]	2.8	0.76	78.1 [150]	98.7 [150]	0.36	173	112	95.7
Read 2[k]	0.1	0	0 [5]	0 [5]	0	662	0	95.3
Read 3[l]	2.8	0.74	79.7 [150]	98.4 [150]	0.42	42	392.6	92.7
Total	5.7	0.75	78.9	98.6	0.39	292	252.3	94.2

Level	Density[m]	Clusters[n]	Reads (M)[o]	Reads PF (M)[p]	% >= Q30	Error Rate 100 cycle (%)[q]
Read 1	1018 ± 20	94 ± 1.2	19.74	18.55	95.7	0.15±0.02
Read 2	1018 ± 20	94 ± 1.2	19.74	18.55	95.3	0.00±0.00
Read 3	1018 ± 20	94 ± 1.2	19.74	18.55	92.7	0.29±0.14

[a] The number of bases sequenced, in gigabases

[b] The percentage of the sample that aligned to the PhiX genome

[c] The percentage of bases in reads that align perfectly, as determined by a spike in PhiX control sample, at the cycle indicated in brackets

[d] The percentage of bases in reads that align with 3 errors or less, as determined by a spike in PhiX control sample, at the cycle indicated in brackets

[e] The calculated error rate, as determined by the PhiX alignment

[f] The average of the four intensities measured at the first cycle averaged over filtered clusters

[g] The corresponding intensity statistic at cycle 20 as a percentage of that at the first cycle

[h] The percentage of bases with a quality score of 30 or higher, respectively, after the 25th cycle

[i] First sequence read

[k] Index read

[l] Second sequence read

[m] The density of clusters (in thousands per mm^2) detected by image analysis

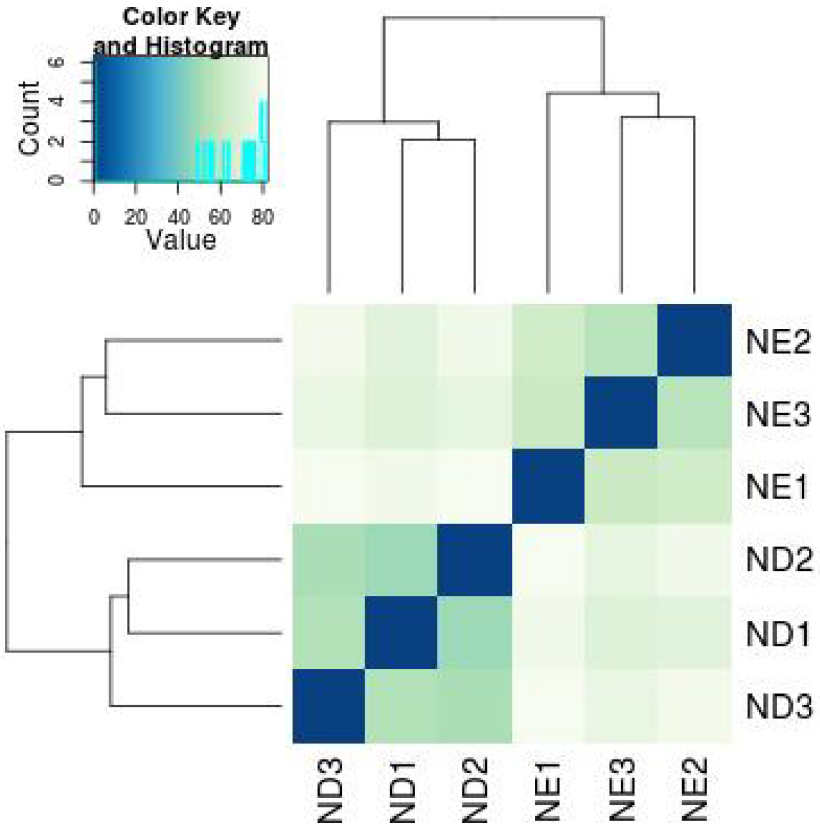
[n] The percentage of clusters passing filtering

[o] The number of clusters (in millions)

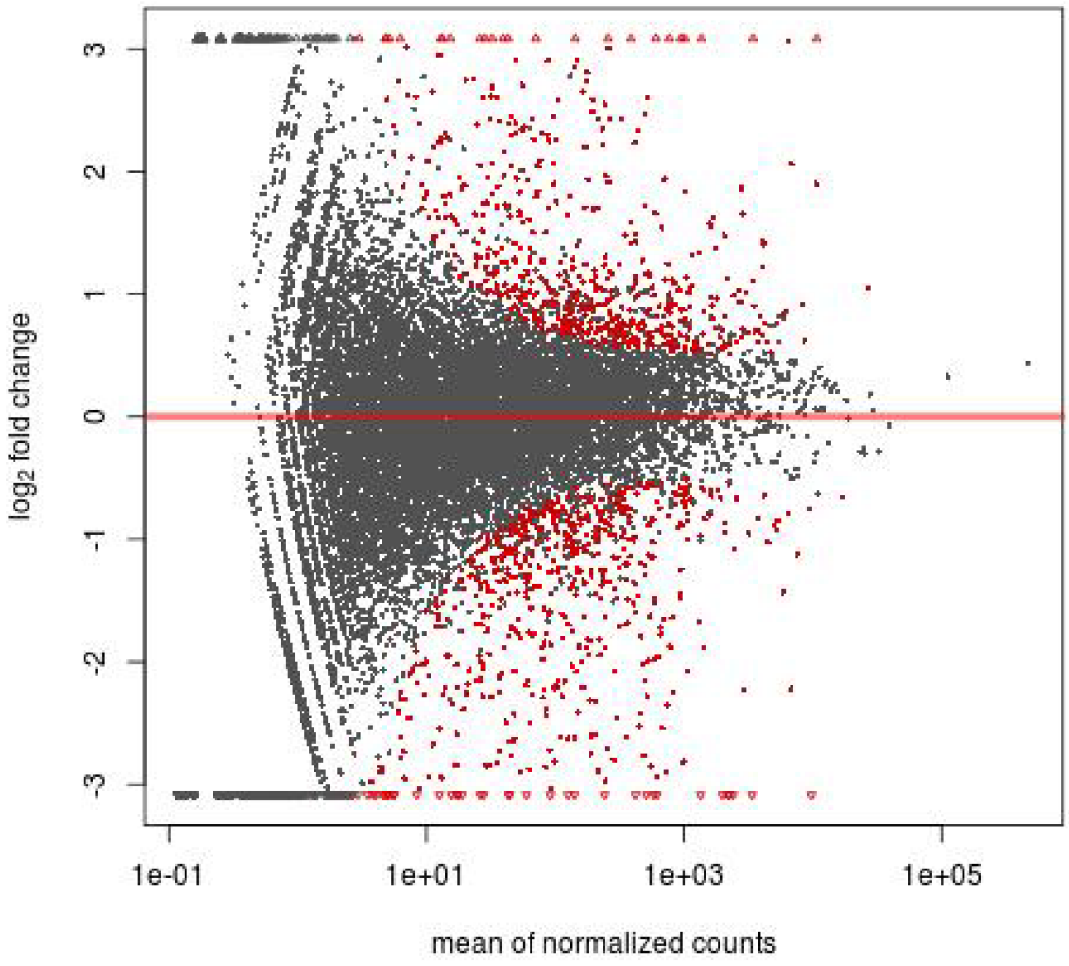
[p] The number of clusters (in millions) passing filtering

[q] The calculated error rate for cycles 1-100, as determined by the PhiX alignment

[A]


Color Key
and Histogram
Count
0
2
4
6
0
20
40
60
80
Value
NE2
NE3
NE1
ND2
ND1
ND3
ND3
ND1
ND2
NE1
NE3
NE2

[B]


log2 fold change
3
2
1
0
-1
-2
-3
1e-01
1e+01
1e+03
1e+05
mean of normalized counts

[C]

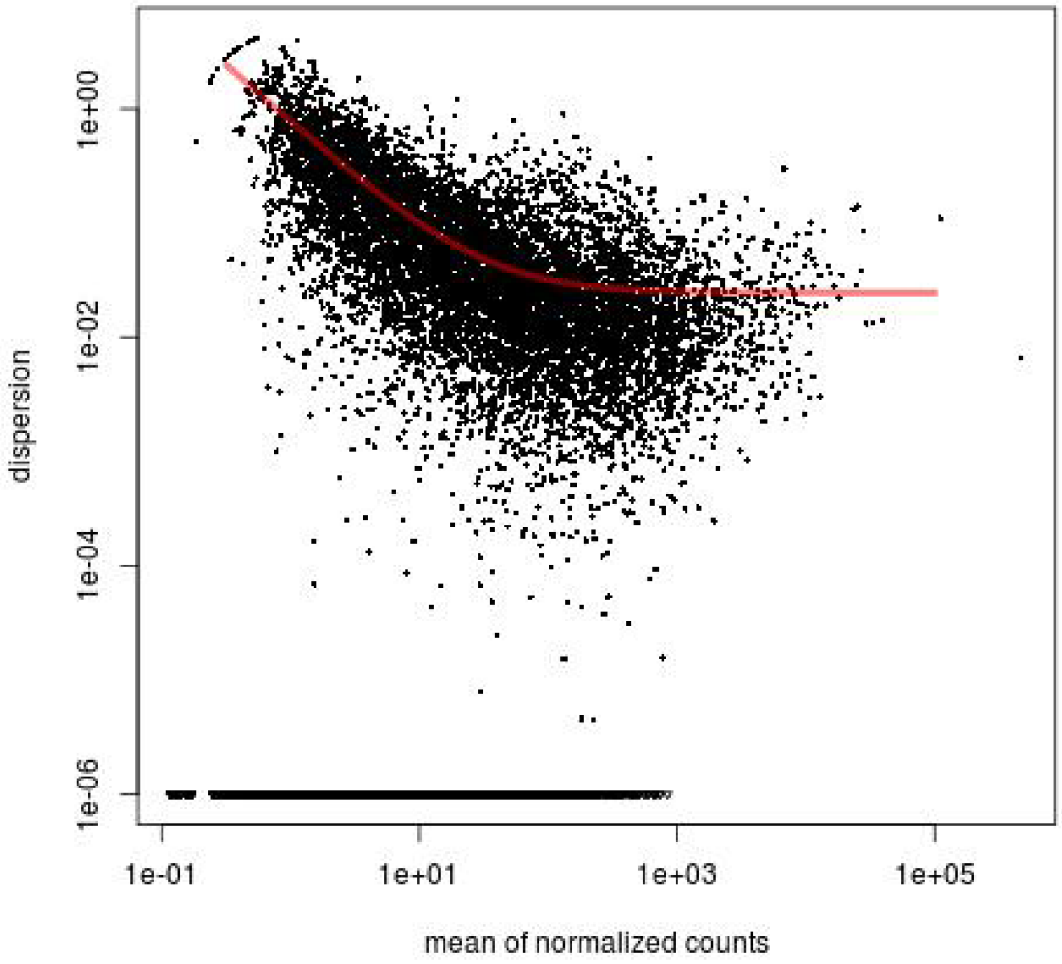


Figure 2.7. Data quality assessment by sample clustering and visualization. [A] Heatmap showing the Euclidean distances between the samples as calculated from the variance stabilising transformation of the count data. **[B]** Plot of normalized mean versus log2 fold change for the contrast nitrogen excess versus nitrogen depletion. **[C]** Empirical (black dots) and fitted (red lines) dispersion values plotted against the mean of the normalized counts.

Chapter 3

Defining the AmtR regulon of *Mycobacterium smegmatis* and its role in urea metabolism

Michael Petridis *, Chelsea Vickers *, Jennifer Robson, Joanna L. McKenzie, Abigail Sharrock, Htin Lin Aung, Vickery L. Arcus † and Gregory M. Cook †.

Article In Preparation: Applied and Environmental Microbiology

Author contributions: Authors are listed in order of magnitude of their contribution in each role. First authors are indicated by an asterisk (*) and corresponding authors are indicated by a cross (†). MP and GMC planned the research. MP performed the research with assistance from JR and HLA. CV solved the crystal structure of AmtR and performed ligand modeling. MP and GMC interpreted data and wrote the manuscript. MP produced Figures 3.1, 3.4, 3.5, 3.6, 3.7, 3.8 and Table 3.1. JR produced Figure 3.2 and 3.3. AS contributed to Figure 3.7.

3.1. Abstract

Some members of the soil-dwelling actinobacteria harbor two distinct global nitrogen regulatory proteins, GlnR and AmtR. The majority of these organisms (e.g. *S. coelicolor* and *C. glutamicum*) encode either for GlnR or AmtR, whereas other organisms (e.g. *M. smegmatis*) encode homologs for both transcriptional regulators in their genome. In *M. smegmatis* GlnR mediates the expression of key genes involved in nitrogen metabolism [43]. Recent studies reported an AmtR-binding site in an operon encoding an urea carboxylase (MSMEG_2184-MSMEG_2189) in *M. smegmatis* [38], however, the function of AmtR in *M. smegmatis* remains unknown. To characterise the regulatory mechanism of AmtR in *M. smegmatis* we performed a combination of molecular, biochemical and structural analyses. We used genome-wide expression profiling comparing the wild-type strain and a Δ*amtR* mutant to define the AmtR regulon and we identified significant changes in the expression of 11 genes including genes involved in uptake, assimilation and metabolism of nitrogen. We showed that AmtR regulated a pathway of urea degradation that is found in the genomes of some soil-dwelling actinobacteria. Moreover, we provided direct evidence for the AmtR-consensus binding motif (CTGTC-N_4-GACAG) in *M. smegmatis*. Using differential scanning fluorimetry we identified xanthine and allantoin as ligands of mycobacterial AmtR that enhance protein stability and showed in *in vitro* DNA binding assays that addition of xanthine and allantoin had no destabilising effect on AmtR:DNA interaction. We conclude that ligand binding enhanced stability of an AmtR:DNA interaction blocking the expression of the operon *msmeg_2184-msmeg_2189* and we propose a novel regulatory mechanism of a TetR-like transcriptional repressor.

3.2. Introduction

Nitrogen is an essential component of almost all macromolecules (e.g. nucleic acids, proteins and cell wall components) and bacteria have developed elaborate mechanisms for the uptake, assimilation and regulation of nitrogen source utilisation. Actinomycetes inhabit diverse environments and some members of the soil-dwelling actinomycetes, e.g. *S. coelicolor*, *C. glutamicum* and *M. smegmatis*, harbor two distinct global nitrogen regulatory proteins, GlnR and AmtR [37,39,43]. GlnR is an OmpR-like transcriptional regulator, which has been characterised in *S. coelicolor* [37,41]. The global regulator of nitrogen metabolism in *C. glutamicum* is AmtR, which mediates the expression of at least 35 genes in response to changing nitrogen levels, including genes involved in ammonium uptake (*amtA*, *amtB*), ammonium assimilation (*glnA*, *gltBD*), urea metabolism (*ureABCEFGD*) and regulatory proteins (*glnK*, *glnD*) [27,39,40].

Some soil-dwelling actinomycetes (e.g. *M. smegmatis* and *S. avermitilis*) harbor homologs for both GlnR and AmtR in their genomes, whereas members of the *M. tuberculosis* complex only encode for GlnR [11,23]. In *M. smegmatis*, GlnR was shown to regulate the expression of more than a hundred genes in response to changing nitrogen levels, including multiple nitrogen uptake systems and ammonium assimilation mechanisms, whereas transcription of the urease-encoding operon is not under regulatory control of GlnR [43]. The AmtR regulon appears to control only a limited number of genes, inparticular the operon *msmeg_2184-msmeg_2189* that is encoding an urea amidolyase [11,38]. Binding of AmtR in the upstream region of *msmeg_2184* was shown using *in vitro* DNA binding assays suggesting a direct regulatory mechanism of AmtR [38]. The binding of AmtR was ligand-independent and no signal transduction pathway was proposed for regulation. Recent findings show that expression of this operon is also under control of GlnR that activates transcription of this operon under nitrogen limitation, but the physiological explanation for this dual control remains unknown [43].

AmtR belongs to the TetR protein family, which is abundant in bacteria that are exposed to environmental changes, such as those found in soil. The TetR family are DNA binding proteins and function generally as transcriptional repressor proteins. This family of proteins are involved in regulating cellular processes such as antibiotic resistance, efflux pumps, osmotic stress and pathogenicity [128]. Members of the TetR-family usually bind small effector molecules, however, in *C. glutamicum*, AmtR interacts with the P_{II} protein, while the molecular mechanism of AmtR in *M. smegmatis* has not been defined [39].

The aim of the current study was to identify the AmtR regulon and the molecular mechanism controlling gene expression in *M. smegmatis*. To address these aims, we performed a combination of molecular, biochemical and structural analyses to understand the function of AmtR in *M. smegmatis*.

3.3. Results

3.3.1. Molecular analysis of mycobacterial AmtR

The AmtR protein encoded by *msmeg_*4300 is 225 amino acids in length and belongs to the TetR/AcrR family of transcriptional regulators (Figure 3.1). Bioinformatic analyses indicated that MSMEG_4300 shared the highest amino acid sequence identity with other *Mycobacterium* species such as *Mycobacterium mageritense* (81%) (Uniprot X5L5Y0) and *Mycobacterium vulneris* (79%) (Uniprot X5LSM6), but shared only 41% sequence identity with the well-studied AmtR from *C. glutamicu*m (Uniprot Q79VH8) (Figure 3.1) [27,39]. The highest sequence similarity (~65%) was observed in the N-terminal DNA-binding domain, while the sequence similarity in the C-terminal ligand-binding domain was only around 20%. This suggests a different ligand for AmtR in *M. smegmatis* and not the P_{II} protein as in *C. glutamicum* [39]. Previous studies identified several amino acids in the AmtR protein of *C. glutamicum* that are important for DNA binding, however, some of these are not conserved in mycobacterial AmtR (H43R, I51V, Y58H) [127]. The closest homologue in *M. tuberculosis* is *Rv3160c* that has very weak identity (27.9%) to the *C. glutamicum* AmtR protein. A BLAST search using *C. glutamicum* AmtR and *M. smegmatis* AmtR against the *M. tuberculosis* complex showed no hits with an identity over 33%, suggesting no AmtR-like protein in the *M. tuberculosis* complex [23].

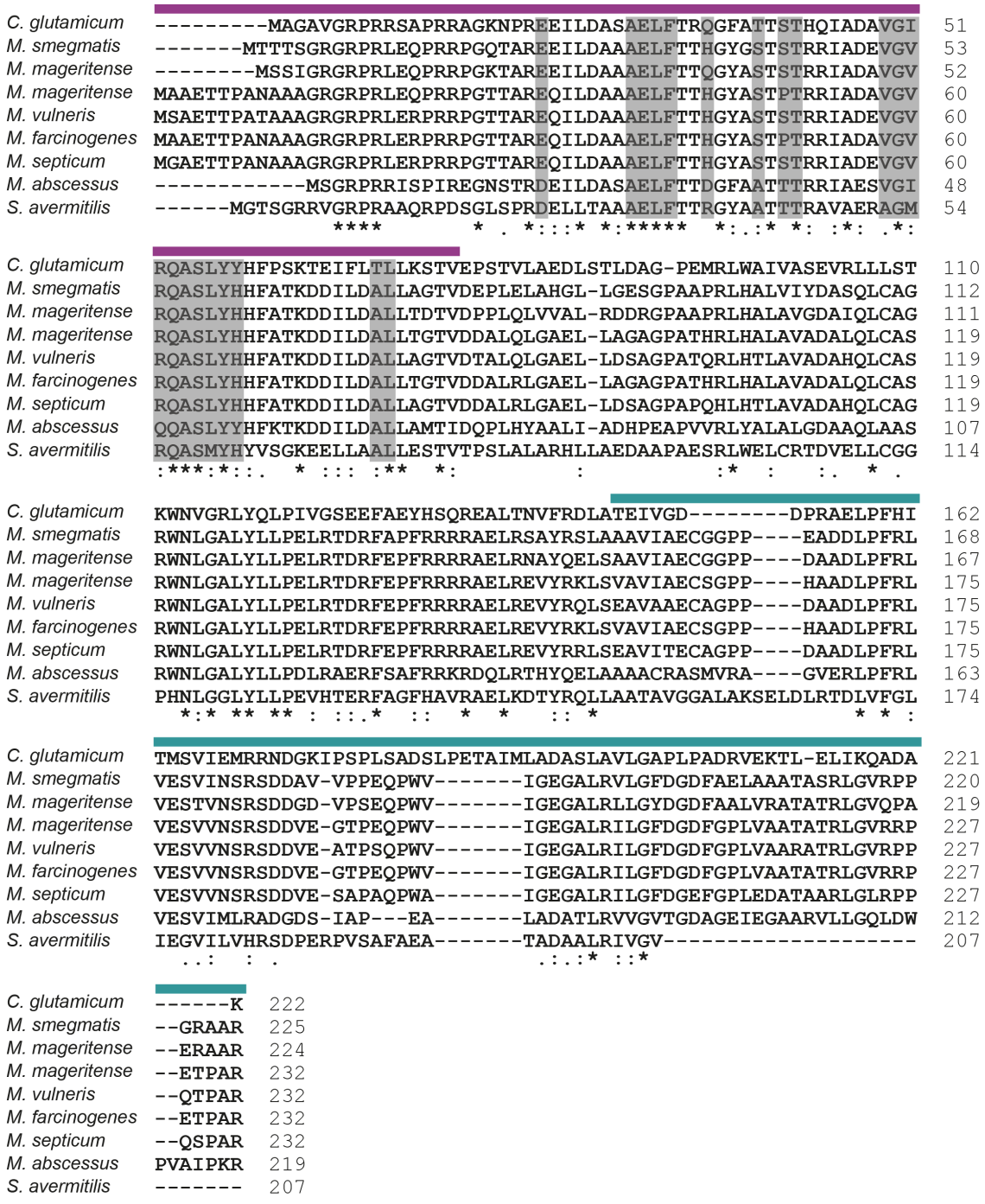


Figure 3.1. Alignment of mycobacterial AmtR proteins with AmtR of *C. glutamicum* and *S. avermitilis*. Residues of AmtR of *C. glutamicum* that are important for DNA binding and protein folding are highlighted in grey. The N-terminal DNA-binding domain is highlighted with a purple bar and the C-terminal ligand-binding domain is highlighted with a cyan bar. (*): identical amino acids; (:) or (.): similar amino acids. *C. glutamicum*: *Corynebacterium glutamicum*; *M. smegmatis*: *Mycobacterium smegmatis*; *M. mageritense*: *Mycobacterium mageritense*; *M. vulneris*: *Mycobacterium vulneris*; *M. farcinogenes*: *Mycobacterium farcinogenes*; *M. septicum*: *Mycobacterium septicum*; *M. abscessus*: *Mycobacterium abscessus*; *S. avermitilis*: *Streptomyces avermitilis*.

In *M. smegmatis*, the *amtR* gene is flanked by genes that encode for an acyl-CoA synthase (*msmeg_4301*) and an enoyl-CoA hydratase/isomerase (*msmeg_4299*), respectively, with both involved in fatty acid and phospholipid metabolism. To investigate the role of AmtR in *M. smegmatis*, we first confirmed whether the *amtR* gene is part of an operon with the flanking genes, which would indicate a putative role in fatty acid/phospholipid metabolism. To determine whether the *amtR* gene is transcribed as a single mRNA or co-transcribed as part of a larger transcript, we performed reverse transcriptase PCR (RT-PCR) with primers that bound within the *amtR* gene alone (Figure 3.2A, PCR product II) and on either side of the gene of interest (Figure 3.2A, PCR product I and product III). We were able to confirm that *amtR* is a single transcript. Primers used to amplify *amtR* resulted in a 480 bp RT-PCR product (Figure 3.2B, lane 2), which was the same size as the control PCR using genomic DNA as template for this primer pair (Figure 3.2B, lane 8). This was further confirmed by sequencing.

3.3.2. Growth of a Δ*amtR* mutant is uncoupled from glycerol consumption

To determine whether *M. smegmatis* AmtR played a role in the regulation of nitrogen metabolism, we created a markerless *amtR* deletion mutant (*M. smegmatis* JR258 Δ*amtR*) and confirmed deletion of 65% of the internal region of *amtR* with PCR (data not shown). We also performed a Southern blot (Figure 3.3), however, this result was inconclusive, as we observed multiple signals that could be the result of a partial digest or unspecific probe binding. The majority of these signals were observed in the digest of both wild-type and Δ*amtR* mutant genomic DNA, while signals corresponding to the approximate fragment size of 1.45 kb (wild-type) and 4.2 kb (Δ*amtR* mutant) were observed only in the respective digests of *M. smegmatis* mc^2155 wild-type and *M. smegmatis* JR258 Δ*amtR* genomic DNA.

To identify a phenotype of *M. smegmatis* JR258 Δ*amtR*, we compared growth between the wild-type strain mc^2155 and strain JR258 Δ*amtR* in HdB minimal medium with different nitrogen sources. While growth of both the wild-type strain mc^2155 and strain JR258 Δ*amtR* was comparable for most of the nitrogen sources (e.g. $(NH_4)_2SO_4$) tested (Figure 3.4A), we only observed a growth phenotype when lysine was added as the sole nitrogen source (Figure 3.4B). *M. smegmatis* JR258 Δ*amtR* revealed a slower doubling time compared to the wild-type strain mc^2155 (wild-type 18.6h ± 1.4 h, Δ*amtR* 28.4h ± 2.7 h) and we verified this phenotype with counts of colony-forming units per mL (CFU/mL) (Figure 3.4C).

To validate the phenotype of the *amtR* deletion, complementation of strain JR258 Δ*amtR* with a tetracycline inducible vector carrying the *amtR* gene (pMind-*amtR*) of *M. smegmatis* was performed. Strain JR258 Δ*amtR* was transformed with either pMind-*amtR* or with the empty plasmid pMind as a control. While the empty plasmid pMind was without effect on microbial growth, pMind-*amtR* restored the growth deficiency of strain JR258 Δ*amtR* back to the wild-type level (Figure 3.4D).

To determine the metabolic basis for the lower growth rate in the strain JR258 Δ*amtR*, we measured the molar growth yield and the dry weights, when cells were grown with glycerol as the carbon source (Figure 3.4E). This analysis revealed that the strain JR258 Δ*amtR* utilised more glycerol to generate the same weight of bacterial biomass compared to the wild-type strain mc^2155. The dry weight for the wild-type (0.93 ± 0.02 g) was the same as for JR258 Δ*amtR* (0.90 ± 0.02 g) (Figure 3.4F), while the glycerol yield ($Y_{glycerol}$) for strain JR258 Δ*amtR* was 35.35 ± 0.46 g [dry weight] cells/mol glycerol utilised and for the wild-type strain 38.17 ± 0.83 g [dry weight] cells/mol glycerol (Figure 3.4.E), indicating that glycerol consumption was uncoupled from biomass generation in *M. smegmatis* JR258 Δ*amtR*.

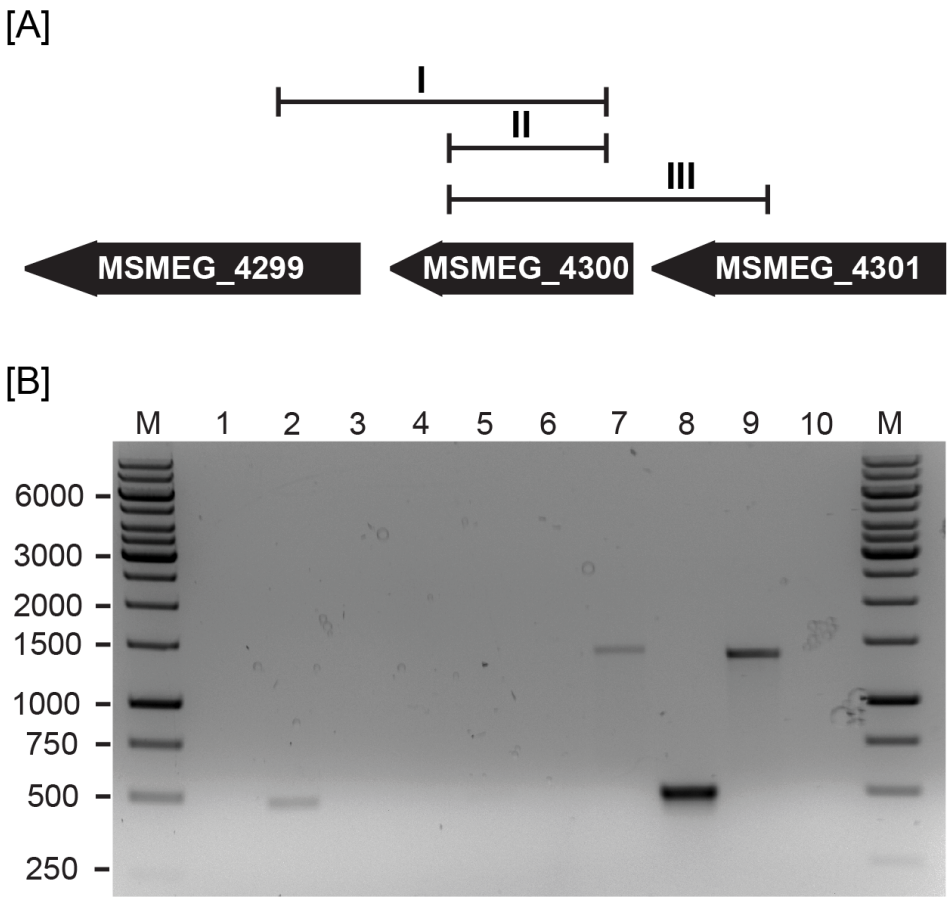
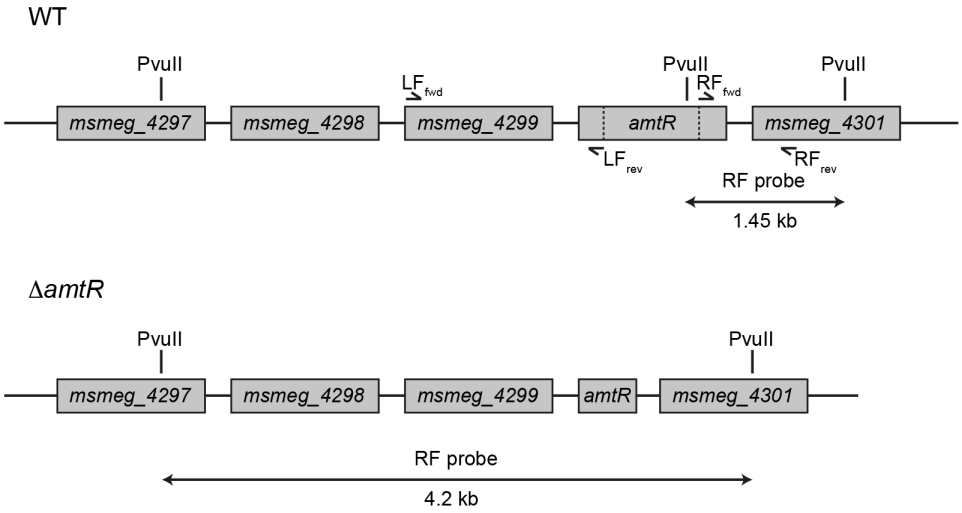


Figure 3.2. The genetic organisation of the *amtR* gene in *M. smegmatis*. **[A]** Schematic overview of the *amtR* genomic region indicated by black arrows, which show the genetic organisation of the *amtR* gene (*msmeg_4300*) and flanking open-reading frames. The downstream gene *msmeg_4299* is annotated as enoyl-CoA hydratase/isomerase and is situated 11 bp downstream of the *amtR* gene stop codon. The upstream gene *msmeg_4301*, annotated as acyl-CoA synthase, terminates 7 bp from the start codon of *msmeg_4300*. Amplified regions are indicated by lines above the open-reading frames numbered I, II and III, which correlate with gene-specific primer pairs (I) AmtR LF F and AmtR inner R, (II) AmtR inner F and R and (III) AmtR inner F and AmtR RF R, respectively. **[B]** Electrophoresed bands from RT-PCR reaction show that no product was present using primer pairs I and III (lanes 1 and 3). A product of expected size (480 base pairs) is shown in lane 2, which correlates to internal primer pairs II for the *amtR* transcript only. No RT-PCR-step controls to check for genomic DNA contamination are shown for primer pairs I, II and III in lanes 4, 5 and 6, respectively. PCR-positive controls using genomic DNA show the expected size fragments for primer pairs I, II and III as 1377 bp (lane 7), 480 bp (lane 8) and 1342 bp (lane 9), respectively. Negative control using water as a template is shown in lane 10. Molecular weight standards with sizes indicated in bases are shown on left. The data for this figure was produced by JR.

[A]



[B]

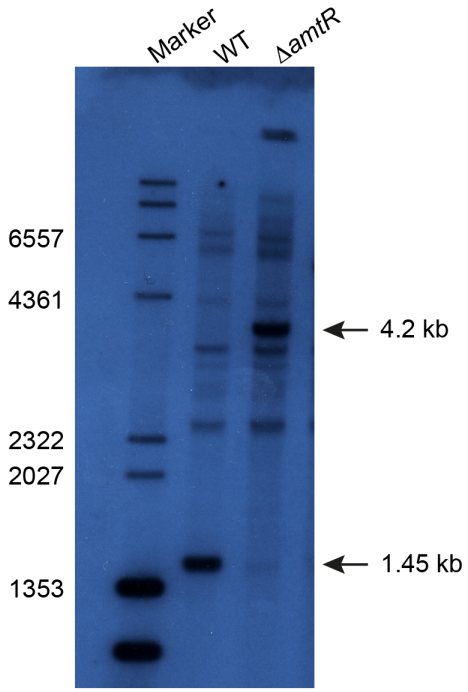


Figure 3.3. Construction of a markerless Δ*amtR* mutant of *M. smegmatis* mc^{2}155. [A] The knockout construct consists of two fragments (left flank spanned by LF_{fwd} and LF_{rev}; right flank spanned by RF_{fwd} and RF_{rev}) on each side of *amtR* in the plasmid pX33. Restriction sites of PvuII and fragment sizes as detected in a Southern hybridisation are indicated. **[B]** PvuII-digested genomic DNA of WT and Δ*amtR* were probed with labelled right flank PCR product of the deletion construct. WT, *M. smegmatis* mc^{2}155 wild-type, Δ*amtR*, *M. smegmatis* JR258 Δ*amtR*. Molecular weight standards with sizes indicated in bases are shown on left. The data for this figure was produced by JR.

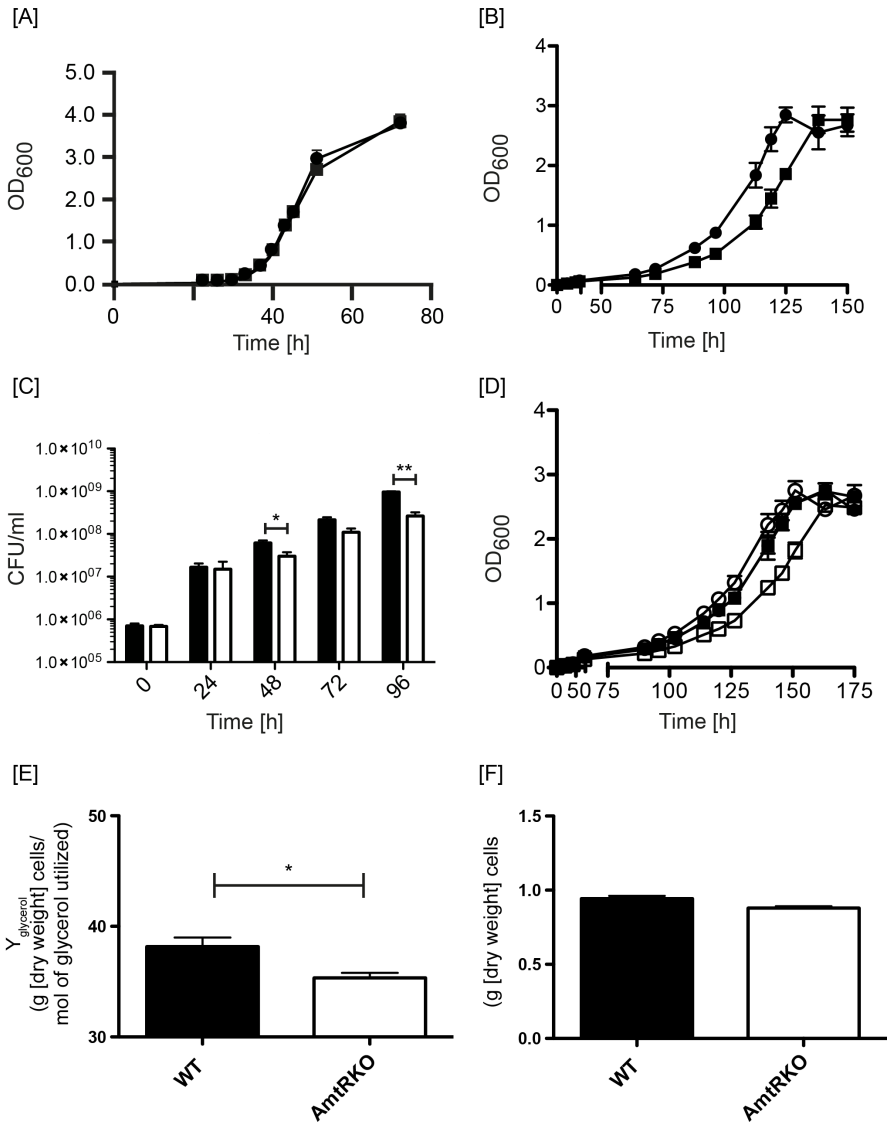


Figure 3.4. Growth comparison of strains mc²155 wild-type and JR258 Δ*amtR*. Comparative growth analyses by OD_{600} measurements were performed by growing the strains in HdB minimal medium containing 0.2% (w/v) glycerol as the carbon source and 1.5 mM $(NH_4)_2SO_4$ **[A]** or 1 mM L-lysine **[B]** as sole nitrogen source (wild-type, closed circles; JR258, closed squares). **[C]** CFU counts with 1 mM L-lysine as sole nitrogen source comparing colony forming units of wild-type strain mc²155 (black bars) and strain JR258 Δ*amtR* (white bars). Error bars represent the standard error of the mean of three independent biological replicates. **[D]** Complementation of the wild-type (closed circle) and JR258 Δ*amtR* (closed square) with a tetracycline-inducible pMind-*amtR* construct. Empty vector controls (pMind) were used as negative controls in the background of both the wild-type (open circle) and JR258 (open square). Complementation was performed in HdB minimal medium containing 0.2% (w/v) glycerol as carbon source and 1 mM L-lysine as sole nitrogen source. Error bars represent the standard error of the mean of three independent biological replicates. Determination of the molar growth yield ($Y_{glycerol}$) **[E]** and dry weight **[F]** of strains mc²155 wild-type and JR258 Δ*amtR*, when cells were grown in HdB minimal medium containing 0.2% (w/v) glycerol as carbon source and 1 mM L-lysine as sole nitrogen source.

3.3.3. AmtR regulates an urea degradation pathway

To gain further insights into the role of AmtR as a regulator of nitrogen metabolism, the transcriptional response of strain JR258 Δ*amtR* was compared to the wild-type mc^{2}155 using microarray analysis. Cells of both strains were grown in HdB minimal medium containing 0.2% (w/v) glycerol as carbon source and 1 mM L-lysine as sole nitrogen source and cells were harvested at an OD_{600} between 0.15 and 0.18 to exclude a general stress response linked to the growth deficit.

Comparison of the transcriptional response of strain JR258 Δ*amtR* with the wild-type mc^{2}155 revealed that a total of eleven genes were differentially expressed (>2.0x change, P<0.05, four independent biological replicates, including two dye swaps). Nine genes were upregulated and two genes were downregulated in strain JR258 Δ*amtR* (Table 3.1). The microarray data were validated by quantitative real-time PCR (qPCR) of eight selected genes (Figure 3.5). Surprisingly, the gene encoding for AmtR was not differentially expressed in the microarray analysis. Because only 65% of the *amtR* gene were deleted, the microarray might pick up truncated message, therefore we included *amtR* in the qPCR validation to validate expression of *amtR* in the wild-type strain (Figure 3.5).

Table 3.1. List of genes that are differentially expressed in strain JR258 Δ*amtR* compared to strain mc^{2}155 wild-type

mc^{2}155 locus[a]	Expression ratio[b]	*p* value[c]	Description
msmeg_2189	56.04	7.50E-02	allophanate hydrolase
msmeg_2186	42.72	3.28E-04	conserved hypothetical protein
msmeg_2184	39.79	1.43E-02	amino acid permease
msmeg_2187	29.86	1.67E-02	urea amidolyase
msmeg_2185	26.15	4.13E-03	conserved hypothetical protein
msmeg_4588	2.90	4.63E-02	ABC nitrate/sulfonate/bicarbonate family protein transporter, inner membrane subunit
msmeg_5594	2.27	3.50E-02	ferredoxin-dependent glutamate synthase
msmeg_4925	2.20	4.77E-02	transcriptional regulator
msmeg_6770	2.07	7.69E-02	conserved hypothetical protein
msmeg_6241	0.50	3.76E-03	ATPase associated with various cellular activites
msmeg_6240	0.46	1.18E-02	conserved hypothetical protein

[a] Locus number of gene in *M. smegmatis* mc^{2}155

[b] Mean gene expression ratio of four biological replicates

[c] *P*-value of gene expression ratio from four biological replicates determined using a *t*-test without false discovery correction

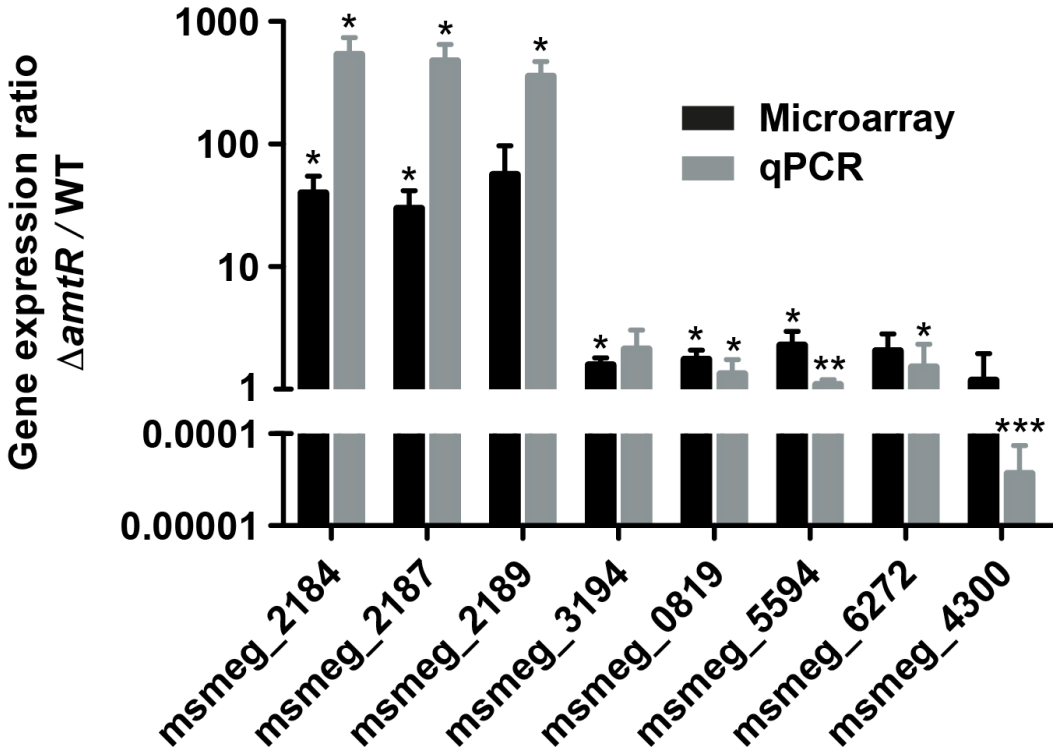


Figure 3.5. Validation of microarray analysis with qPCR. The microarray data were validated by comparing the gene expression changes obtained by microarray analysis with the gene expression changes obtained from quantitative real-time PCR (note the extended log scale on the y-axis). The bars represent mean values for the fold changes in gene expression from the microarray (black) and quantitative real-time PCR (grey). Error bars represent standard deviations from four (microarray) or three (quantitative real-time PCR) biological replicates. The gene *sigA* was used as an internal standard prior to the calculation of the fold change in mRNA levels. Statistically significant differences in gene expression are indicated by stars (* , *P* value <0.05; **, *P* value <0.01; ***, *P* value <0.001). msmeg_2184, amino acid permease; msmeg_2187, urea amidolyase; msmeg_2189, allophanate hydrolase; msmeg_3194, biotin synthase; msmeg_0819, N-carbamoyl-L-amino acid amidohydrolase; msmeg_5594, ferredoxin-dependent glutamate synthase; msmeg_6272, NAD-glutamate dehydrogenase; msmeg_4300, AmtR.

Five of the most highly upregulated genes, *msmeg_2184*, *msmeg_2185*, *msmeg_2186*, *msmeg_2187* and *msmeg_2189* are co-transcribed within a putative operon. The genes *msmeg_2187* and *msmeg_2189* (atzF) are upregulated 30-fold and 56-fold and are annotated as urea amidolyase and allophanate hydrolase, respectively, while *msmeg_2185* and *msmeg_2186* encode for hypothetical proteins that are upregulated 26-fold and 42-fold, respectively, in strain JR258 Δ*amtR*. A BLAST search against the Uniprot database revealed an identity between 70-80% of these two putative proteins to an urea carboxylase-associated protein 1 (*msmeg_2186*) and urea carboxylase-associated protein 2 (*msmeg_2185*) of *Gordonia bronchialis* and some other *Gordonia* species. Interestingly, genes encoding for an urea amidolyase and an allophanate hydrolase are, like AmtR, only present in some soil-dwelling actinobacteria, e.g. *M. smegmatis*, *M. vanbaalenii*, *M. abscessus* and *S. avermitilis*.

Urea amidolyase (UA) consists of two distinct enzyme activities; urea carboxylase (UC) and allophanate hydrolase (AH). This urea carboxylase/allophanate hydrolase (UC/AH) enzyme complex provides a biotin- and energy-dependent pathway of urea degradation, catalysing a two-step reaction to recover ammonium via the intermediate allophanate. Along with the aforementioned genes, we also observed 1.7-fold upregulation of genes encoding for a biotin synthase and a biotin ligase in strain JR258 Δ*amtR*. The urea carboxylase harbors several domains, a biotin carboxylase (BC), a carboxyltransferase (CT) and a biotin carboxyl carrier protein (BCCP). In an ATP-dependent step, the BC domain carboxylates biotin that is covalently linked to the BCCP domain and the CT domain subsequently transfers this carboxyl group from biotin to urea, producing allophanate [179]. However, *M. smegmatis* encodes for a Ni^{2+}-dependent urease enzyme that catalyses a very similar, but energy-independent reaction, converting urea into ammonium [11].

The gene *msmeg_2184* is annotated as amino acid permease and is upregulated 40-fold in strain JR258 Δ*amtR*, however, its specificity is unknown.

3.3.4. Deletion of AmtR affects expression of an NADH-dependent glutamate dehydrogenase

M. smegmatis harbors three copies of GDH enzymes, a class I NADPH-dependent GDH (*msmeg_5442*), a class II NADH-dependent GDH (*msmeg_6272*) and a class III NADH-dependent GDH (*msmeg_4699*). Furthermore, the genome of *M. smegmatis* also harbors multiple homologs of glutamine synthetase and glutamate synthase, but except for *glnA1* and *glnA2*, none of the other genes are under transcriptional control of GlnR [11,43]. In our analysis, we identified upregulation of the class II NADH-dependent GDH (*msmeg_6272*), as well as *msmeg_5594*, a gene encoding for the small subunit of a ferredoxin-dependent glutamate synthase, however, differential expression of both genes was below 2-fold (Figure 3.5). While upregulation of the glutamate synthase could not be verified by qPCR, expression of *msmeg_6272* encoding the class II NADH-dependent GDH was significantly affected in strain JR258 Δ*amtR* (1.5-fold, $p<0.05$), although the microarray data initially suggested no significant effect (2.1-fold, $p=0.068$) (Figure 3.5). Moreover, four other genes were significantly upregulated in strain JR258 Δ*amtR* including genes encoding for a transporter associated protein, a transcriptional regulator and a conserved hypothetical protein.

In total, two genes (*msmeg_6240* and *msmeg_6241*) were downregulated more than 2-fold, that form an operon with *msmeg_6242*, which was downregulated 1.64-fold (Table 3.1). This operon encodes for an alcohol dehydrogenase (*msmeg_6242*), an AAA ATPase (*msmeg_6241*) and a hypothetical protein containing a von Willebrandt factor type A domain (*msmeg_6240*), but the function of these genes in *M. smegmatis* has not been defined.

3.3.5. Xanthine and allantoin are ligands of mycobacterial AmtR

To gain insights into the molecular mechanism of AmtR regulation and to identify putative ligands of AmtR, we used differential scanning fluorimetry (DSF) and combined this with high-throughput screening using the Biolog Plate PM3 from Biolog®. This allowed us to screen 95 different nitrogen sources for potential interactions with AmtR. After optimisation of the DSF assay, we wanted to determine the effect of DNA (promoter region of *msmeg_2184*) on the thermal stability of the AmtR protein. While the melting temperature (T_m) of AmtR in the absence of DNA was around 56°C, we observed a shift of the T_m in presence of DNA to 60.6°C (Figure 3.6A, + DNA = solid grey; - DNA = dashed grey). Next, we included the different compounds in the DSF assay at a final concentration of 2 mM to determine potential effects on the thermal stability of the AmtR protein, which would indicate an interaction of the ligand with AmtR. Addition of the majority of the 95 compounds did not result in a shift of the T_m either in the presence or absence of cognate DNA (the T_m of AmtR or AmtR:DNA remained within a range of 56°C to 60.5°C; Figure 3.6A, + nitrate = solid yellow; + urea = solid green; + glutamate = solid blue). However, we observed a shift in T_m towards increasing temperatures after addition of the two compounds allantoin and xanthine, indicating an increase in the T_m of 64.5°C and 66.5°C, respectively, but only in the presence of DNA (Figure 3.6B, + DNA + xanthine = solid purple; + DNA + allantoin = solid orange), suggesting an interaction of these compounds with mycobacterial AmtR. Subsequently, both xanthine and allantoin were successfully modelled into the ligand binding pocket of a 1.98 Å crystal structure of mycobacterial AmtR[¶]. Xanthine showed putative interactions with several amino acid residues in the ligand binding pocket, suggesting that the ligand cavity is compatible with this ligand.

[¶] Crystal structure and ligand modeling were provided by Chelsea Vickers and Vickery L. Arcus

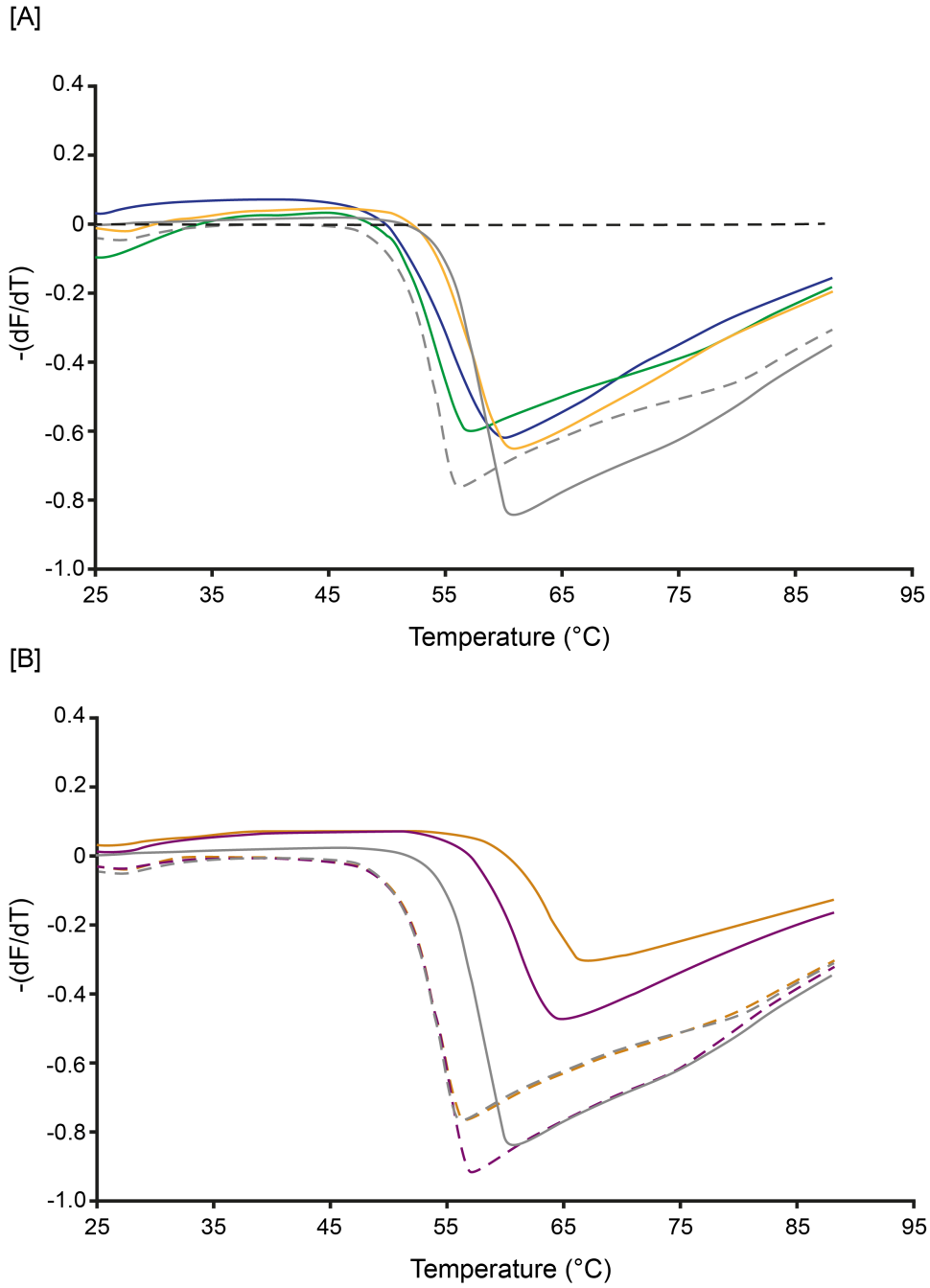


Figure 3.6. Thermal stability of the AmtR protein. [A] Increase in melting temperature (T_m) of the AmtR protein in presence of DNA (solid grey: AmtR + DNA; dashed grey: AmtR – DNA). The effect of various exemplary compounds is shown (solid yellow: AmtR + DNA + nitrate; solid green: AmtR + DNA + urea; solid blue: AmtR + DNA + glutamate). DNA only (dashed black) has no effect. **[B]** Increase in the T_m of the AmtR protein after addition of xanthine and allantoin in presence, but not in absence of DNA (solid grey: AmtR + DNA; dashed grey: AmtR – DNA; solid purple: AmtR + DNA + xanthine; dashed purple: AmtR – DNA + xanthine; solid orange: AmtR + DNA + allantoin; dashed orange: AmtR – DNA + allantoin). The assay was performed in two independent biological replicates and one representative dataset is presented.

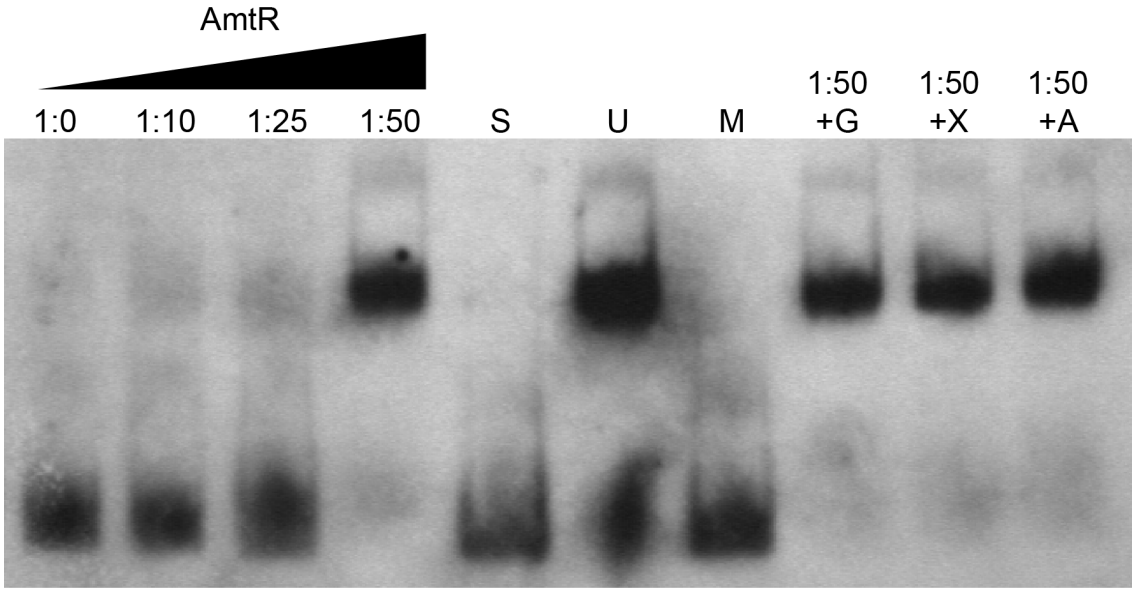
3.3.6. Identification of the AmtR-binding motif in *M. smegmatis*

Previous work has verified an AmtR-binding site in the promoter region of *msmeg_2184* and suggested an inverted repeat (CTGTC-N_4-GACAG) as the AmtR-binding site [38]. To validate this, a 186-bp DIG-labelled PCR product including the promoter region of *msmeg_2184* (P_{2184}) was amplified and purified AmtR protein[¶] from *M. smegmatis* was used for electrophoretic mobility shift assays (EMSA) (Figure 3.7A). The AmtR protein caused a mobility shift of P_{2184} at a molar DNA:protein ratio of 1:25, with a complete shift occurring at a ratio of 1:50 (Figure 3.7, lane 4). To determine specificity of binding to P_{2184}, unlabelled P_{2184} DNA was added in 120x excess (Figure 3.7A, lane 5), resulting in complete abolition of the shift. In lane 6, unlabelled non-specific DNA (promoter region of *phoH2* gene) was added in 103-fold excess, not affecting the mobility shift of P_{2184} (Figure 3.7A, lane 6). This result clearly demonstrated that the DNA retardation was specific for the promoter region of *msmeg_2184*. To verify AmtR-binding at the inverted repeat, we mutated this putative AmtR-binding motif using site-directed mutagenesis (Figure 3.7B). DNA retardation was completely abolished with the mutated promoter region of *msmeg_2184* (Figure 3.7A, lane 7), clearly showing that these nucleotides are essential for DNA-binding of AmtR in mycobacteria. To determine the effect of allantoin, xanthine and guanidine on DNA-binding of AmtR, these ligands were added at a concentration of 10 mM to the EMSA, however, we could not obtain release of AmtR from the promoter region of *msmeg_2184* (Figure 3.7A, lanes 8-10).

Using the Motif Alignment & Search Tool (MAST) of the MEME software, we looked at conservation of an AmtR consensus binding motif in the promoter region of genes in the genome of *M. smegmatis*. We used the confirmed motif in the promoter region of *msmeg_2184* as input, to screen upstream regions (250 bp upstream of +1) of all annotated *M. smegmatis* genes. An AmtR-binding motif was only present upstream of *msmeg_2184*, *msmeg_3966* and *msmeg_0426*, but only expression of *msmeg_2184* and *msmeg_3966* were affected in our microarray analysis. These results indicate that AmtR has only a minor direct regulatory effect in *M. smegmatis*.

[¶] purified AmtR Protein was provided by Chelsea Vickers and Abigail Sharrock

[A]



[B]

Unmutated Sequence: GCGCCCATTGTTATCTGTCATCTGACAGGT*AAGAGCTGGC
Mutated Sequence: GCGCCCATTGTTATAAGAAATCTAAAGGT*AAGAGCTGGC

Figure 3.7. *In vitro* DNA binding of purified AmtR protein. **[A]** Electrophoretic Mobility Shift Assay (EMSA) with AmtR protein and a DIG-labelled 186-bp DNA fragment including the putative AmtR-binding site in P_{msmeg_2184}. The ratios of protein to DNA are shown above. Controls for binding specificity (lane S) using 103-fold excess of unlabelled P_{msmeg_2184} (lane S) and using 103-fold excess of unlabelled promoter region of the unspecific *phoH2* gene (lane U) are shown. EMSA using a mutated promoter region of *msmeg_2184* [B] is shown in lane M. Effect of ligands (guanidine, +G; xanthine, +X; allantoin, +A), added at concentrations of 10 mM, on DNA binding of purified AmtR protein were determined. **[B]** Unmutated AmtR-binding region (underlined) in P_{msmeg_2184}. Nucleotides in the inverted repeats (marked) were mutated (shown in red) using site-directed mutagenesis. AS contributed to this figure.

3.4. Discussion

In *C. glutamicum*, AmtR interacts directly with the P_{II} protein GlnK and mediates the expression of more than 35 genes in response to changing nitrogen levels [27,39,40]. An AmtR:GlnK interaction strongly depends on the nitrogen supply of the cells, which has an immediate effect on the state of GlnK (unmodified form or adenylylated form). AmtR interacts only with GlnK in its adenylylated form that occurs when the cells face nitrogen limitation [39]. However, to date no important amino acid residues for this interaction and no structural data to show the exact mechanistic interaction between GlnK and AmtR have been reported [39]. The mediator of the transcriptomic response to nitrogen limitation in *M. smegmatis* is GlnR, regulating the transcription of important enzymes in nitrogen uptake (e.g. ammonium transporter *amtB*), assimilation (e.g. glutamine synthetase *glnA*) and metabolism (e.g. urease complex *ureABCEFGD*), while the function of AmtR in *M. smegmatis* has not been defined [11,43,71]. In this chapter, we describe the function and the molecular mechanism of nitrogen sensing by AmtR in *M. smegmatis*.

In a genome-wide transcriptomic analysis, we were able to determine the AmtR regulon and used site-directed mutagenesis to define the AmtR-binding site. The regulatory effect of AmtR is restricted to eleven genes, including genes involved in uptake, assimilation and metabolism of nitrogen. We did not observe changes in the transcription of *amtR* comparing the wild-type strain mc^2155 and strain JR258 Δ*amtR*, suggesting that *amtR* is only expressed at low levels. This is supported by recent findings using strand-specific RNA-seq analysis to define the nitrogen regulated transcriptome of *M. smegmatis* (Chapter 2). Furthermore, a non-coding antisense RNA transcript was identified (Chapter 2), suggesting a post-transcriptional regulatory mechanism for AmtR in *M. smegmatis*.

AmtR regulates expression of an alternative urea degradation pathway next to the urease. Urea amidolyase has been mainly described in yeast and algae that lack the urease enzyme, suggesting that urea amidolyase replaces the function of urease in these organisms [180,181]. However, sequencing of

numerous bacterial genomes revealed the presence of an urea amidolyase in prokaryotes, where they co-occur with the urease [182,183]. The urea amidolyase enzyme drives the energy-dependent conversion of urea into ammonium with concomitant ATP consumption, while ammonium production via the urease is energy-independent [181,182,184]. Further biochemical data are required to elucidate the physiological function of the urea amidolyase enzyme in bacteria, organisms that usually utilise the urease enzyme for urea degradation. To date, only one prokaryotic urea amidolyase has been biochemically characterised, while structure and function of the eukaryotic enzyme have been previously reported [179,182,183]. Co-occurrence of these two urea degradation pathways in bacterial genomes may be based on different affinities for their common substrate, however this requires further investigation.

Previous analysis of the genome-wide transcriptomic response to nitrogen limitation revealed that the operon *msmeg_2184-msmeg_2189* is not expressed under nitrogen surplus, but only under nitrogen limitation, whereas basal transcript levels of the urease encoding genes were detected under both conditions (Chapter 2). Furthermore, a GlnR-binding site in the promoter region of *msmeg_2184* was previously identified, suggesting dual regulatory control of the genes encoding urea amidolyase in *M. smegmatis*, while expression of the genes encoding for urease are not under control of GlnR [43].

In strain JR258 Δ*amtR*, expression of a class II NADH-dependent glutamate dehydrogenase was upregulated. In contrast to *C. glutamicum* and *S. coelicolor*, which only encode for one NADPH-dependent GDH, *M. smegmatis* harbors three copies of GDH enzymes in its genome; a class I NADPH-dependent GDH (*msmeg_5442*), a class II NADH-dependent GDH (*msmeg_6272*) and a class III NADH-dependent GDH (*msmeg_4699*). The class I and III enzymes encode for functional GDH enzymes that have been biochemically characterised, while activity of a functional class II GDH has yet be shown [52,59]. The presence of multiple GDH enzymes can have several reasons. The affinity for a common substrate might differ, the genes encoding these enzymes might be transcribed only under specific conditions or they might fulfill different functions *in vivo*. Harper *et al.* reported that expression of

the class II and class III enzymes were elevated 2 hours after exposure to nitrogen starvation [52]. In contrast, transcription of the class I GDH gene *msmeg_5442* was downregulated within the first 30 minutes of nitrogen limitation, whereas its activity increased concurrently [52]. This implies an additional post-translational regulatory mechanism, as previously reported for the GDH enzymes of *C. glutamicum* and *M. tuberculosis* [51,52,59,62]. In *M. smegmatis*, none of the GDH-encoding genes are under control of the global nitrogen regulator GlnR, contrary to *S. coelicolor* and *C. glutamicum*, where *gdhA* is regulated by GlnR or AmtR, respectively [41,63].

We identified the AmtR-binding site in *M. smegmatis* and showed binding of AmtR at the inverted repeat (CTGTC-N_4-GACAG) in the promoter region of *msmeg_2184*, as previously suggested [38,130]. Site-directed mutagenesis of nucleotides within the inverted repeat abolished binding of AmtR completely. Bioinformatic analysis revealed that the AmtR-binding site is absent in the promoter region of the majority of the AmtR regulon, suggesting that differential gene expression of these genes occurs due to indirect effects of an *amtR* deletion. An AmtR-binding site was only detected in the promoter regions of *msmeg_2184*, *msmeg_3966* and *msmeg_0426*, however, we only observed transcriptional changes for *msmeg_2184* and *msmeg_3966*.

To determine the molecular mechanism of AmtR regulation, we used differential scanning fluorimetry to identify potential ligands that cause either an increase or decrease in the thermal stability of AmtR, suggesting conformational changes within the protein in response to ligand binding. The majority of the compounds had little to no effect, however we noted that xanthine and allantoin promoted an increase in stability of an AmtR:DNA interaction. Hypoxanthine has been previously reported to regulate transcription of a xanthine-specific permease in *B. subtilis* by binding to a guanine-responsive riboswitch and to terminate transcription of the *xpt-pbuX* operon that is involved in xanthine metabolism. Furthermore, the transcriptional repressor PurR, a member of the Lac1 family, which regulates purine biosynthesis is dependent on a co-repressor, which can be either hypoxanthine or guanine [185,186]. Based on our results we propose a model in which the AmtR protein

functions as a transcriptional repressor of the urea carboxylase/allophanate hydrolase-encoding operon under nitrogen surplus by sensing intracellular xanthine and/or allantoin levels. Both xanthine and allantoin stabilize an AmtR:DNA interaction and therefore block transcription of the operon *msmeg_2184-msmeg_2189*. Xanthine and allantoin are both metabolites in a pathway of urea biosynthesis and transcription of this pathway is elevated in response to nitrogen limitation, implying a feedback regulatory mechanism that tightly controls urea degradation depending on the intracellular nitrogen levels (Chapter 2). When nitrogen levels drop, concentrations of xanthine and allantoin are likely to decrease, causing destabilisation of an AmtR:DNA interaction, resulting in the release of AmtR from the promoter region of *msmeg_2184*. Then, the transcriptional regulator GlnR binds at the *msmeg_2184* promoter region and activates transcription of the genes encoding UC/AH, an energy-dependent urea degradation pathway (Figure 3.8) [43]. By structurally and functionally characterizing the *M. smegmatis* AmtR transcription factor we have begun to elucidate a novel co-repressor mechanism of nutrient signaling that allows the AmtR regulon to provide an alternative way of nitrogen source acquisition in form of urea metabolism. This is of significance as *M. smegmatis* inhabits an environment of constant nitrogen limitation, so the ability to respond with alternative ammonium production in form of urea metabolism provides a strong adaptive advantage towards other soil-dwelling bacteria.

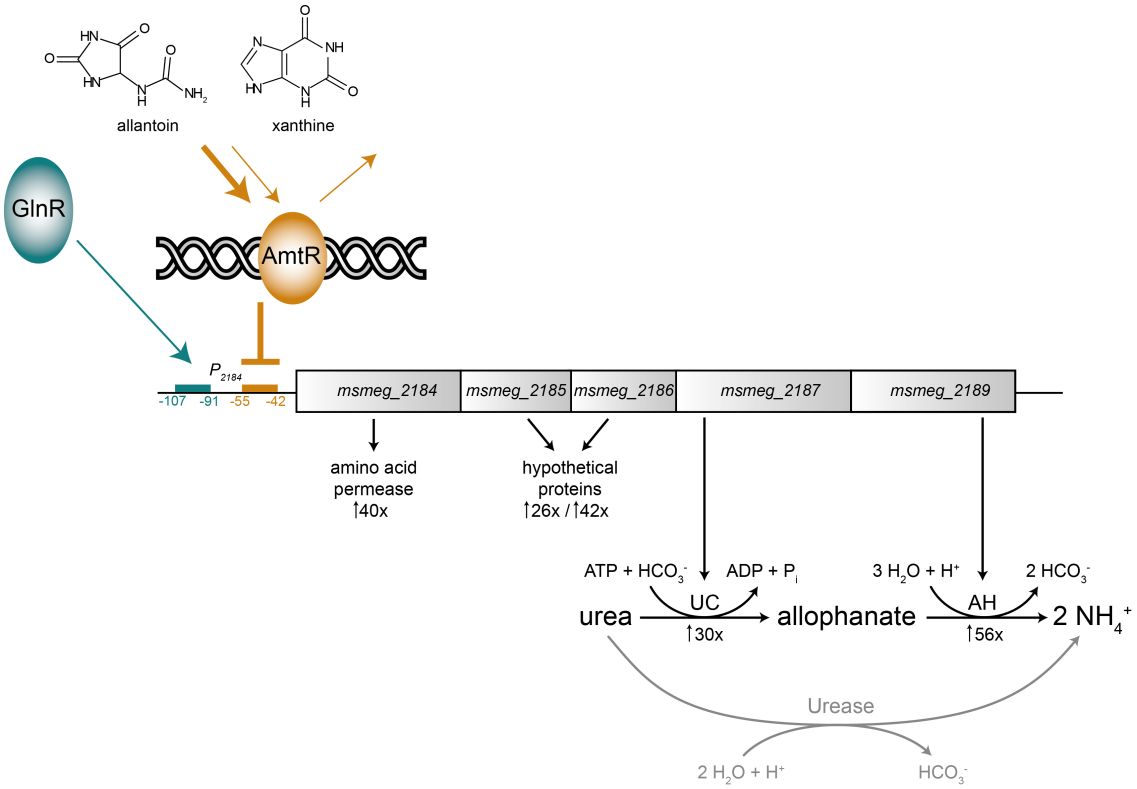


Figure 3.8. Model of function for the regulation of the urea carboxylase and allophanate hydrolase by AmtR and GlnR. The AmtR protein functions as transcriptional repressor of the UC/AH encoding operon under nitrogen surplus by sensing intracelullar xanthine and/or allantoin levels. During nitrogen excess, allantoin and xanthine levels are expected to be high (bold orange arrow), stabilising an AmtR:DNA interaction and therefore blocking transcription of the operon *msmeg_2184-msmeg_2189*. When nitrogen levels drop, the concentrations of xanthine and allantoin are likely to decrease (thin orange arrow), causing destabilisation of the AmtR protein and its release from the *msmeg_2184* promoter region (orange box). Then, the transcriptional regulator GlnR binds at the GlnR-binding region (cyan box) in the promoter region of *msmeg_2184* (thin cyan arrow) and activates transcription of the genes encoding UC/AH, an energy-dependent urea degradation pathway.

3.5. Materials and Methods

3.5.1. Bacterial strains, media and growth conditions

E. coli DH10B was grown in LB medium at 37°C with agitation (200 rpm) or on LB agar plates. *M. smegmatis* mc^2155 and derived strains were grown in LBT or HdB minimal medium supplemented with 0.2% (w/v) glycerol (unless otherwise stated) and 0.05% (w/v) Tween-80 (in LBT) or 0.05% (w/v) Tyloxapol (in HdB) at 37°C with agitation (200 rpm). Growth curves were performed in triplicate. Bacterial cell viability was monitored by cell counts based on CFUs per mL, where serial dilutions of bacterial cell culture in phosphate buffered saline with 0.05% Tween-80 (PBST) were plated on LBT agar plates. Selective media contained ampicillin (100 µg mL^{-1} for *E. coli*), kanamycin (50 µg mL^{-1} for *E. coli*, 20 µg mL^{-1} for *M. smegmatis*), hygromycin B (200 µg mL^{-1} for *E. coli*, 50 µg mL^{-1} for *M. smegmatis*) and gentamycin (25 µg mL^{-1} for *E. coli*, 5 µg mL^{-1} for *M. smegmatis*).

All bacterial strains and plasmids used are listed in Appendices 1 & 2.

3.5.2. Dry weight measurement and glycerol concentration assay

Cultures were grown in HdB with 0.2% (w/v) glycerol as the carbon source and 1 mM L-lysine as the sole nitrogen source until stationary phase and 20 mL of each culture was harvested by rapid filtration through a nitrocellulose filter (pore size 0.22 µm; Millipore) with an applied vacuum of approximately 7.5 psi. Filters were washed twice with 20 mL PBST, dried at 60°C until no change in weight could be determined and dry weights were measured. For glycerol concentration assays, samples were taken after inoculation to determine the starting glycerol concentration for each biological replicate and after entering stationary growth phase. Culture samples were centrifuged (5 min, 14.400 x *g*) to obtain cell-free supernatant. Glycerol concentrations were measured by detecting NADH oxidation (absorbance 340 nm) as previously described [169].

3.5.3. Alignment of AmtR proteins

Amino acid sequences of AmtR proteins from *C. glutamicum* (CAF19569.1), *S. avermitilis* (BAC74412.1), *M. smegmatis* (WP_011729715.1), Mycobacterium mageritense (WP_036438965.1 and WP_019346862.1), *Mycobacterium vulneris* (WP_036449407.1), *Mycobacterium farcinogenes* (WP_036393868.1), *Mycobacterium septicum* (WP_044523525.1) and *Mycobacterium abscessus* (WP_005071231.1) were obtained from the National Center for Biotechnology Information (NCBI). Sequences were aligned using Clustal Omega [187].

3.5.4. Construction of *amtR* deletion mutant and complementation vector

The creation of an *amtR* markerless deletion was performed using the two-step method for integration and excision as previously described [165]. In brief, the regions flanking either side of the *amtR* gene were amplified by PCR using primer pairs AmtR LF F and AmtR LF o/lap R to obtain the left flank (LF; 920 bp) and AmtR RF o/lap F and AmtR RF R to obtain the right flank (RF; 883 bp) regions. To create the markerless deletion insert, both flanks were used as the template for an overlap PCR reaction. The assembled insert containing the left and right flank DNA was then cloned into the SpeI site of pPR23-derived vector pX33 generating pJR500, which was electroporated into *M. smegmatis* mc^2155 background [9,188]. For selection of single recombination events, a culture of *M. smegmatis* mc^2155 harboring pJR500 was grown at 28°C with agitation (200 rpm) to an OD_{600} of approximately 0.6 – 0.8. Aliquots of culture were plated onto LBT Gentamycin and grown at 42°C. Single colonies were selected and grown in LBT at 37°C to an OD_{600} of 0.6. Aliquots of culture were plated onto low-salt (2 g NaCl l^{-1}) LBT plates containing 10% sucrose and incubated 42°C, selecting for double-recombinant events. To confirm that the correct double recombination event had occurred, Southern blot hybridisation was performed on PvuII-digested genomic DNA using the Amersham Gene Images AlkPhos Direct Labelling and Detection System with CDP-Star

detection reagent (GE Healthcare) according to manufacturers instructions, generating *M. smegmatis* JR258 Δ*amtR*.

To create an AmtR overexpression construct, we amplified a PCR product from *M. smegmatis* containing the *amtR* gene using primers AmtR pMind RBS R and AmtR pMind F. The reverse primer introduced a synthetic ribosome-binding sequence (GGAGG) upstream of the *amtR* gene, which was obtained a from nonpolar kanamycin resistance cassette (*aphA-3*) to ensure translation. The PCR product containing the *amtR* gene was then cloned into the tetracycline inducible vector pMind resulting in pMind-*amtR* [189].

All primers are listed in Appendix 3.

3.5.5. Microarray analysis and quantitative real-time PCR

To identify the regulon of the AmtR protein in *M. smegmatis* we performed microarray analysis comparing strains JR258 Δ*amtR* and wild-type mc^2155 grown in HdB minimal medium with 0.2% (w/v) glycerol as the carbon source and 0.05% (w/v) Tyloxapol. We replaced $(NH_4)_2SO_4$ as the sole nitrogen and sulphur source with 11.4 mM K_2SO_4 as sulphur source and 1 mM L-lysine as the sole nitrogen source. Single colonies were used as inoculum into HdB for initial starter cultures and grown O/N at 37°C to an OD_{600} of 0.2-0.3. This was used as an inoculum for HdB to an OD_{600} of 0.001. Cells were harvested at an OD_{600} between 0.17-0.2. Total RNA was extracted using TRIzol reagent (Ambion) according to the manufacturers instructions. The cells were lysed by four cycles in a mini-Beadbeater (Biospec Products) at 4800 rpm for 30 s, with 30 s on ice between each of the cycles. DNA was removed by treatment with 3 U RNase-free DNase using the TURBO DNA-free kit (Ambion) according to the manufacturers instructions. The quality of the RNA was checked with the Bioanalyzer and the concentration was determined using a NanoDrop ND-100 spectrophotometer. RT-PCR was performed using SuperScript III (Invitrogen) according to the manufacturers instructions for cDNA synthesis and Phusion High-Fidelity PCR Kit (New England Biolabs) for PCR. Microarray was performed using TIGR slides as previously described [165]. TIGR Microarray slides are prepared by printing amplified PCR products to an aminosilane-coated glass microscope slide. Each slide contains three probes per gene.

For qPCR, cDNA was synthesised using 800 ng of RNA for each sample using the SuperScript III reverse transcriptase kit (Invitrogen). After cDNA synthesis, qPCR was carried out using Platinum SYBR Green qPCR SuperMix-UDG with ROX (Invitrogen). Primers for 10 reactions were designed with the Primer 3 software and optimisation was performed (for primers see Appendix 2). The qPCR reactions were carried out on a 7500 Fast Real-Time PCR System (Applied Biosystems). As an internal control and for normalisation of results, we used the gene *sigA*.

3.5.6. Electrophoretic mobility shift assay

Purified *M. smegmatis* AmtR was used in *in vitro* DNA binding assays using a DIG Gel-Shift Kit, 2nd Generation (Roche) to 3'-end label target DNA with DIG. The DNA fragments were labelled and used in gel-shift reactions according to the manufacturers instructions. For use in the assay, AmtR was diluted in 20 mM phosphate buffer, pH 7.4 containing 200 mM NaCl. Protein/DNA molar ratios were calculated to achieve different protein concentrations required. Target DNA and AmtR were added in 10 µL reactions and binding buffer (2 µL), poly(dI-dC) (1 mg mL^{-1}) (0.5 µL) and DIG-labelled target DNA were added as per manufacturers instructions. All reactions were incubated at room temperature for 20 min. Electrophoresis, blotting and detection of DIG-labelled DNA were performed as previously described [190].

3.5.7. Differential scanning fluorimetry

Differential scanning fluorimetry assay was performed as previously described by Niesen *et al.*, to detect ligand interactions that promote protein stability [191]. Optimal concentrations of protein and SYPRO® Orange Dye were determined by performing the assay with varying concentrations of protein and dye. After optimisation of the assay, further experiments were performed with final concentrations (f.c.) of 5 µM AmtR protein and 5x SYPRO® Orange Dye. In single wells of an opaque 96-well plate (LightCycler® 480 Multiwell Plate 96, Roche), a 20 µL reaction was conducted by combining 5 µM AmtR protein in PBS buffer, 5x SYPRO® Orange Dye and PBS buffer. When stated,

DNA was added at a molar ratio of 30:1 (AmtR:DNA). To detect ligand interactions, compounds were obtained from the Biolog Plate PM3, resuspended in 50 µL PBS buffer, resulting in an approximate f.c. of 20 mM and 2 µL of solution containing either compounds (f.c. 2 mM) or PBS buffer (as negative control) were added to the wells.

Melting temperature of the AmtR protein was determined using a LightCycler®480 (Roche). Samples were equilibrated for 5 minutes at 25°C and then temperature was increased to 90°C at a rate of 1°C/minute, taking a fluorescence reading every 0.07°C. The melting point of the protein was calculated by the LightCycler®480 software v1.5.1.62.

3.5.8. Identification of AmtR consensus binding motif

The MEME/MAST tool (http://meme-suite.org/) was used to detect a conserved AmtR consensus binding motif in the upstream regions of genes that showed differential gene expression in *M. smegmatis* JR258 Δ*amtR*. The confirmed motif in the promoter region of *msmeg_2184* was used as input to screen upstream regions (250 bp upstream of +1) of all annotated *M. smegmatis* genes [192].

Chapter 4

The ToxR-like transcriptional regulator CadC is involved in the regulation of lysine and diaminopimelate biosynthesis in *Mycobacterium smegmatis*

Michael Petridis *, Zhi Shean Tee and Gregory M. Cook †.

Article In Preparation: Journal of Bacteriology

Author contributions: Authors are listed in order of magnitude of their contribution in each role. First authors are indicated by an asterisk (*) and corresponding authors are indicated by a cross (†). MP and GMC planned research. MP and ZST performed the research. MP and GMC interpreted data and wrote the manuscript. MP and GMC had supervisory roles. MP produced Figures 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7 and Tables 4.1. ZST contributed to Figure 4.4.

4.1. Abstract

Lysine is a proteinogenic amino acid that is associated with different biological functions (e.g. osmoprotectant, proton sink, component of peptidoglycan) and its biosynthesis is non-essential in the presence of external lysine in *M. smegmatis*. Diaminopimelic acid (DAP) is a precursor in lysine biosynthesis and a substantial component of peptidoglycan in mycobacteria. In *M. smegmatis* disruption of the lysine biosynthesis pathway results in lysine and DAP auxotroph mutants with severe defects in growth and survival. However, despite the fact that lysine metabolism is important to adjust the cellular integrity and the metabolic state in mycobacteria, the regulatory mechanism controlling lysine and DAP biosynthesis remain unknown. Using physiological studies, we demonstrate that deletion of the gene *cadC* (MSMEG_3297) in *M. smegmatis* resulted in a severe growth defect manifested as cell lysis during growth on rich medium, while the Δ*cadC* mutant failed to grow on minimal medium. Subsequently, we performed genome-wide expression profiling comparing the Δ*cadC* mutant and the wild-type and showed elevated expression of several genes in the Δ*cadC* mutant that are involved in the DAP and lysine biosynthesis pathway, including aspartokinase, dihydrodipicolinate synthase and diaminopimelate decarboxylase. Supplementation of either LBT or minimal medium with DAP or lysine could not rescue the growth phenotype of the Δ*cadC* mutant. *M. smegmatis* has a high-affinity lysine uptake system and *M. smegmatis* wild-type exhibited high rates of lysine transport during growth in minimal medium. However, lysine transport in LBT medium was highly reduced in both the wild-type and the Δ*cadC* mutant. Our data suggest that *M. smegmatis* Δ*cadC* is defective in the generation or replenishment of intracellular lysine and DAP levels that are essential for growth and survival in mycobacteria.

4.2. Introduction

Soil-dwelling actinomycetes face rapidly changing environments and require mechanisms to adapt to environmental stress conditions (e.g. nutrient limitation [15,41,43,193], high salinity [194], oxidative stress [165,195] and soil acidification [196]). Bacteria have developed different adaptation mechanisms to alter intracellular physiological processes. This includes the genetic adaptation involving transcriptional changes in gene expression [15,41,43,193], but also the involvement of amino acids and their derivatives in various protective mechanisms [165,194]. Amino acids are mainly used as building blocks in protein biosynthesis, however, under certain conditions amino acids and their derivatives can serve as compatible solutes (e.g. L-proline, N-acetyl-β-lysine, betaine) [194,197], exhibit protective effects against methylglyoxal (e.g. L-proline) [165], serve as protective mechanism in acidic environments (e.g. L-lysine) [198] or sustain cellular integrity (e.g. diaminopimelic acid) [199].

Lysine is a proteinogenic amino acid that is associated with different biological functions. N-acetyl-β-lysine exhibits a role as osmoprotectant in methanogenic archaea [197,200], while in enteric bacteria lysine serves as a protective mechanism under acid stress that is regulated by CadC [198,201,202]. The ToxR-like protein CadC is a membrane-anchored orphan transcriptional regulator that combines sensory mechanisms, signal transduction and DNA-binding functions. In *E. coli*, the CadC protein regulates lysine-dependent acid adaptation [201-203]. After dimerization of its periplasmic domain, CadC regulates expression of the *cadBA* operon that encodes for a lysine decarboxylase (CadA) and a lysine/cadaverine antiporter (CadB) (Figure 4.1) [204]. CadA converts lysine in a proton-requiring reaction to carbon dioxide and cadaverine that is then exported via CadB [205]. Transcription of the *cadBA* operon in *E. coli* is activated, when the external pH drops below 6.6 and also depends on the external availability of lysine that is sensed by a separately regulated lysine permease LysP, which transmits this signal to CadC (Figure 4.1) [198,202,204]. *M. smegmatis* encodes for a CadC-like protein in its genome that was upregulated 2.78-fold under nitrogen

limitation (Chapter 2), while genes encoding for CadB and CadA are missing. Therefore, the role of CadC in the regulation of lysine uptake or its metabolism, as well as its role in mediating the transcriptomic response to nitrogen limitation in *M. smegmatis* are unclear.

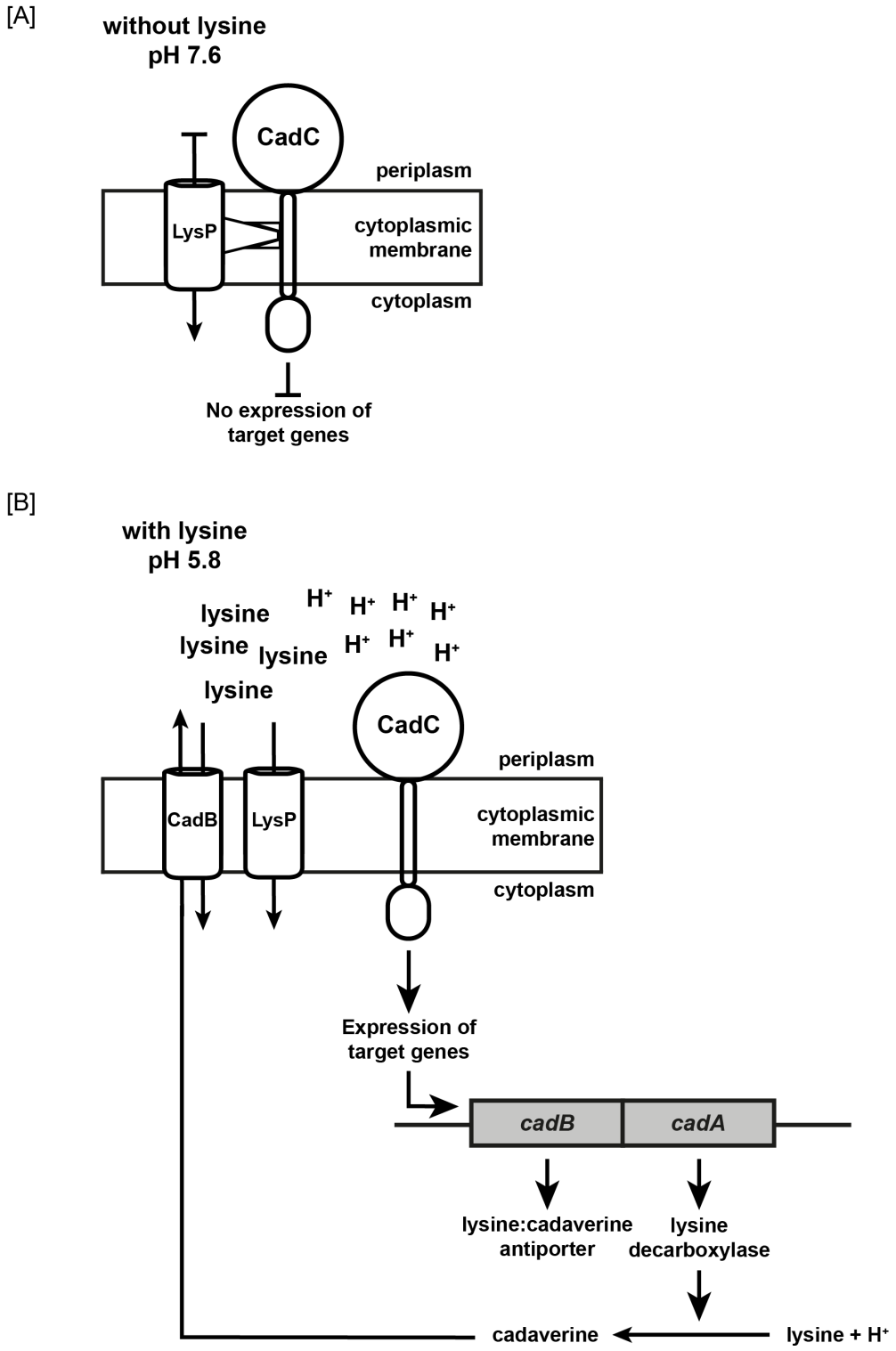


Figure 4.1. Model for the regulatory interaction between CadC and LysP in *E. coli*. [A] Under non-inducing conditions (e.g. pH 7.6, without external lysine) LysP inhibits CadC via transmembrane and periplasmic interactions, resulting in an inhibition of target gene expression. **[B]** Conformational changes induced in LysP upon binding of lysine weaken its contacts with CadC, allowing CadC activation by low pH and resulting in expression of the *cadBA* operon. Adapted from [198].

A key intermediate in lysine metabolism is diaminopimelic acid (DAP) that is an essential component of peptidoglycan in the bacterial cell wall (Figure 4.2) [206]. DAP is responsible for cross-linking of the peptidoglycan chains that consist of N-acetylglucosamine and N-acetylmuramic acid in the peptidoglycan of mycobacteria [207]. Biosynthesis of DAP is non-essential in the presence of external lysine in *M. smegmatis* [199]. Disruption of the enzyme aspartate kinase (*msmeg_6257*), which catalyses the first step in this pathway, results in lysine auxotrophy with a severe growth defect manifested as cell lysis [199]. Consaul *et al*. have characterised DAP auxotrophs in *M. smegmatis* that grew in absence of DAP due to replacement of DAP with lanthionine in the peptidoglycan. This phenotype resulted from a mutation in the putative ribosome binding site of the *cbs* gene, encoding cystathionine-β-synthase, an enzyme that is part of the cysteine biosynthesis pathway [208,209]. A lysine auxotroph *M. tuberculosis* strain required 7 mM lysine for growth and had an absolute dependency on supplementation of the growth medium with Tween 80 [210]. This detergent is believed to increase the permeability of the mycobacterial cell envelope, perhaps allowing some diffusion of lysine into the cell [210]. An *in vivo* mouse model has further shown that a lysine auxotrophic mutant of *M. tuberculosis* is unable to survive in the host [211].

In mycobacteria, intracellular lysine accumulation is observed during growth on fatty acids [166], however, no lysine uptake system has been functionally characterised in mycobacteria [210]. Previous studies suggested that *M. tuberculosis* is unable to transport extracellular available lysine at physiological concentrations, but relies on lysine biosynthesis to replenish its intracellular lysine pool [210]. Despite the fact that lysine metabolism is important to adjust the cellular integrity and the metabolic state in mycobacteria, the regulatory mechanisms controlling lysine and DAP biosynthesis remain unknown. In this study, we used genome-wide expression profiling and biochemical assays in combination with a physiological characterisation of a Δ*cadC* deletion mutant to provide initial evidence that CadC is involved in the regulation of lysine and DAP biosynthesis in *M. smegmatis*.

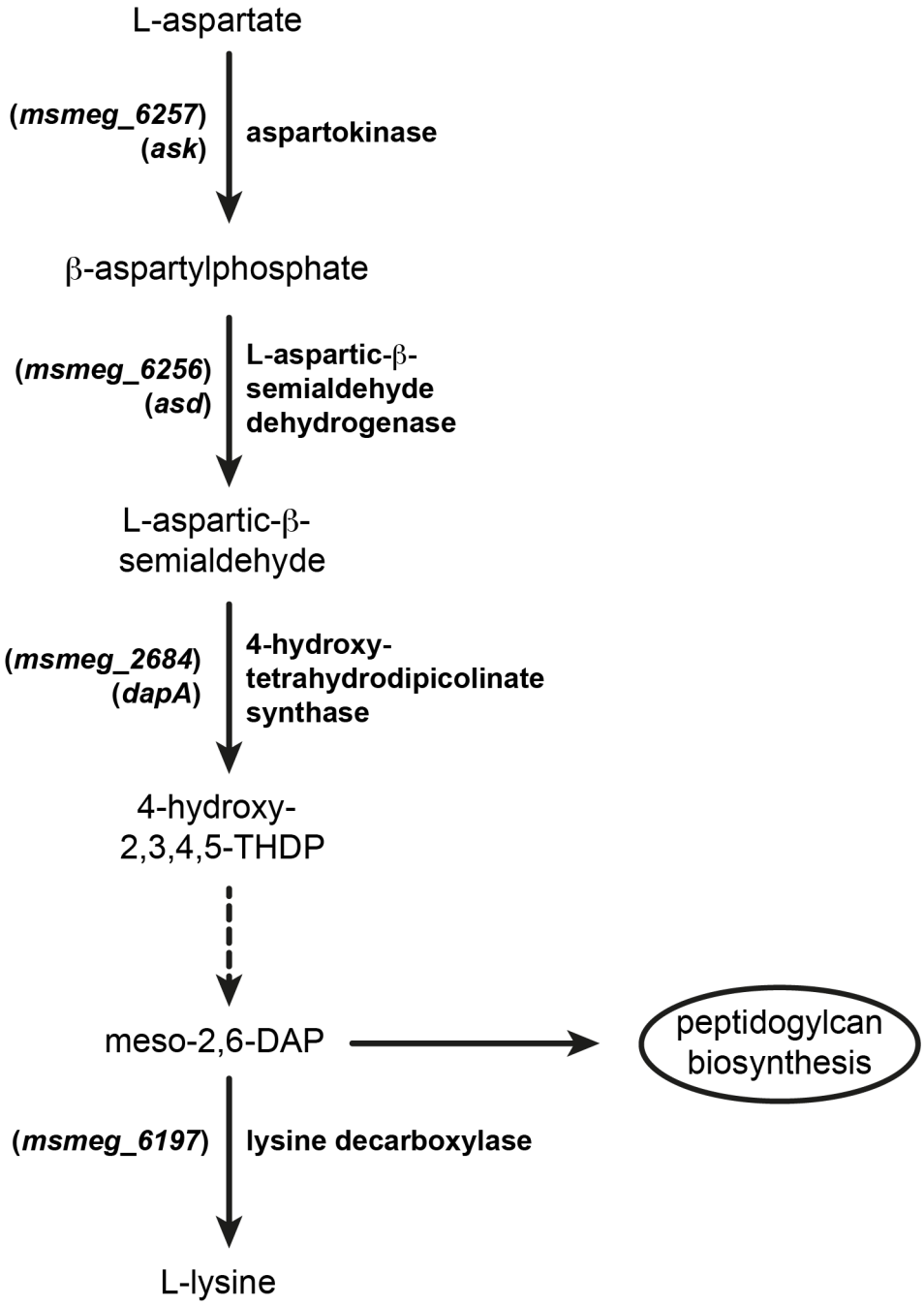


Figure 4.2. The biosynthesis pathway of DAP and lysine from aspartate. The first reaction in this pathway is the phosphorylation of L-aspartate by an aspartokinase to β-aspartylphosphate, followed by the conversion of β-aspartylphosphate to L-aspartic-β-semialdehyde by a dehydrogenase. L-aspartic-β-semialdehyde is converted to meso-2,6-DAP (meso-diamino-pimelate) in a multi-step reaction and meso-2,6-DAP is converted to L-lysine by a lysine decarboxylase. The dotted line indicates a multi-step reaction from 4-hydroxy-2,3,4,5-tetrahydrodipicolinate to meso-2,6-DAP. Gene numbers and gene names are indicated in brackets. 4-OH-THPP: 4-hydroxy-2,3,4,5-tetrahydrodipicolinate; meso-2,6-DAP: meso-2,6-diaminopimelate. Adapted from [199].

4.3. Results

4.3.1. A Δ*cadC* mutant is impaired in growth and survival

In previous work in this thesis (Chapter 2, Table 2.3) the ToxR-like transcriptional regulatory protein CadC (MSMEG_3297) was upregulated 2.8-fold in *M. smegmatis* in response to nitrogen limitation. In addition, enzymes of the lysine biosynthesis pathway (dihydrodipicolinate reductase (*msmeg_3317*) and diaminopimelate decarboxylase (*msmeg_6197*)) were upregulated 3.1-fold and 2.4-fold (Chapter 2, Table 2.2), respectively, under the same conditions of nitrogen limitation. To determine, whether CadC is involved in the regulation of lysine metabolism in *M. smegmatis*, we created a markerless *cadC* deletion mutant (*M. smegmatis* MP3297 Δ*cadC*) and confirmed deletion of 80% of the internal region of *cadC* via PCR and Southern hybridisation (Figure 4.3). In the Southern Blot, we observed signals corresponding to the approximate fragment size of 2.2 kb (wild-type) and 11.7 kb (Δ*cadC* mutant), however, we also observed one weakly hybridising band at around 4 kb in the digest of both wild-type and Δ*cadC* mutant genomic DNA, likely due to non-specific or weakly hybridising probe binding (Figure 4.3B).

Initial growth studies with strain MP3297 Δ*cadC* revealed a severe growth defect during growth on minimal medium (HdB) (Figure 4.4C) and rich medium (LBT) (Figure 4.4A). Growth of strain MP3297 Δ*cadC* on LBT was initially comparable to the wild-type strain mc^2155, however, when cells reached an OD_{600} between 1.4 – 1.8 we observed a continuous decrease in OD_{600} and visual inspection of the cultures indicated clearing of the medium, suggesting cell lysis of the Δ*cadC* mutant (Figure 4.4A). The colony forming units (CFU) of strain MP3297 Δ*cadC* decreased between 48 and 72 hours by 1-log and comparison of CFU of the strains MP3297 Δ*cadC* and mc^2155 wild-type revealed a significant difference of 1-log after 48 hours and 2-log after 72 hours, respectively (Figure 4.4B). The strain MP3297 Δ*cadC* failed to grow in minimal medium that were supplemented with various carbon sources (e.g. either glycerol, glucose or arabinose) (Figure 4.4C).

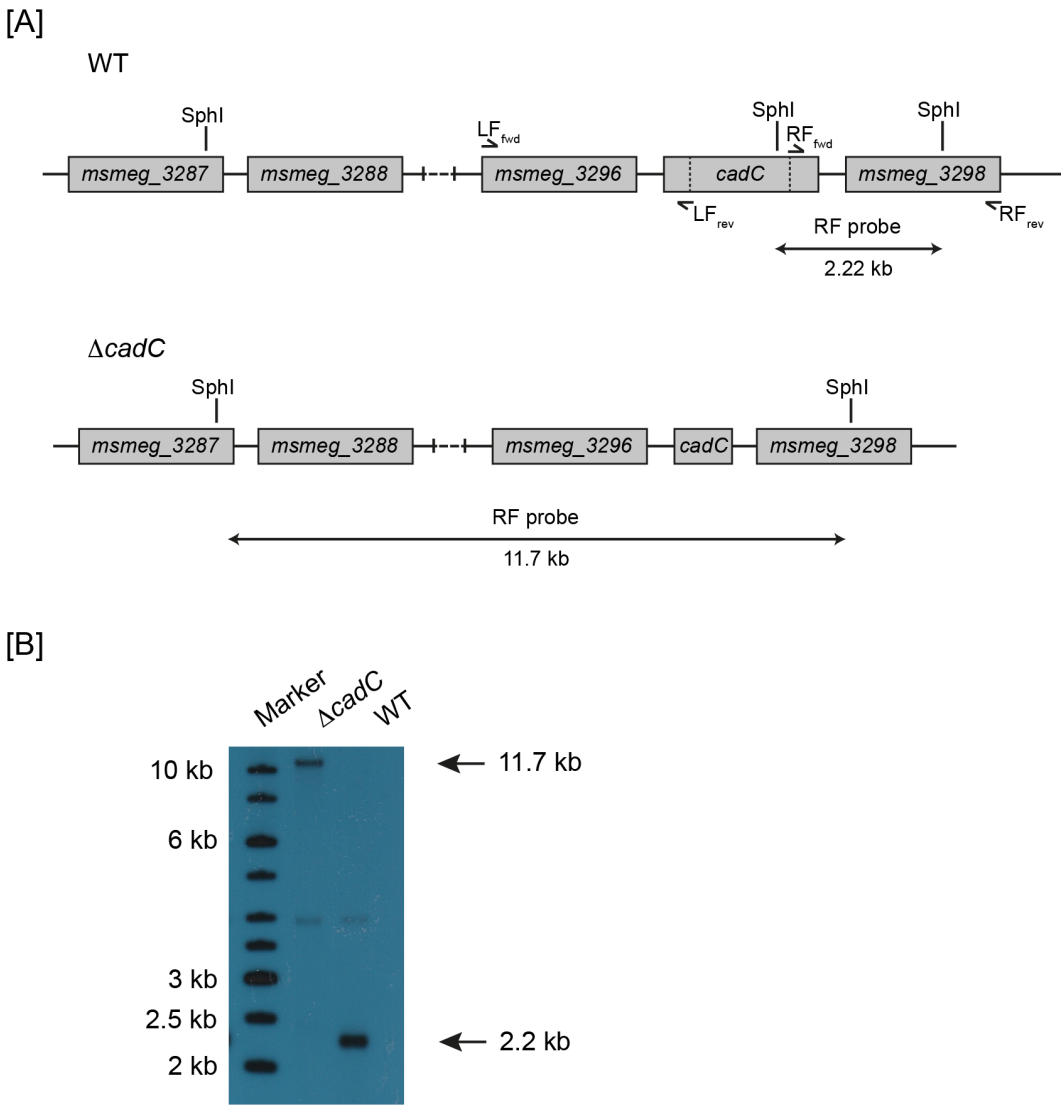


Figure 4.3. Construction of a markerless Δ*cadC* mutant of *M. smegmatis* mc^{2}155. [A] The knockout construct consists of two fragments (left flank spanned by LF_{fwd} and LF_{rev}; right flank spanned by RF_{fwd} and RF_{rev}) on each side of *cadC* in the plasmid pX33. Restriction sites of SphI and fragment sizes as detected in a Southern hybridisation are indicated. **[B]** SphI-digested genomic DNA of WT and a colony of Δ*cadC* were probed with labelled right flank PCR product of the deletion construct. Molecular weight standards with sizes indicated in kilobases are shown on left. WT, *M. smegmatis* mc^{2}155 wild-type; Δ*cadC*, *M. smegmatis* MP3297 Δ*cadC*.

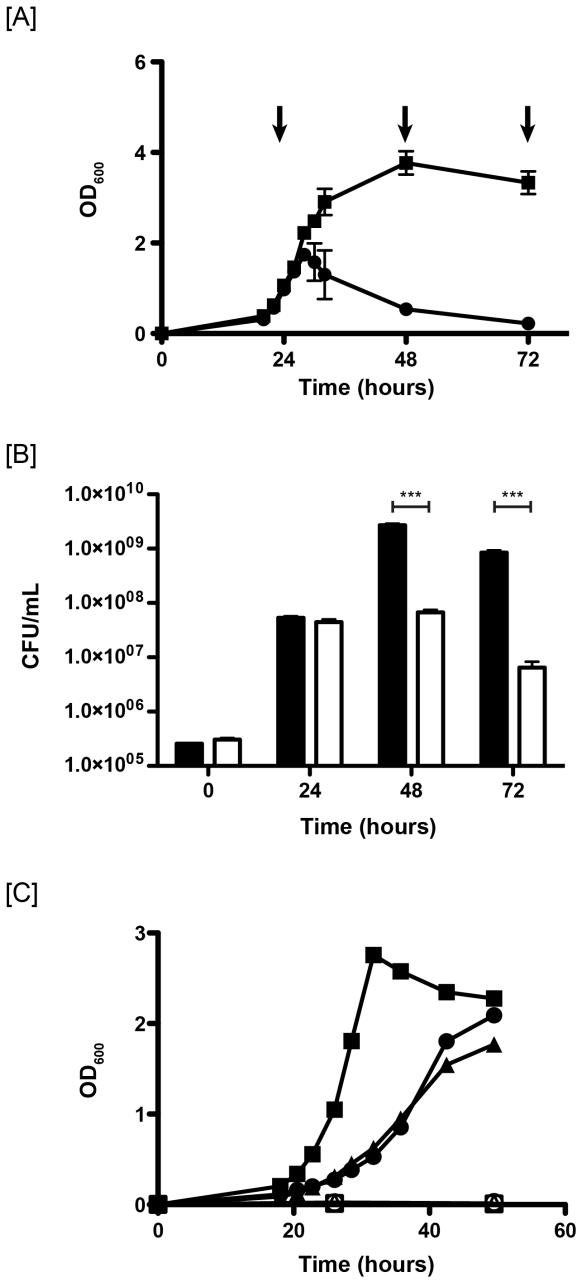


Figure 4.4. Growth and viability of *M. smegmatis* mc^{2}155 wild-type and strain MP3297 Δ*cadC*. Comparative growth analysis was performed by growing the strains mc^{2}155 wild-type and strain MP3297 Δ*cadC* in LBT medium and monitoring **[A]** OD_{600} of mc^{2}155 wild-type (closed squares) and MP3297 Δ*cadC* (closed circles) and **[B]** CFU per mL comparing strains mc^{2}155 wild-type (black bars) and MP3297 Δ*cadC* (white bars) at selected time points (indicated by arrows in panel [A]). **[C]** Comparative growth analyses of mc^{2}155 wild-type (closed symbols) and MP3297 Δ*cadC* (open symbols) in HdB minimal medium containing glycerol (squares), glucose (circles) or arabinose (triangles) as sole carbon source. Error bars represent the standard error of the mean of three independent biological replicates. A two-tailed Student *t*-test was performed to determine a statistical significant difference between mc^{2}155 wild-type and MP3297 Δ*cadC* (p >0.001 = ***). ET contributed to this figure.

4.3.2. Deletion of *cadC* in *M. smegmatis* results in the upregulation of enzymes involved in lysine biosynthesis

To gain further insights into the regulatory function of CadC we conducted a whole-genome RNA-seq analysis comparing strain MP3297 Δ*cadC* to mc^2155 wild-type. For each strain, three independent biological replicates were grown in LBT medium and harvested at an OD_{600} between 1.4 and 1.6 (Figure 4.4A). The RNA-seq analysis of the transcriptome revealed a total of 480 differentially expressed genes in strain MP3297 Δ*cadC* (>2-fold change, $P<0.05$) i.e. approximately 7% of the *M. smegmatis* genome. In total, 203 genes were upregulated and 277 genes downregulated (Table 4.2, supplementary data). Functional classification of these genes showed major changes in expression of proteins involved in transcription and translation and in the group of hypothetical and uncharacterised proteins. In total, 45 genes involved in nitrogen metabolism, twenty genes involved in amino acid metabolism and five genes involved in nucleotide metabolism were differentially expressed in strain MP3297 Δ*cadC* (Figure 4.5).

In this analysis, we identified elevated expression of six genes that are involved in the biosynthesis of lysine (Figure 4.5), (Table 4.1). Additionally, ten genes were upregulated including adenylosuccinate synthetase (*msmeg*_0759), adenylosuccinate lyase (*msmeg*_5847), 4-aminobutyrate aminotransferase (*msmeg*_2959) and succinate-semialdehyde dehydrogenase (*msmeg*_5912), that link the metabolism of the amino acids glutamate and aspartate via lysine to the central carbon and energy metabolism (Figure 4.5), (Table 4.1). Apart from these genes involved in amino acid metabolism, we found several genes of the central carbon metabolism with reduced expression in the *M. smegmatis* Δ*cadC* mutant (Table 4.1).

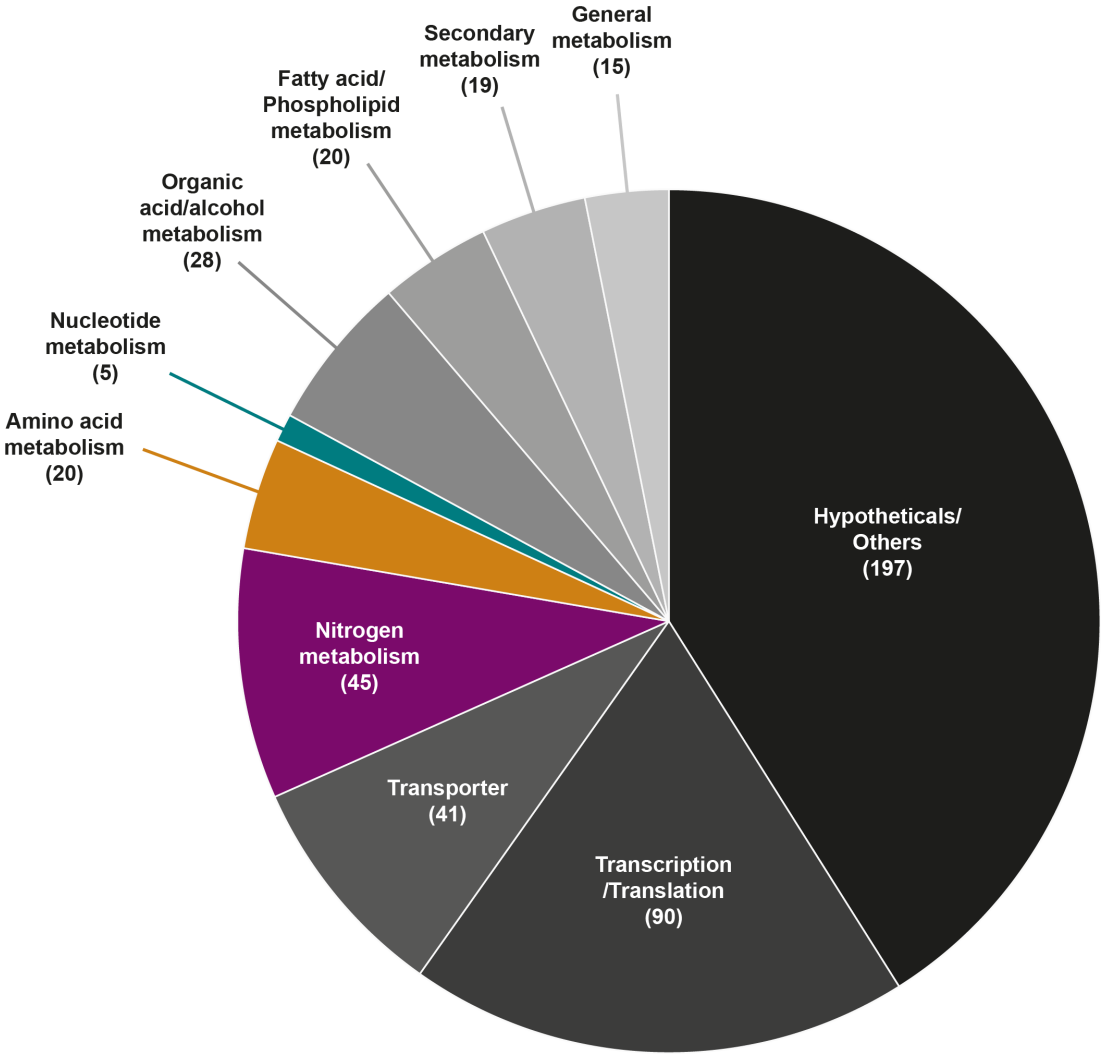


Figure 4.5. Functional categorisation of differentially expressed genes in strain MP3297 Δ*cadC*. Genes were assigned to functional categories using the Database for Annotation, Visualization and Integrated Discovery v6.7 (DAVID) and the NCBI database. Number of genes that belong to a particular category and were differentially expressed in strain MP3297 Δ*cadC* are indicated in brackets. For a detailed list of genes that belong to the group "amino acid metabolism" see Table 4.1.

Table 4.1. Selected genes with differential gene expression in *M. smegmatis* MP3297 Δ*cadC* compared to *M. smegmatis* mc^{2}155 wild-type

	mc^{2}155 locus[a]	Expression ratio[b]	FDR[c]	Description
Lysine metabolism	msmeg_6197	2.69	7.82E-04	diaminopimelate decarboxylase
	msmeg_5286	2.66	4.95E-08	dihydrodipicolinate reductase
	msmeg_6094	1.83	6.27E-07	lysine-tRNA synthetase
	msmeg_2684	1.72	3.16E-02	dihydrodipicolinate synthase
	msmeg_0733	1.69	3.56E-02	dihydrodipicolinate reductase
	msmeg_6257	1.59	9.00E-03	aspartokinase
	msmeg_1764	1.35	2.86E-02	L-lysine aminotransferase
	msmeg_3796	0.49	1.95E-04	lysine-tRNA synthetase
Ala/Asp/Glu metabolism	msmeg_0759	3.56	5.05E-10	adenylosuccinate synthetase
	msmeg_5912	2.73	1.31E-06	succinate-semialdehyde dehydrogenase
	msmeg_3046	2.11	1.11E-05	carbamoyl-phosphate synthase; small subunit
	msmeg_5119	1.90	4.00E-02	1-pyrroline-5-carboxylate dehydrogenase
	msmeg_5800	1.68	1.44E-03	amidophosphoribosyltransferase
	msmeg_2959	1.63	4.22E-02	4-aminobutyrate aminotransferase
	msmeg_3173	1.60	4.98E-02	L-asparaginase
	msmeg_2659	1.50	2.55E-03	alanine dehydrogenase
	msmeg_5847	1.42	4.93E-02	adenylosuccinate lyase
	msmeg_3047	1.41	4.73E-02	carbamoyl phosphate synthase large subunit
Central carbon metabolism	msmeg_1986	8.00	7.20E-14	tartrate dehydrogenase
	msmeg_1670	0.71	3.08E-02	succinate dehydrogenase, flavoprotein

	msmeg_0255	0.69	1.14E-02	phosphoenolpyruvate carboxykinase
	msmeg_5939	0.66	1.30E-02	acetaldehyde dehydrogenase
	msmeg_6785	0.41	4.80E-04	triosephosphate isomerase
	msmeg_5826	0.23	1.61E-06	pyruvate decarboxylase

[a] Locus number of gene in M. smegmatis mc^2155

[b] Mean gene expression ratio of three biological replicates

[c] *P*-values of gene expression ratio from three biological replicates were corrected for multiple testing using the Benjamini and Hochberg False Discovery Rate

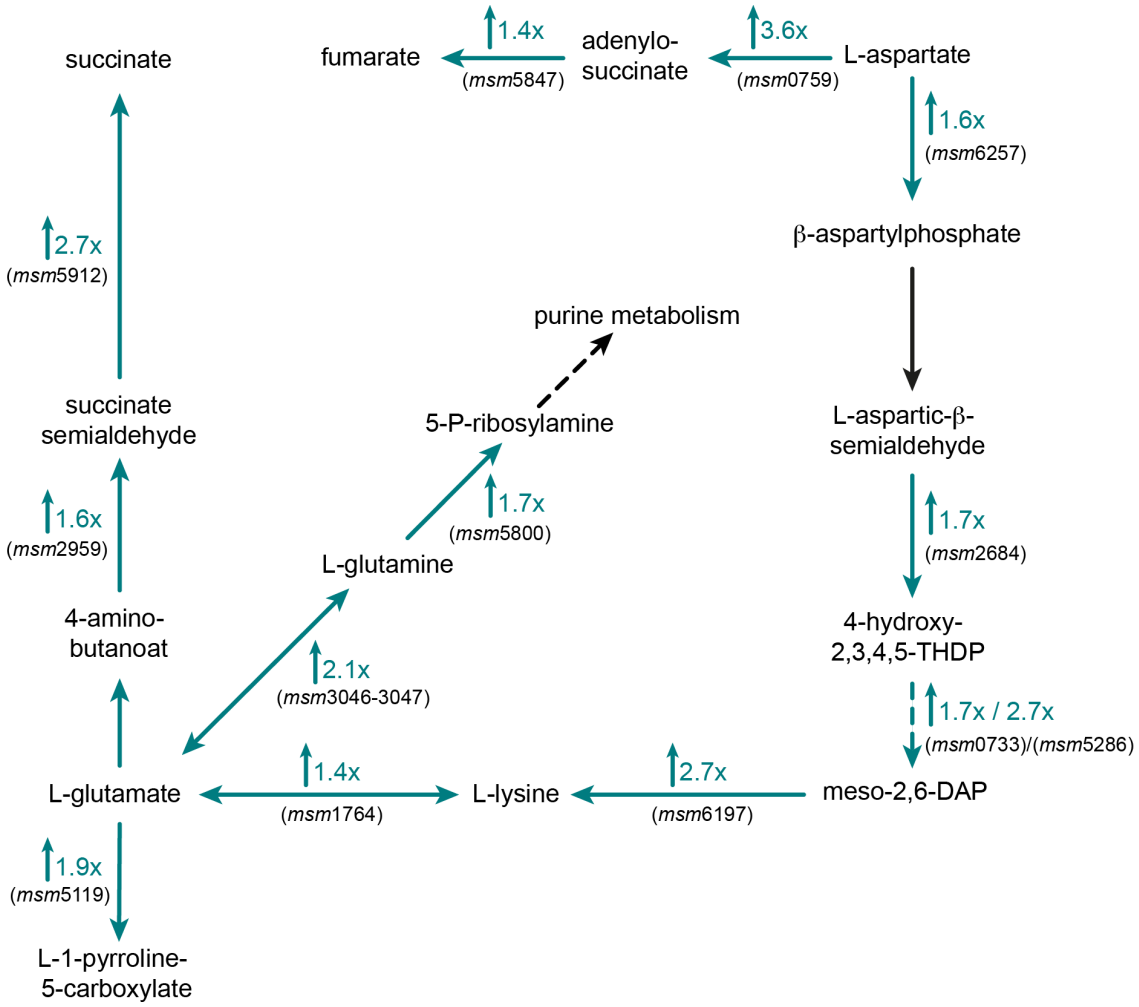


Figure 4.6. Diagram of the differentially expressed genes in strain MP3297 Δ*cadC* involved in the metabolism of lysine, glutamate and aspartate. Fold change in gene expression and predicted directionality of reaction are indicated by numbers and arrows between metabolites. Small arrows indicate upregulation (arrow up) or downregulation (arrow down) of gene expression. Gene IDs are indicated as abbreviations (msm6197, msmeg_6197) in brackets. (4-hydroxy-2,3,4,5-THDP: 4-hydroxy-2,3,4,5-tetrahydrodipicolinate; meso-2,6-DAP: meso-2,6-diaminopimelate).

4.3.3. *M. smegmatis* has a high-affinity lysine uptake system

Our whole-genome analysis revealed upregulation of three metabolic pathways that are linked to each other via the amino acid lysine; degradation of aspartate to DAP, which is a key component of bacterial peptidoglycan, decarboxylation of DAP to lysine, and amination of lysine to glutamate that is directed into several pathways (e.g. purine metabolism, succinate production) form two metabolic routes that are connected via succinate and fumarate (Figure 4.6). Next, we wanted to define the capability of different mycobacterial strains to transport externally available lysine to meet their intracellular lysine demand. We further wanted to know, whether the transcriptional changes in lysine metabolism affected the ability of a Δ*cadC* mutant to transport lysine.

To gain insight into the ability of lysine transport in different mycobacterial species, we used a lysine transport assay (Figure 4.7). Transport studies were conducted with radioactively labelled L-[^{14}C]lysine (final concentration 0.75 µM L-[^{14}C]lysine) in *M. smegmatis* mc^{2}155 wild-type, *M. smegmatis* MP3297 Δ*cadC* and members of the *M. tuberculosis* complex (*M. bovis* BCG (Pasteur) and *M. tuberculosis* mc^{2}6206 Δ*leuCD* Δ*panCD*).

Our results clearly show that *M. smegmatis* harbors a high-affinity lysine uptake system. Lysine uptake is only observed in minimal medium with a rate of 3.91 ± 0.17 nmol.min^{-1} (mg protein)$^{-1}$, while this rate drops significantly to 0.02 ± 0.01 nmol.min^{-1} (mg protein)$^{-1}$ in rich medium (e.g. LBT) (Figure 4.7). To test, whether lysine transport is dependent on nutrient starvation to induce expression of the lysine transport system, we starved cells of *M. smegmatis* wild-type for 24 hours in PBST buffer and resuspended the cells in HdB (Figure 4.6A). However, we could not detect a significant increase in the lysine uptake rate of *M. smegmatis* after exposure to nutrient starvation (Figure 4.7).

Subsequently, we wanted to determine whether the lysine uptake rate of the Δ*cadC* mutant was affected by the transcriptional changes in the lysine biosynthesis and lysine degradation pathways. The lysine uptake rate of strain MP3297 Δ*cadC* was 0.04 ± 0.01 nmol.min^{-1} (mg protein)$^{-1}$, when grown in LBT medium, which was comparable to the wild-type strain mc^{2}155.

In *M. tuberculosis*, lysine accumulation is closely associated with growth on fatty acids [166], therefore we also wanted to test members of the *M. tuberculosis* complex for their ability to transport lysine. We found that both *M. bovis* BCG (Pasteur) and *M. tuberculosis* $mc^2$6206 showed a low lysine uptake rate (0.012 $nmol.min^{-1}$ $(mg\ protein)^{-1}$ and 0.099 $nmol.min^{-1}$ $(mg\ protein)^{-1}$, respectively) under the conditions tested (Figure 4.7).

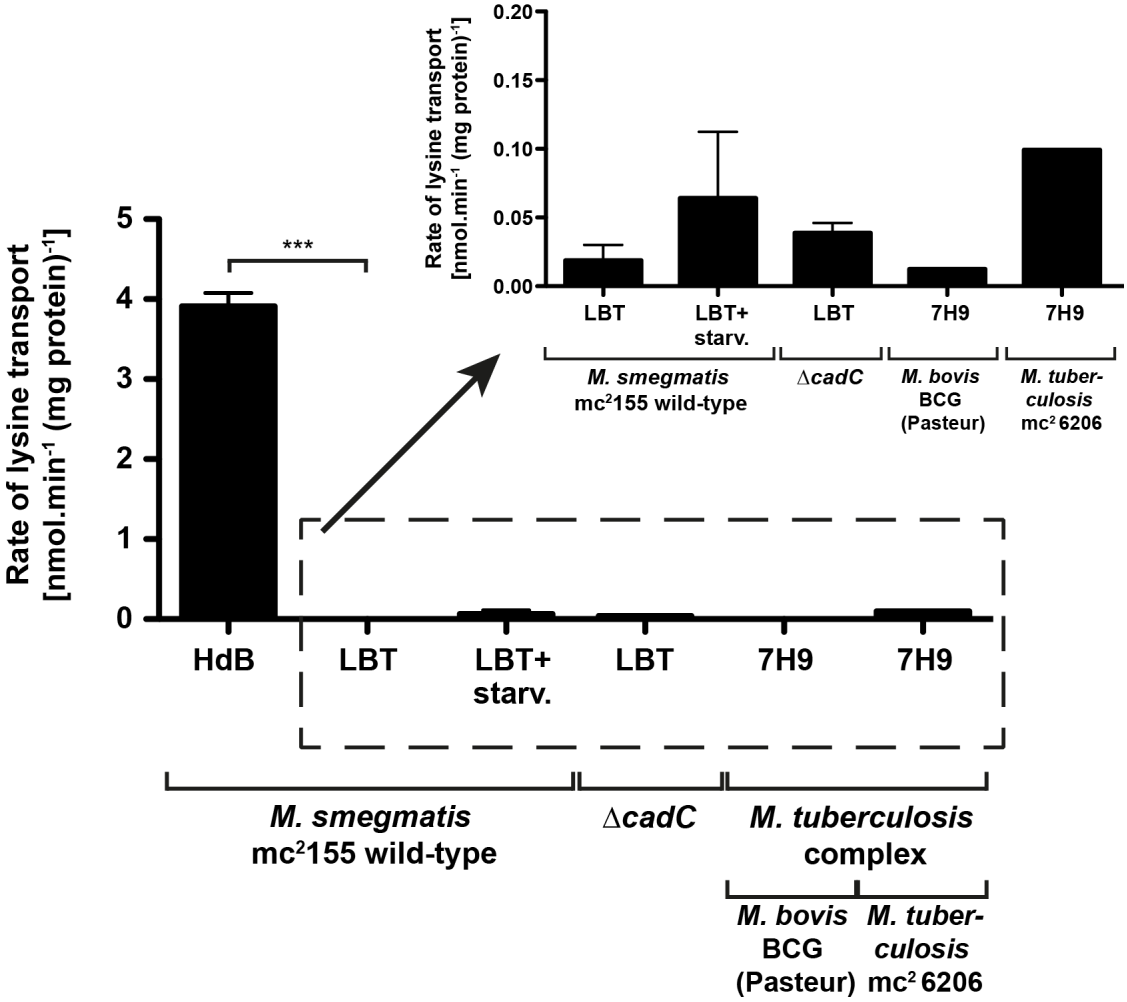


Figure 4.7. Determination of the lysine transport rate in *M. smegmatis* and in members of the *M. tuberculosis* complex. Indicated is the rate of L-[^{14}C]lysine transport in strain mc^{2}155 wild-type, strain MP3297 Δ*cadC*, *M. bovis* BCG (Pasteur) and *M. tuberculosis* mc^{2}6206 Δ*leuCD* Δ*panCD* grown in minimal medium (HdB), rich medium (LBT), LBT followed by 24 hours starvation in PBST (LBT+starv.) and Middlebrook 7H9. Inset: Highlighted data are shown; note the different y-axis. The final concentration of L-[^{14}C]lysine in each assay was 0.75 µM. The error bars indicate the standard deviation of three biologically independent experiments. Lysine transport assay with *M. bovis* BCG (Pasteur) and *M. tuberculosis* mc^{2}6206 Δ*leuCD* Δ*panCD* were performed in single biological replicates. A two-tailed Student *t*-test was performed to determine a statistical significant difference between lysine transport of strain mc^{2}155 wild-type in HdB and LBT medium (p >0.001 = ***).

4.3.4. DAP and lysine do not rescue the growth defect of a ΔcadC *mutant*

DAP is not only required for the synthesis of lysine, but also important for the cellular integrity as key component of peptidoglycan. We wanted to test, whether the decrease in OD_{600} in LBT medium is caused by depletion of DAP. For this, a mid-exponential-phase culture of strain MP3297 Δ*cadC* was used to subculture fresh LBT medium and the OD_{600} was monitored over time. At an OD_{600} of 0.5 the medium was supplemented with either DAP or lysine and compared to a control without any supplementation (Figure 4.8A). Growth was under all conditions comparable and began to plateau after 30 hours, when a decrease in OD_{600} and drop in culture turbidity started to occur for the non-supplemented control and the DAP- and lysine-supplemented cell cultures (FIgure 4.8A). Supplementation of the growth medium with either DAP or lysine did not rescue the Δ*cadC* mutant and we observed a comparable decrease in OD_{600} and CFU as for the non-supplemented cell culture (Figure 4.8A). The non-supplemented culture of strain MP3297 Δ*cadC* revealed a decrease of 58% in OD_{600} within the first five hours and both DAP- and lysine-supplemented cultures showed a decrease of 45% in the OD_{600} the same time (Figure 4.8A). We further determined the cell survival rate of a lysine-supplemented culture of strain MP3297 Δ*cadC* and compared this to the non-supplemented control. Ten hours after cell lysis started to occur, we observed a cell survival rate of 72.4% the lysine-supplemented cell culture and a cell survival rate of 51.3% for the non-supplemented cell culture (Figure 4.8B).

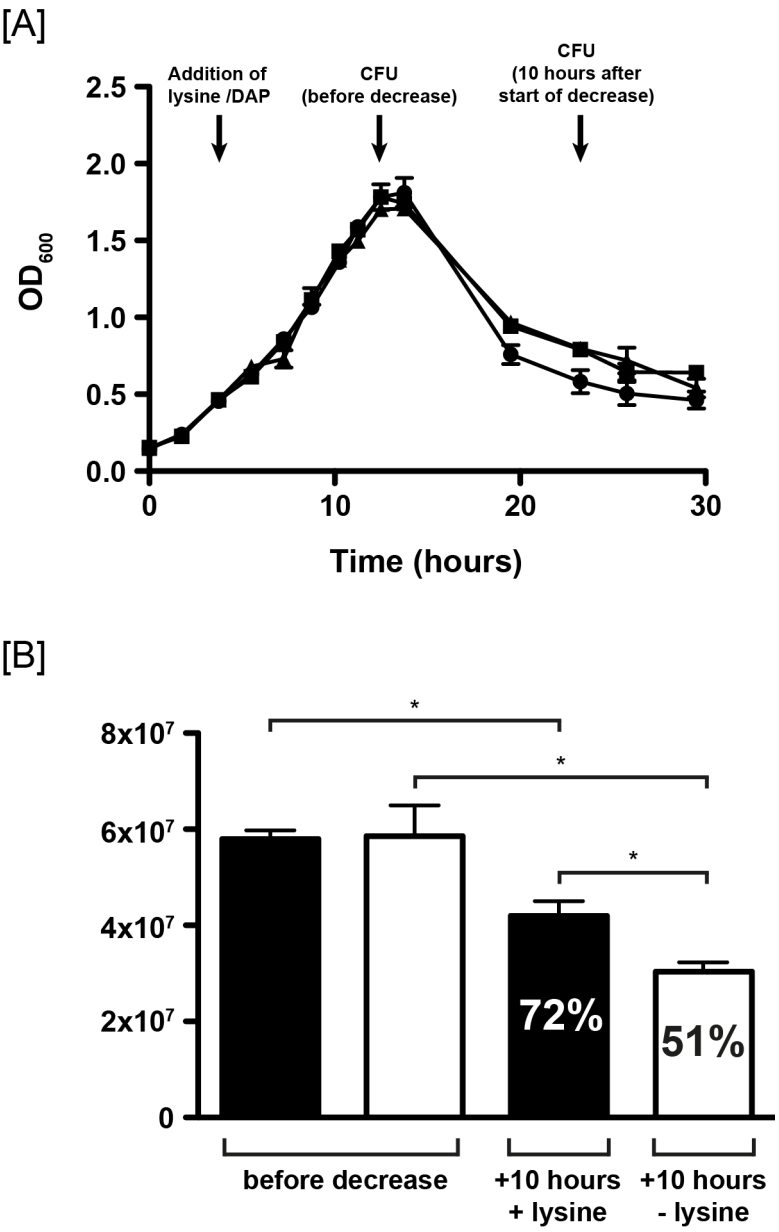


Figure 4.8. DAP and lysine do not rescue the growth defect of strain MP3297 Δ*cadC*. To investigate the effect of DAP and lysine on cell growth and survival, 1 mM DAP (closed squares) and 1 mM lysine (closed triangles) were added to a mid-exponential-phase culture (OD_{600} = 0.5; indicated with an arrow) and cell lysis was compared to a non-supplemented control (closed circles). Cell lysis was determined as **[A]** decrease in OD_{600} over time and **[B]** CFU measurements with calculation of the cell survival rate (indicated as percentage in the bars) of a culture supplemented with 1 mM lysine compared to a non-supplemented culture ten hours after cell lysis started to occur.

4.4. Discussion

In mycobacteria, lysine accumulation is observed during growth on fatty acids and a lysine auxotrophic mutant of *M. tuberculosis* is unable to survive in a mouse model [166,211]. The amino acid L-lysine is associated with different biological functions (e.g. peptidoglycan biosynthesis, osmo-protectant, acid-protectant) [197-202,206]. However, the knowledge surrounding lysine metabolism and its regulation in mycobacteria is scarce. In this study, we provide initial evidence that sustainment of intracellular lysine and DAP levels in *M. smegmatis* is dependent on CadC and show the importance of CadC for mycobacterial growth and survival.

Our whole-genome RNA-seq analysis revealed that expression of the lysine biosynthesis pathway was elevated in *M. smegmatis* MP3297 Δ*cadC*. Lysine biosynthesis from aspartate is a six-step reaction and the two enzymes aspartate kinase (*ask*) and dihydrodipicolinate synthase (*dapA*) are responsible for the control of the metabolic flux during the conversion of aspartate to lysine [212]. This pathway includes biosynthesis of DAP that is required for peptidoglycan biosynthesis and DAP biosynthesis is conditionally essential in *M. smegmatis* [199]. Disruption of the enzyme aspartate kinase, that is catalysing the first step in this pathway produced a lysine auxotroph with a severe growth defect ultimately resulting in cell lysis, similar to *S. typhimurium* [199,213]. Our data clearly show that a deletion of the *cadC* gene in *M. smegmatis* had a comparable effect on growth and survival. After initial growth in LBT, we observed a strong decrease in OD_{600} and CFU measurements suggested that cell death was occuring during exponential growth phase. The *M. smegmatis* strain MP3297 Δ*cadC* mutant failed to grow on minimal medium supplemented with various carbon sources. This growth deficiency on defined minimal medium may be explained due to the limited nutrient availability and a lack in DAP precursors necessary for permanent peptidoglycan biosynthesis in replicating cells. The lysine uptake rate of *M. smegmatis* wild-type was significantly increased during growth in minimal medium even in absence of external lysine, indicating that expression of the high-affinity lysine uptake system is not dependent on the availability of external lysine, but occurs in response to a yet unidentified stimulus.

M. smegmatis might prefer uptake of lysine to generate DAP rather than to run DAP biosynthesis from aspartate, while the requirement for lysine uptake in rich media like LBT that contains about 5-7 mM L-lysine would be minor due to an abundant nutrient availability allowing sufficient DAP biosynthesis [214]. Furthermore, we could not detect a significant increase in the lysine uptake rate after cells were exposed to nutrient starvation, suggesting that expression of the high-affinity lysine uptake system is not regulated by nutrient starvation. However, our lysine transport experiments did not exclude the possibility that the genome of *M. smegmatis* encodes for multiple lysine transport systems. Lysine transport in HdB occured via a high-affinity lysine transport system that allowed lysine transport at concentrations below 1 µM lysine, although kinetic parameters such as the Michaelis constant (K_m) and the maximum velocity (V_{max}) of this lysine transport system were not determined in this work. Further work is required to identify the lysine transport system(s) in *M. smegmatis*, followed by their functional characterisation, to get insights into the essentiality of lysine for the balance of the metabolic state in mycobacteria.

Our results demonstrate that external addition of DAP or lysine fail to rescue cell lysis that is likely caused by DAP deprivation. In previous studies, supplementation of LBT with DAP for growth of an aspartate kinase deficient strain of *M. smegmatis* resulted in growth, while the aspartate kinase deficient strain of *M. smegmatis* failed to grow in non-supplemented LBT medium [201]. DAP biosynthesis in the Δ*ask* mutant was completely abolished, therefore supplementation with DAP resulted in growth of the *ask* mutant. However, in strain MP3297 Δ*cadC* mutant the DAP biosynthesis pathway was still intact, although regulation of this pathway through CadC is likely to be abolished. The general presence of a functional DAP biosynthesis pathway in *M. smegmatis* strain MP3297 Δ*cadC* would explain, why we observe initial growth and only a moderate effect after supplementation with DAP or lysine on cell growth and survival compared to an *ask* mutant of *M. smegmatis* [199]. However, we could not identify the cause for the suddenly occuring cell lysis and further work is now required to understand the molecular mechanism of CadC and its regulatory impact on lysine and DAP metabolism.

We show that *M. smegmatis* harbors a functional lysine uptake system, while we did not observe lysine transport for members of the *M. tuberculosis* complex under the conditions tested. This would suggest that these organisms rely on intracellular lysine biosynthesis mechanisms rather than lysine transport to replenish their intracellular lysine pool. This is consistent with previous studies by Pavelka and Jacobs (1999) demonstrating that a lysine auxotrophic mutant of *M. smegmatis* can grow on lysine-supplemented minimal medium, while a *M. tuberculosis* lysine auxotroph required 25-fold more lysine than *M. smegmatis* and furthermore required supplementation of the growth medium with Tween 80. This detergent is believed to increase the permeability of the mycobacterial cell envolpe, perhaps allowing slight diffusion of lysine [210]. Both soil-dwelling and pathogenic mycobacteria harbor numerous transport systems for amino acids. The importance of these uptake systems for the pathogenicity of *M. tuberculosis* is under investigation and until now only a small number has been functionally characterised [18,19]. Although members of the *M. tuberculosis* complex seem to lack a functional lysine uptake system, they harbor two genes encoding for putative lysine exporters in their genome that show a high sequence similarity to their homologs in *M. smegmatis* [45]. Future work will focus on the biochemical and biophysical characterisation of the lysine import and lysine export systems in mycobacteria and their importance for mycobacterial physiology.

Our data demonstrate that deletion of *cadC* in *M. smegmatis* resulted in a severe growth defect manifested as cell lysis during growth on rich medium, suggesting depletion of DAP under these conditions that has a direct impact on peptidoglycan biosynthesis and cellular integrity. However, the molecular mechanism of the CadC-dependent maintenance of intracellular lysine and DAP levels remains to be elucidated and subsequent work is required to ultimately understand lysine metabolism in mycobacteria.

4.5. Materials and Methods

4.5.1. Bacterial strains, media and growth conditions

E. coli DH10B was grown in LB medium at 37°C with agitation (200 rpm) or on LB agar plates. *M. smegmatis* mc^2155 and derived strains were grown in LBT or HdB minimal medium supplemented with 0.2% (w/v) glycerol (unless otherwise stated) and 0.05% (w/v) Tween-80 (in LBT) or 0.05% (w/v) Tyloxapol (in HdB) at 37°C with agitation (200 rpm). Growth curves were performed in triplicate. Bacterial cell viability was monitored by cell counts based on CFUs per mL, where serial dilutions of bacterial cell culture in phosphate buffered saline with 0.05% (w/v) Tween-80 (PBST) were plated on LBT agar plates. Selective media contained kanamycin (50 μg mL^{-1} for *E. coli*, 20 μg mL^{-1} for *M. smegmatis*), gentamycin (25 μg mL^{-1} for *E. coli*, 5 μg mL^{-1} for *M. smegmatis*). *M. tuberculosis* mc^26206 and *M. bovis* BCG (Pasteur) were grown in Middlebrook 7H9 medium, supplemented with 0.5% glycerol (w/v) and 0.05% (w/v) tyloxapol. Medium for growth of *M. tuberculosis* mc^26206 Δ*leuCD* Δ*panCD* was further supplemented with 50 μg mL^{-1} L-leucine and 50 μg mL^{-1} pantothenic acid.

All bacterial strains and plasmids used are listed in Appendices 1 & 2.

4.5.2. Construction of a markerless Δ*cadC* deletion mutant

To create a *cadC* markerless deletion we performed the two-step method for integration and excision as previously described [165]. The regions flanking either side of the *cadC* gene were amplified by PCR using primer pairs 3297_LF and 3297_LR to obtain the left flank (LF; 852bp) and 3297_RF and 3297_RR to obtain the right flank (RF; 842bp) regions. Both flanks were used as the template for an overlap PCR reaction. The assembled insert was then cloned into the SpeI site of the vector pX33 generating pMP3297, which was electroporated into *M. smegmatis* mc^2155 background [9,188]. For selection of single recombination events *M. smegmatis* mc^2155 (pMP3297) was grown at 28°C with agitation (200 rpm) to an optical density (OD_{600}) of approximately 0.8. Aliquots of culture were plated onto LBT Gentamycin and grown at 42°C.

Single colonies were selected and grown in LBT at 37°C to an OD_{600} of approximately 0.6. For selection of double recombination events, aliquots of culture were plated onto low-salt (2 g NaCl l^{-1}) LBT plates containing 10% sucrose and incubated at 42°C. For confirmation of the correct double recombination, Southern hybridisation was performed on SphI-digested genomic DNA using the Amersham Gene Images AlkPhos Direct Labelling and Detection System with CDP-Star detection reagent (GE Healthcare) according to manufacturers protocol, generating *M. smegmatis* MP3297 Δ*cadC*.

All primers are listed in Appendix 3.

4.5.3. RNA extraction and library preparation

Cells were lysed by four cycles at 4800 rpm for 30 s in a mini-Beadbeater (Biospec Products), with 30 s on ice between each of the cycles. Total RNA was extracted using TRIzol® reagent (Ambion) according to the manufacturers instructions. DNA was removed by treatment with 3 U RNase-free DNase using the TURBO DNA-free kit (Ambion) according to the manufacturers instructions. The quality of the RNA was checked with the Bioanalyzer (RIN >8) and the concentration was determined using a NanoDrop ND-100 spectrophotometer. Next, depletion of ribosomal RNA and strand-specific library preparation were performed, using the EpiCentre ScriptSEq™ Complete Kit for Bacteria according to manufacturers instructions. Quantification of nucleic acids was performed using a Qubit® Fluorometer to ensure DNA contamination of the samples was less than 10% and libraries were created. Prepared libraries have undergone a quality control using an Agilent 2100 Bioanalyzer: DNA 1000 Labchip, Quant-iT dSDNA HS Assay for quantification and Quant-iT RNA Assay and Quant-iT Protein Assay for percentage contamination check, using an Invitrogen Qubit® Fluorometer. Libraries were run on a Illumina MiSeq 150 cycle Kit_v3 with a single-end (SE) read length of 1x100 and a PhiX control library was also loaded and used as control for the run.

4.5.4. Analysis of RNA-sequencing data

Adapter sequences were removed from raw fastq files using Flexbar [171], and reads shorter than 40 bp were removed. Sequences were then mapped against the *M. smegmatis* genome (GenBank NC_008596.1) using Bowtie2 with default settings. Counts for each gene were calculated with HtSeq-count [215]. Differential expression was performed on normalized read counts using DESeq2 [216]. GO term enrichment was analysed using the Database for Annotation, Visualization and Integrated Discovery v6.7 (DAVID) [217,218] and gene functions were assigned using public databases [175,178]. The RNA sequencing run report is shown in Table 4.3.

4.5.5. Lysine transport assay

To measure lysine transport in *M. smegmatis* mc^2155 wild-type and *M. smegmatis* MP3297 Δ*cadC*, cells were grown in either HdB medium or LBT medium. For measurements in *M. bovis* BCG (Pasteur) and *M. tuberculosis* mc^26206, cells were grown in 7H9 medium supplemented with 0.5% (w/v) glycerol. The pH was adjusted to pH 7 and cells were grown to exponential growth stage unless otherwise stated. Cells were harvested by centrifugation (17.000 *g*, 2 min, RT) and resuspended to a final OD_{600} of 0.7 – 1.2. Lysine transport was measured by adding L-[^{14}C]lysine (specific activity, 320 mCi $mmol^{-1}$, 0.75 µM final concentration) to 200 µL of cell suspension at 37°C for various time intervals up to 120 seconds to observe linear lysine uptake. The reaction was stopped by adding 2 mL of ice-cold 0.1 M LiCl and rapid filtration through a nitrate cellulose filter (pore size 0.45 µm; Millipore), using a vacuum manifold (Millipore) with an applied vacuum of approximately 7.5 psi. Filters were washed once with 2 mL 0.1 M LiCl, dried at 60°C for 1 hour in scintillation vials and then covered in 2mL of scintillation fluid (Amersham). The amount of radioactivity taken up by the cells was determined with a TriCarb 2910 TR Liquid Scintillation Analyzer (PerkinElmer). Rates of lysine uptake were expressed as $nmol.min^{-1}$ $(mg\ protein)^{-1}$ and when necessary corrected for dilution of L-[^{14}C]lysine with cold L-lysine. Protein concentrations were determined as previously described [165].

4.6. Supplementary Data

Table 4.2. Complete list of genes that are differentially expressed in strain MP3297 Δ*cadC* compared to strain mc^{2}155 wild-type

mc^{2}155 locus[a]	Expression ratio[b]	FDR[c]
MSMEG_0004	0.486	4.06E-03
MSMEG_0051	3.620	2.13E-06
MSMEG_0052	2.163	5.51E-05
MSMEG_0062	0.424	2.89E-07
MSMEG_0071	0.476	4.56E-06
MSMEG_0079	0.383	3.58E-06
MSMEG_0080	0.364	7.24E-07
MSMEG_0081	0.351	1.61E-03
MSMEG_0082	0.421	5.52E-05
MSMEG_0091	2.783	6.20E-08
MSMEG_0108	0.256	1.04E-16
MSMEG_0109	0.417	5.49E-06
MSMEG_0134	0.488	3.45E-05
MSMEG_0137	0.476	7.35E-04
MSMEG_0139	0.453	3.05E-07
MSMEG_0144	0.449	1.52E-02
MSMEG_0147	0.482	5.08E-03
MSMEG_0172	2.985	1.35E-04
MSMEG_0194	2.734	2.26E-04
MSMEG_0221	0.492	7.75E-05
MSMEG_0234	0.439	5.34E-09
MSMEG_0278	0.495	1.57E-02
MSMEG_0282	0.492	4.87E-02
MSMEG_0311	2.140	5.93E-04
MSMEG_0315	2.259	4.60E-05
MSMEG_0317	2.300	3.59E-04
MSMEG_0333	0.331	1.14E-04
MSMEG_0334	0.453	1.25E-02
MSMEG_0335	0.331	1.40E-03
MSMEG_0336	0.402	4.70E-03
MSMEG_0337	0.391	1.33E-03
MSMEG_0338	0.421	1.66E-03
MSMEG_0351	0.358	7.11E-06
MSMEG_0382	0.468	3.97E-05
MSMEG_0400	0.460	1.57E-06
MSMEG_0401	0.404	7.34E-07
MSMEG_0403	0.357	3.59E-06
MSMEG_0404	0.346	4.98E-08
MSMEG_0406	2.037	1.13E-03
MSMEG_0413	2.503	1.32E-02
MSMEG_0422	2.331	5.49E-06
MSMEG_0425	0.475	1.37E-02
MSMEG_0429	2.543	1.77E-05
MSMEG_0438	0.497	3.40E-02
MSMEG_0441	0.430	1.38E-07
MSMEG_0443	0.405	1.04E-02
MSMEG_0446	0.258	5.85E-14
MSMEG_0449	0.475	3.57E-03
MSMEG_0460	0.468	5.70E-04
MSMEG_0523	0.345	4.63E-05
MSMEG_0557	0.365	1.35E-06
MSMEG_0558	0.449	1.58E-03
MSMEG_0559	0.375	9.37E-13
MSMEG_0580	0.476	4.63E-05
MSMEG_0639	0.408	6.49E-06
MSMEG_0640	0.471	4.73E-03
MSMEG_0641	0.387	2.55E-06
MSMEG_0642	0.435	2.74E-05
MSMEG_0643	0.497	2.61E-06
MSMEG_0649	2.717	1.73E-03
MSMEG_0669	0.420	4.28E-03
MSMEG_0671	0.402	1.95E-04
MSMEG_0708	2.113	8.35E-03
MSMEG_0750	2.105	2.13E-06
MSMEG_0753	2.258	1.89E-04
MSMEG_0757	2.352	4.66E-03
MSMEG_0759	3.559	5.05E-10
MSMEG_0760	2.245	1.60E-03
MSMEG_0762	0.353	6.66E-07
MSMEG_0765	0.453	3.26E-03
MSMEG_0775	0.491	1.38E-04
MSMEG_0776	2.152	2.98E-02
MSMEG_0789	0.474	3.91E-02
MSMEG_0817	0.467	8.19E-03
MSMEG_0866	2.390	1.01E-08
MSMEG_0882	0.430	5.29E-04
MSMEG_0883	0.447	3.64E-04
MSMEG_0902	0.372	2.94E-03
MSMEG_0909	2.008	1.44E-07
MSMEG_0911	2.750	1.40E-02

MSMEG_0912	2.092	2.64E-04	MSMEG_1474	0.443	1.57E-04
MSMEG_0933	0.395	4.40E-05	MSMEG_1523	0.496	1.61E-06
MSMEG_0934	0.464	4.47E-03	MSMEG_1530	2.375	1.52E-03
MSMEG_0948	2.007	6.85E-03	MSMEG_1611	0.490	1.12E-02
MSMEG_0956	0.442	2.31E-04	MSMEG_1612	2.099	2.75E-02
MSMEG_0968	0.442	4.47E-04	MSMEG_1620	2.205	2.09E-02
MSMEG_1073	0.365	2.07E-06	MSMEG_1629	2.226	2.68E-05
MSMEG_1112	0.489	8.12E-03	MSMEG_1636	2.304	7.55E-03
MSMEG_1119	0.384	4.68E-03	MSMEG_1650	2.674	2.62E-04
MSMEG_1130	2.277	1.50E-04	MSMEG_1651	2.791	1.77E-05
MSMEG_1167	2.096	2.67E-03	MSMEG_1652	5.386	2.44E-10
MSMEG_1169	2.619	2.53E-04	MSMEG_1653	2.130	9.58E-05
MSMEG_1236	2.987	2.70E-05	MSMEG_1677	0.391	1.21E-06
MSMEG_1237	2.121	4.68E-03	MSMEG_1751	2.209	2.06E-04
MSMEG_1253	0.487	9.98E-06	MSMEG_1771	0.424	3.71E-04
MSMEG_1261	0.438	1.16E-04	MSMEG_1784	0.249	1.04E-06
MSMEG_1263	0.472	4.01E-03	MSMEG_1805	0.500	2.36E-02
MSMEG_1264	2.202	7.55E-03	MSMEG_1810	0.467	3.14E-03
MSMEG_1272	2.275	3.27E-04	MSMEG_1823	0.429	3.94E-03
MSMEG_1286	0.442	5.08E-03	MSMEG_1837	2.248	1.50E-07
MSMEG_1301	2.064	2.53E-02	MSMEG_1868	0.323	7.35E-06
MSMEG_1312	2.227	1.30E-03	MSMEG_1883	2.020	7.19E-03
MSMEG_1316	2.800	6.49E-06	MSMEG_1906	2.237	9.21E-05
MSMEG_1328	2.225	1.32E-05	MSMEG_1911	2.407	1.37E-04
MSMEG_1338	0.323	1.58E-09	MSMEG_1917	2.284	1.23E-04
MSMEG_1339	0.445	7.77E-07	MSMEG_1953	2.158	8.54E-03
MSMEG_1364	0.249	1.00E-23	MSMEG_1978	0.383	4.00E-03
MSMEG_1365	0.295	4.49E-09	MSMEG_1984	3.636	2.38E-04
MSMEG_1369	3.646	1.20E-08	MSMEG_1986	8.002	7.20E-14
MSMEG_1379	2.211	3.08E-05	MSMEG_1987	5.277	7.02E-09
MSMEG_1403	2.154	6.11E-05	MSMEG_1999	3.362	5.03E-05
MSMEG_1440	0.430	2.68E-11	MSMEG_2021	2.821	1.49E-04
MSMEG_1441	0.437	1.10E-10	MSMEG_2025	0.450	6.11E-05
MSMEG_1442	0.358	2.30E-19	MSMEG_2045	0.495	3.99E-05
MSMEG_1443	0.321	2.70E-13	MSMEG_2046	0.308	1.98E-05
MSMEG_1444	0.337	1.50E-06	MSMEG_2111	0.479	1.99E-05
MSMEG_1445	0.427	8.64E-06	MSMEG_2142	2.107	8.98E-05
MSMEG_1453	2.131	2.57E-02	MSMEG_2157	2.463	2.81E-03
MSMEG_1465	0.398	1.57E-10	MSMEG_2179	2.063	1.29E-02
MSMEG_1467	0.415	9.16E-08	MSMEG_2182	0.467	4.24E-02
MSMEG_1468	0.344	2.04E-16	MSMEG_2271	2.684	1.01E-08
MSMEG_1469	0.354	6.78E-12	MSMEG_2272	2.607	8.57E-08
MSMEG_1470	0.315	4.70E-16	MSMEG_2273	3.215	7.20E-14
MSMEG_1471	0.386	2.80E-13	MSMEG_2274	2.644	2.30E-08
MSMEG_1472	0.322	7.76E-09	MSMEG_2275	2.017	4.58E-07
MSMEG_1473	0.205	2.72E-26	MSMEG_2342	2.117	2.80E-02

MSMEG_2380	6.208	1.93E-08
MSMEG_2386	2.755	3.88E-03
MSMEG_2387	3.769	1.45E-05
MSMEG_2388	2.408	1.48E-03
MSMEG_2433	0.494	3.57E-02
MSMEG_2437	0.421	2.68E-05
MSMEG_2438	0.478	1.22E-04
MSMEG_2439	2.340	5.79E-05
MSMEG_2544	2.466	5.08E-03
MSMEG_2585	2.064	2.62E-04
MSMEG_2645	2.106	6.80E-04
MSMEG_2661	0.340	6.90E-06
MSMEG_2726	0.417	2.69E-04
MSMEG_2727	0.479	1.13E-04
MSMEG_2728	0.494	9.98E-06
MSMEG_2742	2.078	8.53E-03
MSMEG_2777	0.411	8.64E-06
MSMEG_2780	2.408	1.18E-06
MSMEG_2897	0.478	2.78E-03
MSMEG_2900	2.488	6.36E-06
MSMEG_2939	0.433	6.45E-03
MSMEG_2940	2.788	1.16E-05
MSMEG_3015	2.257	8.35E-04
MSMEG_3032	0.484	3.45E-05
MSMEG_3042	5.566	4.40E-16
MSMEG_3043	6.364	6.53E-33
MSMEG_3044	5.244	4.57E-22
MSMEG_3045	2.719	3.23E-10
MSMEG_3046	2.105	1.11E-05
MSMEG_3058	0.495	9.87E-07
MSMEG_3068	2.088	2.70E-02
MSMEG_3083	0.349	2.39E-04
MSMEG_3157	2.169	1.30E-02
MSMEG_3181	2.109	3.99E-03
MSMEG_3187	2.032	1.45E-05
MSMEG_3216	0.480	8.07E-04
MSMEG_3232	0.495	1.82E-03
MSMEG_3244	0.491	3.15E-04
MSMEG_3248	0.476	6.17E-05
MSMEG_3250	0.401	8.14E-06
MSMEG_3251	0.427	6.05E-04
MSMEG_3264	2.356	1.82E-03
MSMEG_3283	2.866	2.05E-05
MSMEG_3287	2.158	3.25E-05
MSMEG_3297	0.250	1.75E-11
MSMEG_3313	2.529	6.25E-03
MSMEG_3325	2.607	1.22E-04
MSMEG_3326	0.435	7.84E-03
MSMEG_3343	0.485	4.92E-03
MSMEG_3344	2.264	1.01E-06
MSMEG_3363	2.343	1.42E-04
MSMEG_3378	2.003	8.43E-03
MSMEG_3380	2.068	2.51E-03
MSMEG_3411	2.048	2.59E-02
MSMEG_3447	2.434	1.17E-02
MSMEG_3450	0.497	1.17E-02
MSMEG_3457	2.517	9.74E-10
MSMEG_3499	0.454	4.55E-04
MSMEG_3522	2.726	3.25E-05
MSMEG_3531	2.286	6.90E-06
MSMEG_3546	3.863	3.79E-10
MSMEG_3562	0.490	1.85E-03
MSMEG_3582	2.032	1.82E-03
MSMEG_3619	2.004	2.41E-04
MSMEG_3623	0.406	2.77E-03
MSMEG_3637	3.054	5.11E-12
MSMEG_3638	3.241	4.79E-16
MSMEG_3649	0.417	7.55E-05
MSMEG_3650	0.487	1.30E-03
MSMEG_3651	0.397	1.34E-06
MSMEG_3678	2.744	1.03E-03
MSMEG_3680	2.288	1.37E-02
MSMEG_3687	0.491	2.89E-02
MSMEG_3704	2.740	2.76E-05
MSMEG_3759	2.210	1.66E-02
MSMEG_3768	2.173	3.40E-04
MSMEG_3769	0.493	1.94E-03
MSMEG_3792	0.413	5.20E-11
MSMEG_3796	0.489	1.60E-04
MSMEG_3800	2.428	6.68E-04
MSMEG_3828	0.436	2.25E-02
MSMEG_3871	2.275	2.68E-05
MSMEG_3933	0.455	2.02E-02
MSMEG_3935	0.413	4.01E-03
MSMEG_3942	0.408	8.54E-03
MSMEG_3943	0.444	1.21E-02
MSMEG_3946	0.420	1.29E-02
MSMEG_3950	0.388	8.13E-04
MSMEG_3952	0.485	2.37E-02
MSMEG_3961	0.381	2.13E-03
MSMEG_4014	0.425	1.48E-03
MSMEG_4082	0.449	2.38E-03

MSMEG_4083	0.499	3.51E-03	MSMEG_4741	4.278	3.96E-25
MSMEG_4085	0.269	7.64E-11	MSMEG_4753	2.288	5.99E-04
MSMEG_4086	0.322	1.51E-06	MSMEG_4767	3.424	5.45E-07
MSMEG_4087	0.345	7.02E-06	MSMEG_4829	0.467	2.51E-02
MSMEG_4108	0.289	2.63E-04	MSMEG_4854	0.486	9.94E-03
MSMEG_4110	0.330	3.57E-06	MSMEG_4855	0.378	5.03E-07
MSMEG_4111	0.470	1.48E-03	MSMEG_4890	2.505	6.36E-06
MSMEG_4114	0.489	1.15E-04	MSMEG_4891	2.489	4.08E-04
MSMEG_4179	2.000	3.08E-03	MSMEG_4914	0.498	2.04E-04
MSMEG_4182	2.122	5.03E-03	MSMEG_4930	2.222	1.73E-02
MSMEG_4183	2.212	5.08E-04	MSMEG_4933	0.415	5.21E-05
MSMEG_4207	0.493	4.94E-03	MSMEG_4935	0.432	1.32E-04
MSMEG_4215	2.425	2.41E-04	MSMEG_4936	0.456	6.20E-05
MSMEG_4220	0.360	3.95E-13	MSMEG_4937	0.408	7.51E-06
MSMEG_4221	0.441	3.59E-04	MSMEG_4938	0.467	2.64E-07
MSMEG_4242	0.454	8.62E-06	MSMEG_4973	2.212	5.21E-05
MSMEG_4264	2.648	2.41E-04	MSMEG_4985	0.379	3.31E-05
MSMEG_4265	2.251	6.25E-05	MSMEG_4999	0.447	7.51E-06
MSMEG_4270	2.070	6.66E-07	MSMEG_5042	2.228	9.40E-11
MSMEG_4291	2.207	1.50E-02	MSMEG_5052	2.161	1.62E-05
MSMEG_4292	0.304	2.64E-05	MSMEG_5068	2.394	3.24E-11
MSMEG_4305	0.397	3.15E-05	MSMEG_5128	2.001	3.39E-04
MSMEG_4307	0.438	1.77E-05	MSMEG_5129	2.041	1.23E-06
MSMEG_4315	2.031	4.82E-03	MSMEG_5136	0.415	3.13E-03
MSMEG_4349	2.211	6.92E-05	MSMEG_5148	2.633	8.63E-03
MSMEG_4365	2.305	3.18E-05	MSMEG_5149	2.445	3.50E-06
MSMEG_4450	2.342	9.69E-04	MSMEG_5170	0.422	5.80E-07
MSMEG_4474	0.439	1.15E-06	MSMEG_5188	0.431	3.65E-03
MSMEG_4486	2.247	1.21E-03	MSMEG_5198	0.493	7.02E-06
MSMEG_4492	2.750	1.08E-06	MSMEG_5243	0.287	5.72E-04
MSMEG_4497	2.372	1.14E-06	MSMEG_5244	0.437	1.49E-03
MSMEG_4532	2.034	8.34E-03	MSMEG_5245	0.350	1.87E-03
MSMEG_4567	0.494	1.13E-02	MSMEG_5246	0.264	2.29E-06
MSMEG_4569	0.465	1.61E-02	MSMEG_5255	2.015	5.27E-04
MSMEG_4582	0.311	5.84E-10	MSMEG_5257	2.148	1.85E-05
MSMEG_4583	0.396	1.17E-05	MSMEG_5263	0.491	5.49E-06
MSMEG_4619	2.053	2.70E-02	MSMEG_5274	2.160	3.73E-05
MSMEG_4649	2.089	1.82E-02	MSMEG_5280	4.637	3.19E-06
MSMEG_4650	2.285	6.18E-04	MSMEG_5286	2.662	4.95E-08
MSMEG_4651	2.764	1.85E-05	MSMEG_5298	2.930	4.47E-05
MSMEG_4664	2.051	3.43E-03	MSMEG_5316	0.455	2.05E-02
MSMEG_4678	3.014	3.12E-06	MSMEG_5318	0.424	3.04E-04
MSMEG_4729	0.434	5.61E-03	MSMEG_5322	2.048	2.25E-02
MSMEG_4730	0.289	2.64E-05	MSMEG_5328	0.274	1.17E-15
MSMEG_4731	0.348	1.06E-04	MSMEG_5339	2.435	2.60E-04
MSMEG_4732	0.341	2.41E-04	MSMEG_5349	2.604	3.13E-03

MSMEG_5355	2.092	1.04E-02	MSMEG_6138	0.288	5.05E-10
MSMEG_5359	0.464	3.35E-02	MSMEG_6143	2.939	1.64E-06
MSMEG_5400	0.458	1.61E-02	MSMEG_6197	2.689	7.82E-04
MSMEG_5425	0.355	3.52E-04	MSMEG_6198	2.915	7.95E-11
MSMEG_5452	2.274	1.89E-05	MSMEG_6209	2.953	1.49E-04
MSMEG_5487	0.441	1.28E-06	MSMEG_6213	0.385	3.35E-05
MSMEG_5489	2.252	7.01E-05	MSMEG_6223	2.019	4.71E-03
MSMEG_5526	2.940	1.25E-04	MSMEG_6233	0.422	6.00E-03
MSMEG_5534	2.023	1.65E-05	MSMEG_6234	0.409	4.52E-04
MSMEG_5540	0.346	4.90E-05	MSMEG_6253	3.135	4.15E-06
MSMEG_5549	2.045	1.12E-02	MSMEG_6292	0.487	5.51E-05
MSMEG_5556	4.056	7.82E-07	MSMEG_6293	0.466	5.40E-03
MSMEG_5557	2.277	1.15E-03	MSMEG_6312	0.458	1.87E-06
MSMEG_5570	0.364	1.90E-04	MSMEG_6321	0.360	1.16E-03
MSMEG_5606	0.357	1.30E-03	MSMEG_6322	0.231	2.64E-07
MSMEG_5641	2.212	8.54E-03	MSMEG_6345	0.313	5.84E-10
MSMEG_5661	0.476	5.44E-03	MSMEG_6346	0.454	5.98E-04
MSMEG_5680	2.192	4.42E-07	MSMEG_6347	0.420	1.75E-02
MSMEG_5686	2.170	1.11E-02	MSMEG_6350	2.208	1.30E-03
MSMEG_5687	2.805	9.65E-05	MSMEG_6355	0.451	1.29E-02
MSMEG_5705	2.654	6.27E-08	MSMEG_6375	2.312	1.18E-02
MSMEG_5717	0.496	4.43E-05	MSMEG_6376	2.574	9.31E-04
MSMEG_5733	0.332	9.21E-05	MSMEG_6378	2.836	2.01E-04
MSMEG_5734	0.467	5.93E-04	MSMEG_6385	2.214	6.11E-05
MSMEG_5747	0.458	3.61E-02	MSMEG_6392	0.417	3.74E-07
MSMEG_5774	0.474	2.64E-04	MSMEG_6400	0.474	1.48E-04
MSMEG_5781	2.157	3.39E-04	MSMEG_6406	0.427	1.75E-04
MSMEG_5782	2.655	5.37E-04	MSMEG_6414	2.299	9.00E-06
MSMEG_5798	2.172	4.93E-09	MSMEG_6422	2.337	7.86E-04
MSMEG_5821	0.472	2.90E-03	MSMEG_6451	2.272	2.45E-02
MSMEG_5823	0.461	1.03E-05	MSMEG_6499	0.410	6.05E-04
MSMEG_5826	0.239	1.61E-06	MSMEG_6500	0.430	3.08E-03
MSMEG_5834	0.422	3.76E-04	MSMEG_6501	0.327	9.85E-04
MSMEG_5835	0.494	9.00E-06	MSMEG_6563	2.164	2.98E-02
MSMEG_5852	2.433	3.19E-09	MSMEG_6567	0.296	9.44E-13
MSMEG_5868	4.010	3.27E-07	MSMEG_6569	0.274	1.71E-08
MSMEG_5897	0.426	1.40E-10	MSMEG_6570	0.268	6.61E-23
MSMEG_5904	2.048	2.74E-05	MSMEG_6571	0.338	1.23E-09
MSMEG_5912	2.727	1.31E-06	MSMEG_6580	0.482	1.29E-02
MSMEG_5924	2.179	3.42E-05	MSMEG_6581	0.426	2.63E-04
MSMEG_5966	0.405	1.42E-04	MSMEG_6583	0.322	2.72E-08
MSMEG_5970	0.415	9.13E-03	MSMEG_6584	0.393	2.58E-07
MSMEG_5971	0.390	1.30E-02	MSMEG_6585	0.428	1.02E-04
MSMEG_5978	0.301	1.73E-05	MSMEG_6587	2.579	2.79E-09
MSMEG_6009	0.479	3.02E-03	MSMEG_6590	2.134	4.55E-04
MSMEG_6137	0.485	2.62E-04	MSMEG_6607	0.458	5.97E-03

MSMEG_6608	0.339	1.76E-03
MSMEG_6609	0.321	8.12E-04
MSMEG_6610	0.344	9.86E-03
MSMEG_6612	0.384	8.56E-05
MSMEG_6615	0.417	6.29E-03
MSMEG_6627	0.454	3.21E-03
MSMEG_6655	0.485	2.39E-02
MSMEG_6666	0.389	7.55E-05
MSMEG_6699	0.422	6.19E-04
MSMEG_6727	0.496	4.79E-03
MSMEG_6728	0.246	2.29E-06
MSMEG_6741	0.441	9.89E-04
MSMEG_6761	0.494	3.32E-03
MSMEG_6767	0.440	1.77E-03
MSMEG_6769	0.274	9.98E-06
MSMEG_6785	0.412	4.81E-04
MSMEG_6787	0.434	5.33E-04
MSMEG_6793	2.014	3.72E-03
MSMEG_6801	0.439	9.97E-03
MSMEG_6821	2.507	5.57E-05
MSMEG_6822	0.357	5.81E-05
MSMEG_6888	0.446	1.30E-02
MSMEG_6889	0.293	8.95E-06
MSMEG_6894	0.484	1.97E-04
MSMEG_6895	0.349	2.34E-12
MSMEG_6896	0.456	2.26E-05
MSMEG_6911	0.415	7.84E-03
MSMEG_6939	0.363	7.14E-12
MSMEG_6940	0.474	4.19E-04
MSMEG_6941	0.433	1.20E-10
MSMEG_6942	0.479	5.21E-05
MSMEG_6944	0.472	7.81E-05

[a] Locus number of gene in M. smegmatis mc^2155

[b] Mean gene expression ratio of three biological replicates

[c] *P*-values of gene expression ratio from three biological replicates were corrected for multiple testing using the Benjamini and Hochberg False Discovery Rate

Table 4.3. RNA-sequencing run summary report

Level	Yield Total (G)[a]	Aligned (%)[b]	% Perfect [Num Cycles][c]	% <=3 errors [Num cycles][d]	Error rate (%)[e]	Intensity Cycle 1[f]	% Intensity Cycle 20[g]	% >= Q30[h]
Read 1[i]	1.9	1.06	89.6 [100]	99.1 [100]	0.28	263	80.6	97.0
Read 2[k]	0.1	0	0 [5]	0 [5]	0	696	0	93.7
Total	2.0	1.06	89.6	99.1	0.28	479	80.6	96.8

Level	Density[l]	Clusters[m]	Reads (M)[n]	Reads PF (M)[o]	% >= Q30	Error Rate 100 cycle (%)[p]
Read 1	875 ± 24	93.1 ± 0.3	20.67	19.25	97.0	0.28±0.04
Read 2	875 ± 24	93.1 ± 0.3	20.67	19.25	93.7	0.00±0.00

[a] The number of bases sequenced, in gigabases

[b] The percentage of the sample that aligned to the PhiX genome

[c] The percentage of bases in reads that align perfectly, as determined by a spike in PhiX control sample, at the cycle indicated in brackets

[d] The percentage of bases in reads that align with 3 errors or less, as determined by a spike in PhiX control sample, at the cycle indicated in brackets

[e] The calculated error rate, as determined by the PhiX alignment

[f] The average of the four intensities measured at the first cycle averaged over filtered clusters

[g] The corresponding intensity statistic at cycle 20 as a percentage of that at the first cycle

[h] The percentage of bases with a quality score of 30 or higher, respectively, after the 25th cycle

[i] First sequence read

[k] Second sequence read

[l] The density of clusters (in thousands per mm^2) detected by image analysis

[m] The percentage of clusters passing filtering

[n] The number of clusters (in millions)

[o] The number of clusters (in millions) passing filtering

[p] The calculated error rate for cycles 1-100, as determined by the PhiX alignment

Chapter 5

Conclusions and Future Perspectives

5.1. Understanding Nitrogen Metabolism in Mycobacteria

Nitrogen is an essential component of all living cells and prokaryotes have developed elaborate mechanisms for the uptake, assimilation and metabolism of nitrogen. Despite the importance of nitrogen for cell growth and maintenance of intracellular metabolism, there have been only a few studies on nitrogen metabolism of free-living saprophytes and pathogenic species of mycobacteria (Figure 5.1). The primary goals of this thesis were to address the following questions: “How do mycobacteria regulate nitrogen metabolism?“, “Why does *M. smegmatis* encode for two global nitrogen regulators, GlnR and AmtR?“ and “How do mycobacteria adapt to nitrogen limitation in their environment?“. The work presented in this thesis has contributed new knowledge and insight into addressing these questions in the fast-growing saprophytic species *Mycobacterium smegmatis* (Figure 5.2). We show a wider transcriptional response and greater metabolic versatility of *M. smegmatis* in response to nitrogen limitation compared to other actinobacteria. We have analysed the molecular mechanism of AmtR regulation in *M. smegmatis* and show the essentiality of CadC for cell growth and survival, suggesting CadC as regulator of lysine and DAP biosynthesis in *M. smegmatis*.

M. smegmatis appears to have a broader network of regulatory systems that together mediate the response to nitrogen limitation and enable *M. smegmatis* to adapt its nitrogen metabolism to rapidly changing environments. These mechanisms are likely to be very important in soil ecosystems that are known for their high variability and competitiveness.

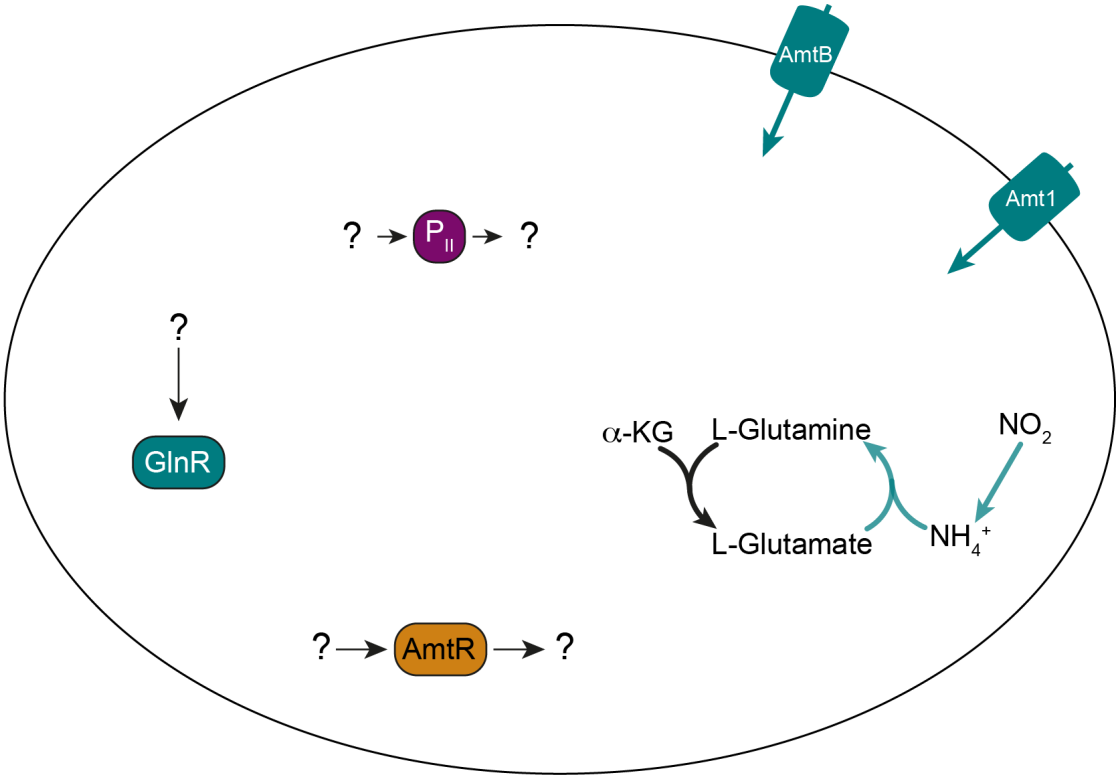


Figure 5.1. Knowledge about nitrogen metabolism and its regulation in *M. smegmatis* at the start of this study. In 2012, our knowledge about nitrogen metabolism and regulatory mechanisms involved in the adaptation to nitrogen limitation was limited to a few studies. Three potential global nitrogen regulatory proteins, GlnR (cyan box), AmtR (orange box) and a P_{II} protein (purple box) are encoded in the genome of *M. smegmatis.* It was shown that GlnR is the transcriptional regulator of *amt1*, *amtB*, *nirBD* and *glnA* encoding for two ammonium transporters, nitrite reductase and glutamine synthetase. Catalytic mechanisms and transport mechanisms driven by these enzymes are indicated by cyan arrows and cyan cylinders. The function of AmtR and the P_{II} protein was unknown. Adapted from [71].

M. smegmatis reveals a versatile transcriptomic response to nitrogen limitation

In the current study, the nitrogen regulated transcriptome of *M. smegmatis* was analysed using a nitrogen-limited continuous culture. Dissection of the nitrogen-regulated transcriptome in the absence of growth rate effects, using RNA-sequencing, revealed multiple adaptive mechanisms on a transcriptional level in metabolic pathways of *M. smegmatis* in response to nitrogen limitation. Expression of genes involved in nitrogen scavenging, nucleotide catabolism and amino acid metabolism were greatly affected and revealed a far more complex response to nitrogen limitation than previously reported for other actinobacteria (e.g. *C. glutamicum* and *S. coelicolor*) [141,144,153]. These results highlight the capability of *M. smegmatis* to adapt its metabolism to nitrogen-depleted growth conditions and to transport various nitrogen sources followed by their concomitant feeding into nitrogen metabolic processes (Figure 5.2). Comparison of our results to a concurrently performed study by Williams *et al.* that defined the genome-wide transcription profile of *M. smegmatis* in response to nitrogen limitation in batch culture highlighted differences in the transcriptomic response of *M. smegmatis* to nitrogen limitation [142]. Nutrient depletion studies are hard to perform in batch culture, because nutrient run-out will have an immediate effect on the growth rate and therefore might pose difficulties to assign transcriptomic changes to the exact stimulon (nutrient or growth rate). We confirmed that continuous culture can indeed reduce the number of differentially expressed genes by 83% that are not associated with nitrogen metabolism and its regulation but rather respond to changes in growth rate.

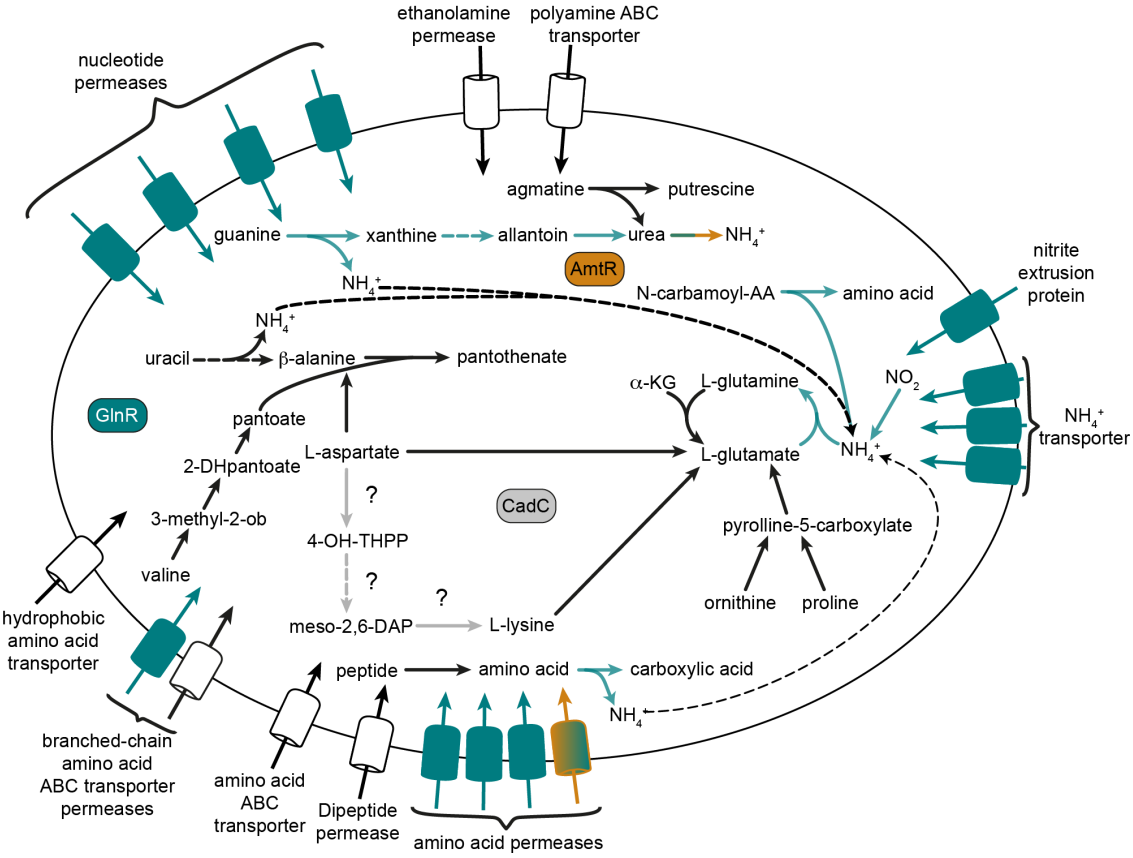


Figure 5.2. Summary of nitrogen metabolism and its regulation in *M. smegmatis*. The nitrogen regulated transcriptome was defined in continuous culture (this work) and in batch culture [142] and showed major changes in expression of genes involved in nitrogen scavenging, nucleotide catabolism, ammonium recovery and amino acid metabolism. The GlnR regulon (cyan box) was characterised by Jenkins *et al.* [43]. Both the catalytic and the transport mechanisms driven by enzymes that are subject to GlnR-regulated gene expression are indicated by cyan arrows and cyan cylinders. The GlnR regulon includes several nitrogen uptake systems, nucleotide degradation and the urea carboxylase/allophanate hydrolase enzymes. The AmtR regulon (orange box) was characterised in this work. The AmtR regulon includes an amino acid permease and the urea carboxylase/allophanate hydrolase enzymes. Both the catalytic and the transport mechanisms driven by enzymes that are subject to AmtR-regulated gene expression are indicated by orange arrows and orange cylinders.

The nitrogen regulated transcriptome of pathogenic Mycobacteria

The development of a nitrogen-limited continuous culture described in this thesis allowed a clear dissection of the nitrogen regulated transcriptome of *M. smegmatis*. However, elevated gene expression does not always correlate to an increase in protein levels. Therefore it is essential to complete the picture of the molecular response to nitrogen limitation in *M. smegmatis* with metabolomic analyses. It will be interesting to define the metabolome of nitrogen-depleted and nitrogen-replete cells of *M. smegmatis* and investigate whether changes in the transcriptome overlap with changes in the metabolome.

The next milestone in the field of nitrogen metabolism in mycobacteria will be to define the nitrogen regulated transcriptome in members of the *M. tuberculosis* complex. Recent findings start to explore nitrogen metabolism as a potential area for antimicrobial development [16,219-222]. The glutamine synthetase enzyme GS-I is a key player in ammonium assimilation and in *M. tuberculosis* the majority of this enzyme is secreted into the culture medium [92,93]. Deletion of the glutamine synthetase encoding gene *glnA1* resulted in a glutamine auxotrophy and attenuation in both macrophage infection and guinea pig models, indicating its essentiality for ammonium assimilation in *M. tuberculosis* [220]. In recent studies several potent GS-I enzyme inhibitors with *in vivo* antibacterial activities were reported [221,222].

Further potential targets might be found in the metabolism of nitrate, nitrite and amino acids. Nitrate respiration has been suggested to play a crucial role in the pathophysiology of *M. tuberculosis* [44,223-225]. Sohaskey *et al.* reported that nitrate enhanced survival of *M. tuberculosis* during inhibiton of aerobic respiration [224]. Furthermore, nitrite generation is directly associated to a decrease in growth and ATP consumption as well as several alterations in the transcriptome of *M. tuberculosis* [225]. Targeting these mechanisms could disturb essential physiological processes influencing the success of mycobacterial defence against host stresses [226].

The catabolism of the amino acids asparagine and aspartate is important for virulence of *M. tuberculosis* [18,19]. Virulence of an aspartate transporter mutant was impaired *in vivo* pointing to the necessity for aspartate

uptake in *M. tuberculosis* [18]. The asparaginase AnsA is crucial for mediating resistance to acid stress in *M. tuberculosis* through asparagine hydrolysis and ammonia release and a mutant was attenuated in macrophages in mice [19]. These findings are particularly interesting considering the results from our work showing several alterations in the amino acid metabolism of *M. smegmatis*, suggesting that amino acids and their derivatives might exhibit additional functions beyond being used as building blocks in protein biosynthesis, as previously reported for proline [165].

Initial work has been done to grow members of the *M. tuberculosis* complex (*M. bovis* BCG (Pasteur) and *M. tuberculosis* H37Ra) in nitrogen-depleted continuous cultures to analyze their transcriptional response to nitrogen limitation. This will enable us to establish a clear view on how *M. tuberculosis* adapts to nitrogen limitation, what metabolic pathways are affected in this adaptation mechanism and together with defining the metabolome in nitrogen-depleted cells of *M. tuberculosis* this will help us to identify potential drug targets in nitrogen metabolism.

The genome of *M. tuberculosis* harbors two GlnR homologs and further work should concentrate on the molecular characterisation and the physiological role of these two potential nitrogen regulatory systems. It will be interesting to find out to what extend *M. tuberculosis* GlnR is involved in the regulation of nitrogen uptake considering its highly reduced number of e.g. amino acid uptake systems and ammonium uptake systems compared to *M. smegmatis*.

5.2. Molecular Adaptation of Mycobacterial Cells to Nitrogen Limitation

Our RNA-seq analysis of the nitrogen-regulated transcriptome showed differential expression of multiple transcriptional regulators and several putative non-coding RNAs. It will be a challenging task for the future to identify the role of the single differentially expressed regulatory proteins and to define a complete regulatory network. Several reports have described transcriptional regulators that link nitrogen metabolism to carbon and phosphate metabolism or to various environmental stress conditions (e.g. nutrient limitation, pH variation, oxidative stress, osmotic stress) highlighting the importance of amino acids in these processes that can act as nutrient but also fulfill a secondary function in a bacterial cell. This would require a tight regulation of the metabolism of these amino acids, pointing to sophisticated regulatory mechanisms to prevent depletion of these where deemed necessary.

Importance of CadC for lysine metabolism in Mycobacteria

In a whole-genome RNA-seq analysis we showed elevated expression of genes involved in the metabolism of lysine, aspartate and glutamate in a Δ*cadC* mutant including lysine biosynthesis from aspartate, a pathway that comprises DAP biosynthesis (Figure 5.2). In LBT medium we observed a strong decrease in OD_{600} and CFU measurements, suggesting that cell lysis occured during exponential growth phase, while minimal medium growth experiments demonstrated that a *M. smegmatis* strain MP3297 Δ*cadC* mutant failed to grow. Based on previous reports by Pavelka *et al.*, we hypothesize that the observed cell lysis is caused by DAP deprivation [199]. However, preliminary results demonstrated that external addition of DAP or lysine had only a moderate effect on cell lysis. We confirmed the presence of a functional high-affinity lysine uptake system in *M. smegmatis*. The lysine uptake rate is significantly increased during growth in minimal medium, suggesting that *M. smegmatis* prefers uptake of lysine to generate DAP rather than to run DAP biosynthesis from aspartate, while the requirement for lysine uptake in rich

media like LBT is minor. However, members of the *M. tuberculosis* complex failed to transport lysine, suggesting that they solely rely on intracellular lysine biosythesis mechanisms to replenish their intracellular lysine pool.

Further work is now required to understand the exact mechanism of CadC-dependent regulation of these processes. Alterations in the lysine biosynthetis pathway would ultimately affect intracellular DAP levels in a mycobacterial cell and have an immediate effect on the cellular integrity because the bacteria would lose their ability to synthesise peptidoglycan (Figure 5.2). Initial experiments should focus on genetic complementation with a non-chromosomal copy of *cadC* and the identification of growth conditions that restore growth (minimal medium) or prevent cell lysis (rich medium) of a *M. smegmatis* Δ*cadC* deletion mutant. It would also be necessary to describe the expression profile of the *cadC* gene during growth in different media. Our transcriptomic analysis revealed only small changes in expression of genes involved in the metabolism of lysine, aspartate and glutamate. However, these changes might have a major effect on the metabolite level, which would need to be investigated in a metabolomic analysis.

The ToxR-like protein CadC of *E. coli* is a membrane-bound orphan transcriptional regulator that combines sensory mechanisms, signal transduction and DNA-binding functions. However, no structural data are available for the mycobacterial homolog of CadC. Preliminary work was done to compare the CadC proteins of *M. smegmatis* and *E. coli* showing that the membrane-located domain is likely to be missing in *M. smegmatis*. Expression and purification of CadC would be the starting point for many biochemical, biophysical and structural analyses to identify its molecular mechanism. We have created an expression construct to express CadC using the expression vector pQE80-L for use in a heterologous expression system in *E. coli*. *In vitro* DNA binding assays with purified CadC protein and promoter regions of genes that were differentially expressed in our genome-wide transcriptome analysis would be important to confirm a direct regulatory effect of CadC on transcription of genes involved in lysine metabolism and would further allow the identification of the CadC-binding site.

Non-coding RNAs as regulatory mechanism in response to changing nitrogen levels

Further work should focus on the confirmation of the regulatory effect of the identified antisense RNAs and small RNAs on transcript levels of their target genes. To our knowledge there is no report of regulation of nitrogen metabolism by non-coding RNAs in actinobacteria. This field is completely unexplored and certainly harbors many interesting and important findings. In our RNA-seq analysis of the nitrogen regulated transcriptome we not only identified antisense RNAs and small RNAs of genes involved in nitrogen metabolism, but also of transcriptional regulators and genes that encode for proteins involved in carbon and fatty acid metabolism. This would be the first report of regulatory non-coding RNAs that are mediating the response to nitrogen limitation in mycobacteria.

In our RNA-seq analysis defining the nitrogen-regulated transcriptome, we confirmed that AmtR was expressed only at low levels under both nitrogen-depleted and nitrogen-replete conditions. Our data further suggested that AmtR was subject to post-transcriptional regulation by an antisense transcript of the AmtR-encoding gene. This antisense transcript *msmeg_4300* revealed a strong downregulation (127-fold) under nitrogen limitation with a concurrent downregulation of *amtR* by 1.7-fold under the same conditions. Although an immediate effect of this antisense transcript needs to be verified, this observation would explain why previous studies failed to observe significant transcriptional changes in *amtR* gene expression.

5.3. Regulation and Physiological Properties of Urea Degradation in *M. smegmatis*

In a genome-wide transcriptomic analysis we were able to determine the AmtR regulon and showed that the regulatory effect of AmtR is restricted to eleven genes, including genes involved in uptake, assimilation and metabolism of nitrogen (Figure 5.2). AmtR regulates a pathway of urea degradation by an urea carbolyase (Figure 5.2). In a physiological characterisation we showed that growth of a Δ*amtR* mutant is uncoupled from carbon source. Using a combination of biochemical and biophysical analyses we further determined the AmtR binding site and identified two ligands of mycobacterial AmtR. Based on our results we proposed a model for the dual transcriptional regulation of the urea carboxylase/allophanate hydrolase-encoding operon by AmtR and GlnR. We propose fine-tuning transcriptional control where AmtR functions as transcriptional repressor of this operon under nitrogen surplus by sensing xanthine and/or allantoin and GlnR as transcriptional activator of this operon under nitrogen limitation.

Further work is now required to validate our proposed mechanism of the dual regulation of the urea amidolyase-encoding operon by GlnR and AmtR. We have shown that *M. smegmatis* can metabolise urea, however, to understand the physiological necessity of two urea degradation pathways in *M. smegmatis*, both urease and urea amidolyase would need to be expressed and purified to allow their biochemical characterisation and determination of kinetic parameters (e.g. K_m and V_{max}). To assess the importance of each enzyme for urea degradation, deletion mutants of urea-degradation enzymes would need to be created in *M. smegmatis*. Urea degradation is under-studied in actinobacteria and we lack general knowledge about the physiological importance of the urea amidolyase in prokaryotes. Therefore, further biochemical characterisation of urea-degrading enzymes is absolutely required to strengthen our understanding of urea metabolism, an ecologically important substrate, in environmental actinobacteria.

Appendices

Appendix 1. Bacterial strains used in this study. JR made deletion mutant JR258 Δ*amtR* prior to this work. The strains *M. smegmatis* mc^{2}155, *M. bovis* BCG (Pasteur) and *M. tuberculosis* mc^{2}6206 were kindly provided by Bill Jacobs. MP made deletion mutant MP3297 Δ*cadC*.

Strain	Description	Reference
Escherichia coli		
DH10B	F$^-$ *mcr*A Δ (*mrr-hsd*RMS-*mcr*BC) Φ80, *lac*Z ΔM15 Δ*lac*X74 *deo*R *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L *end*A1 *nup*G	[227]
Mycobacterium smegmatis		
mc^{2}155	electrocompetent wild-type strain of *M. smegmatis*	[228]
JR258 Δ*amtR*	mc^{2}155 with markerless deletion of *msmeg_4300*	This study
MP3297 Δ*cadC*	mc^{2}155 with markerless deletion of *msmeg_3297*	This study
Mycobacterium bovis		
Bacillus Calmette-Guérin (Pasteur)	wild-type strain of *M. bovis* BCG Pasteur	[229]
Mycobacterium tuberculosis		
mc^{2}6206	Δ*leuCD* Δ*panCD*	[230]

Appendix 2. Plasmids used in this study. JR made deletion construct pJR500 and complementation vector pMind-AmtR. MP made deletion construct pMP3297.

Plasmid	Description	Reference
pX33	pPR23 carrying a constitutive *xylE* marker; Gm^r Sac^s ts	[9,188]
pMP3297	pX33 carrying a truncated copy of the *cadC* gene	This study
pJR500	pX33 carrying a truncated copy of the *amtR* gene	This study
pMind	Tetracycline inducible expression vector; Km^r Hyg^r	[189]
pMind-AmtR	pMind carrying the *amtR* gene	This study

Appendix 3. Primers used in this study. JR designed primers for the determination of the genetic organization of *amtR*, for the *amtR* deletion construct, the *amtR* complementation construct and primers for amplification of the *pho*H2 gene. MP designed primers for quantitative RT-PCR, Differential Scanning Fluorimetry and the *cadC* deletion construct.

Primer	Sequence	Description
AmtR LF F	AAATTTACTAGTCTAGGCGAATTCGGGTCTGC	genetic organization of *amtR* gene and *amtR* deletion construct; SpeI cut site
AmtR inner R	CTGTTCACCACCCACGGGTAC	genetic organization of *amtR* gene
AmtR inner F	CTCTCCGATCACCCAGGGTTG	genetic organization of *amtR* gene
AmtR RF R	AAATTTACTAGTCTGTTCGTGTCGACCATCGG	genetic organization of *amtR* gene and *amtR* deletion construct; SpeI cut site
AmtR LF o/lap R	CGGGTACCAACCCTGGGTGATCGGAGAG	*amtR* deletion construct
AmtR RF o/lap F	AGGGTTGGTACCCGTGGGTGGTGAACAG	*amtR* deletion construct
AmtR pMind F	AAATTTACTAGTAGACCGGCACTCTAGCG	complementation; SpeI cut site
AmtR pMind RBS R	AAATTTGGATCCGGAGGAATAATGACGACCAC CTCCGGCCGT	complementation; BamHI cut site
msmeg_0819_left	ATGGACGAAGACGAGATGCT	quantitative RT-PCR
msmeg_0819_right	TCGAGTTTCTGCGTGTACCA	quantitative RT-PCR
msmeg_2184_left	GCACGTTGGCAATTCAGAC	quantitative RT-PCR
msmeg_2184_right	ACGCCAGGTAGATCAGGATG	quantitative RT-PCR
msmeg_2187_left	ATCACCTCGCACTGGAACAT	quantitative RT-PCR
msmeg_2187_right	GAGTAGGCGTTGTCGTGGAT	quantitative RT-PCR
msmeg_2189_left	AATACATCTCGGGCGGTTC	quantitative RT-PCR
msmeg_2189_right	TCGTCACGCAGTCATAGCTC	quantitative RT-PCR
msmeg_3194_left	CCATCAACGGCACAGATGTA	quantitative RT-PCR
msmeg_3194_right	GAGCAGAAGTGGCAGTCCTC	quantitative RT-PCR
msmeg_5594_left	CAACACCGACATCCATGAAC	quantitative RT-PCR
msmeg_5594_right	ACTGACGAACAGCGGAATGT	quantitative RT-PCR
msmeg_6272_left	TACAGCCTGCTCGACATCAT	quantitative RT-PCR
msmeg_6272_right	CGACCGGTACAGATCCTCAC	quantitative RT-PCR
MSMEG_2184_F	GGGCATACCGGCGATTG	EMSA[a] & DSF[b]
MSMEG-2184_R	CGCTGCTGGTCATGGGT	EMSA & DSF
phoH2_fwd	ATGGCTAGCGACCTGCTCTGCTGTC	EMSA
phoH2_rev	TCAGTCACGTGGCGCTCCTAGGGGA	EMSA

3297_LF	TTACTAGTATGATCCGCGAGGAGTTC	*cadC* deletion construct; SpeI cut site
3297_LR	TGAGCATGAACACCTGCGGTTCCAC	*cadC* deletion construct
3297_RF	CCACAAGATGCTCATCCCGCAGTCG	*cadC* deletion construct
3297_RR	TTACTAGTTCACAGGCCCAGTTCGCT	*cadC* deletion construct; SpeI cut site

Underlined: enzyme cut site

[a] Electrophoretic Mobility Shift Assay

[b] Differential Scanning Fluorimetry

References

1. Hartmans S, Bont JAM, Stackebrandt E (2006) The Genus *Mycobacterium*– Nonmedical. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. The Prokaryotes: Archaea. Bacteria: *Firmicutes*, *Actinomycetes*. 3rd edition. Springer, New York. 889–918.

2. Cook GM, Berney M, Gebhard S, Heinemann M, Cox RA, et al. (2009) Physiology of Mycobacteria. Adv Microb Physiol. 55: 81–182.

3. Greening C, Berney M, Hards K, Cook GM, Conrad R (2014) A soil actinobacterium scavenges atmospheric H_2 using two membrane-associated, oxygen-dependent [NiFe]-hydrogenases. Proc Natl Acad Sci USA 111: 4257–4261.

4. Greening C, Villas-Bôas SG, Robson JR, Berney M, Cook GM (2014) The growth and survival of *Mycobacterium smegmatis* is enhanced by co-metabolism of atmospheric H_2. PLoS ONE 9: e103034. doi:10.1371/journal.pone.0103034.

5. Pecsi I, Hards K, Ekanayaka N, Berney M, Hartman T, et al. (2014) Essentiality of succinate dehydrogenase in *Mycobacterium smegmatis* and its role in the generation of the membrane potential under hypoxia. mBio 5: e01093-14. doi:10.1128/mBio.01093-14.

6. Niederweis M (2008) Nutrient acquisition by mycobacteria. Microbiology 154: 679–692.

7. Braibant M, Lefèvre P, de Wit L, Peirs P, Ooms J, et al. (1996) A *Mycobacterium tuberculosis* gene cluster encoding proteins of a phosphate transporter homologous to the *Escherichia coli* Pst system. Gene 176: 171–176.

8. Wooff E, Michell SLI, Gordon SV, Chambers MA, Bardarov S, et al. (2002) Functional genomics reveals the sole sulphate transporter of the *Mycobacterium tuberculosis* complex and its relevance to the acquisition of sulphur in vivo. Mol Microbiol 43: 653–663.

9. Gebhard S, Tran SL, Cook GM (2006) The Phn system of *Mycobacterium smegmatis*: a second high-affinity ABC-transporter for phosphate. Microbiology 152: 3453–3465.

10. Titgemeyer F, Amon J, Parche S, Mahfoud M, Bail J, et al. (2007) A genomic view of sugar transport in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. J Bacteriol 189: 5903–5915.

11. Amon J, Titgemeyer F, Burkovski A (2009) A genomic view on nitrogen metabolism and nitrogen control in mycobacteria. J Mol Microbiol Biotechnol 17: 20–29.

12. Paula S, Volkov AG, Van Hoek AN, Haines TH, Deamer DW (1996) Permeation of protons, potassium ions, and small polar molecules through phospholipid bilayers as a function of membrane thickness. Biophys J 70: 339–348.

13. Beste D, Peters J, Hooper T, Avignone-Rossa C, Bushell ME, et al. (2005) Compiling a molecular inventory for *Mycobacterium bovis* BCG at two growth rates: evidence for growth rate-mediated regulation of ribosome biosynthesis and lipid metabolism. J Bacteriol 187: 1677–1684.

14. Bacon J, Hatch KA (2008) Continuous culture of mycobacteria. In: Parish T and Brown AC, editors. Mycobacteria Protocols. 2nd edition. Springer, New York.153–171.

15. Berney M, Cook GM (2010) Unique flexibility in energy metabolism allows mycobacteria to combat starvation and hypoxia. PLoS ONE 5: e8614. doi:10.1371/journal.pone.0008614.

16. Gouzy A, Poquet Y, Neyrolles O (2014) Nitrogen metabolism in *Mycobacterium tuberculosis* physiology and virulence. Nat Rev Microbiol 12: 729–737.

17. Peirs P, Lefèvre P, Boarbi S, Wang X-M, Denis O, et al. (2005) *Mycobacterium tuberculosis* with disruption in genes encoding the phosphate binding proteins PstS1 and PstS2 is deficient in phosphate uptake and demonstrates reduced *in vivo* virulence. Infect Immun 73: 1898–1902.

18. Gouzy A, Larrouy-Maumus G, Wu T-D, Peixoto A, Levillain F, et al. (2013) *Mycobacterium tuberculosis* nitrogen assimilation and host colonization require aspartate. Nat Chem Biol 9: 674–676.

19. Gouzy A, Larrouy-Maumus G, Bottai D, Levillain F, Dumas A, et al. (2014) *Mycobacterium tuberculosis* exploits asparagine to assimilate nitrogen and resist acid stress during infection. PLoS Pathog 10: e1003928. doi:10.1371/journal.ppat.1003928.

20. Walter B, Küspert M, Ansorge D, Krämer R, Burkovski A (2008) Dissection of ammonium uptake systems in *Corynebacterium glutamicum*: mechanism of action and energetics of AmtA and AmtB. J Bacteriol 190: 2611–2614.

21. Andrade SLA, Einsle O (2007) The Amt/Mep/Rh family of ammonium transport proteins. Mol Membr Biol 24: 357–365.

22. Leigh JA, Dodsworth JA (2007) Nitrogen regulation in bacteria and archaea. Annu Rev Microbiol 61: 349–377.

23. Harper C, Hayward D, Wiid I, van Helden P (2008) Regulation of nitrogen metabolism in *Mycobacterium tuberculosis*: a comparison with mechanisms in *Corynebacterium glutamicum* and *Streptomyces coelicolor*. IUBMB Life 60: 643–650.

24. Schultz AC, Nygaard P, Saxild HH (2001) Functional analysis of 14 genes that constitute the purine catabolic pathway in *Bacillus subtilis* and evidence for a novel regulon controlled by the PucR transcription activator. J Bacteriol 183: 3293–3302.

25. Nishijyo T, Haas D, Itoh Y (2001) The CbrA-CbrB two-component regulatory system controls the utilization of multiple carbon and nitrogen sources in *Pseudomonas aeruginosa*. Mol Microbiol 40: 917–931.

26. Jakoby MM, Krämer RR, Burkovski AA (1999) Nitrogen regulation in *Corynebacterium glutamicum*: isolation of genes involved and biochemical characterization of corresponding proteins. FEMS Microbiol Lett 173: 303–310.

27. Jakoby M, Nolden L, Meier-Wagner J, Krämer R, Burkovski A (2000) AmtR, a global repressor in the nitrogen regulation system of *Corynebacterium glutamicum*. Mol Microbiol 37: 964–977.

28. Meier-Wagner J, Nolden L, Jakoby M, Siewe R, Krämer R, et al. (2001) Multiplicity of ammonium uptake systems in *Corynebacterium glutamicum*: role of Amt and AmtB. Microbiology 147: 135–143.

29. Thomas G, Coutts G, Merrick M (2000) The glnKamtB operon. A conserved gene pair in prokaryotes. Trends Genet 16: 11–14.

30. Javelle A, Severi E, Thornton J, Merrick M (2004) Ammonium sensing in *Escherichia coli*. Role of the ammonium transporter AmtB and AmtB-GlnK complex formation. J Biol Chem 279: 8530–8538.

31. Strösser J, Lüdke A, Schaffer S, Krämer R, Burkovski A (2004) Regulation of GlnK activity: modification, membrane sequestration and proteolysis as regulatory principles in the network of nitrogen control in *Corynebacterium glutamicum*. Mol Microbiol 54: 132–147.

32. Sant'Anna FH, Trentini DB, de Souto Weber S, Cecagno R, da Silva SC, et al. (2009) The PII superfamily revised: a novel group and evolutionary insights. J Mol Evol 68: 322–336.

33. Hesketh A, Fink D, Gust B, Rexer H-U, Scheel B, et al. (2002) The GlnD and GlnK homologues of *Streptomyces coelicolor* A3(2) are functionally dissimilar to their nitrogen regulatory system counterparts from enteric bacteria. Mol Microbiol 46: 319–330.

34. Williams KJ, Bennett MH, Barton GR, Jenkins VA, Robertson BD (2013) Adenylylation of mycobacterial Glnk (PII) protein is induced by nitrogen limitation. Tuberculosis 93: 198–206.

35. Conroy MJ, Durand A, Lupo D, Li X-D, Bullough PA, et al. (2007) The crystal structure of the *Escherichia coli* AmtB-GlnK complex reveals how GlnK regulates the ammonia channel. Proc Natl Acad Sci USA 104: 1213–1218.

36. Gruswitz F, O'Connell J, Stroud RM (2007) Inhibitory complex of the transmembrane ammonia channel, AmtB, and the cytosolic regulatory protein, GlnK, at 1.96 Å. Proc Natl Acad Sci USA 104: 42–47.

37. Fink D, Weissschuh N, Reuther J, Wohlleben W, Engels A (2002) Two transcriptional regulators GlnR and GlnRII are involved in regulation of nitrogen metabolism in *Streptomyces coelicolor* A3(2). Mol Microbiol 46: 331–347.

38. Jessberger N, Lu Y, Amon J, Titgemeyer F, Sonnewald S, et al. (2012) Nitrogen starvation-induced transcriptome alterations and influence of transcription regulator mutants in *Mycobacterium smegmatis*. BMC Res Notes 6: 482. doi:10.1186/1756-0500-6-482.

39. Beckers G, Strösser J, Hildebrandt U, Kalinowski J, Farwick M, et al. (2005) Regulation of AmtR-controlled gene expression in *Corynebacterium glutamicum*: mechanism and characterization of the AmtR regulon. Mol Microbiol 58: 580–595.

40. Buchinger S, Strösser J, Rehm N, Hänssler E, Hans S, et al. (2009) A combination of metabolome and transcriptome analyses reveals new targets of the *Corynebacterium glutamicum* nitrogen regulator AmtR. J Biotechnol 140: 68–74.

41. Tiffert Y, Supra P, Wurm R, Wohlleben W, Wagner R, et al. (2008) The *Streptomyces coelicolor* GlnR regulon: identification of new GlnR targets and evidence for a central role of GlnR in nitrogen metabolism in *actinomycetes*. Mol Microbiol 67: 861–880.

42. Beckers G, Bendt AK, Krämer R, Burkovski A (2004) Molecular identification of the urea uptake system and transcriptional analysis of urea transporter- and urease-encoding genes in *Corynebacterium glutamicum*. J Bacteriol 186: 7645–7652.

43. Jenkins VA, Barton GR, Robertson BD, Williams KJ (2013) Genome wide analysis of the complete GlnR nitrogen-response regulon in *Mycobacterium smegmatis*. BMC Genomics 14: 301. doi:10.1186/1471-2164-14-301.

44. Malm S, Tiffert Y, Micklinghoff J, Schultze S, Joost I, et al. (2009) The roles of the nitrate reductase NarGHJI, the nitrite reductase NirBD and the response regulator GlnR in nitrate assimilation of *Mycobacterium tuberculosis*. Microbiology 155: 1332–1339.

45. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, et al. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393: 537–544.

46. Merrick MJ, Edwards RA (1995) Nitrogen control in bacteria. Microbiol Rev 59: 604–622.

47. Reitzer L (2003) Nitrogen assimilation and global regulation in *Escherichia coli*. Annu Rev Microbiol 57: 155–176.

48. Engel PC (2014) Glutamate dehydrogenases: the why and how of coenzyme specificity. Neurochem Res 39: 426–432.

49. Coulton JW, Kapoor M (1973) Studies on the kinetics and regulation of glutamate dehydrogenase of *Salmonella typhimurium*. Can J Microbiol 19: 439–450.

50. Sakamoto N, Kotre AM, Savageau MA (1975) Glutamate dehydrogenase from *Escherichia coli*: purification and properties. J Bacteriol 124: 775–783.

51. Börmann ER, Eikmanns BJ, Sahm H (1992) Molecular analysis of the *Corynebacterium glutamicum gdh* gene encoding glutamate dehydrogenase. Mol Microbiol 6: 317–326.

52. Harper CJ, Hayward D, Kidd M, Wiid I, van Helden P (2010) Glutamate dehydrogenase and glutamine synthetase are regulated in response to nitrogen availability in *Myocbacterium smegmatis*. BMC Microbiol 10: 138. doi:10.1186/1471-2180-10-138.

53. Ruiz JL, Ferrer J, Camacho M, Bonete MJ (1998) NAD-specific glutamate dehydrogenase from *Thermus thermophilus* HB8: purification and enzymatic properties. FEMS Microbiol Lett 159: 15–20.

54. Belitsky BR, Sonenshein AL (1998) Role and regulation of *Bacillus subtilis* glutamate dehydrogenase genes. J Bacteriol 180: 6298–6305.

55. Miñambres B, Olivera ER, Jensen RA, Luengo JM (2000) A new class of glutamate dehydrogenases (GDH). Biochemical and genetic characterization of the first member, the AMP-requiring NAD-specific GDH of *Streptomyces clavuligerus*. J Biol Chem 275: 39529–39542.

56. LéJohn HB, McCrea BE (1968) Evidence for two species of glutamate dehydrogenases in *Thiobacillus novellus*. J Bacteriol 95: 87–94.

57. Janssen DB, op den Camp HJ, Leenen PJ, Van der Drift C (1980) The enzymes of the ammonia assimilation in *Pseudomonas aeruginosa*. Arch Microbiol 124: 197–203.

58. Camardella L, Di Fraia R, Antignani A, Ciardiello M, di Prisco G, et al. (2001) The antarctic *Psychrobacter* sp. TAD1 has two cold-active glutamate dehydrogenases with different cofactor specificities. Characterisation of the NAD^+-dependent enzyme. Comp Biochem Physiol A 131: 559–567.

59. O'Hare HM, Durán R, Cerveñansky C, Bellinzoni M, Wehenkel AM, et al. (2008) Regulation of glutamate metabolism by protein kinases in mycobacteria. Mol Microbiol 70: 1408–1423.

60. Veronese FM, Nyc JF, Degani Y, Brown DM, Smith EL (1974) Nicotinamide adenine dinucleotide-specific glutamate dehydrogenase of *Neurospora*. I. Purification and molecular properties. J Biol Chem 249: 7922–7928.

61. Tesch M, Eikmanns BJ, de Graaf AA, Sahm H (1998) Ammonia assimilation in *Corynebacterium glutamicum* and a glutamate dehydrogenase-deficient mutant. Biotechnol Lett 20: 953–957.

62. Niebisch A, Kabus A, Schultz C, Weil B, Bott M (2006) Corynebacterial protein kinase G controls 2-oxoglutarate dehydrogenase activity via the phosphorylation status of the OdhI protein. J Biol Chem 281: 12300–12307.

63. Hänssler E, Müller T, Palumbo K, Patek M, Brocker M, et al. (2009) A game with many players: control of *gdh* transcription in *Corynebacterium glutamicum*. J Biotechnol 142: 114–122.

64. Hänssler E, Müller T, Jessberger N, Völzke A, Plassmeier J, et al. (2007) FarR, a putative regulator of amino acid metabolism in *Corynebacterium glutamicum*. Appl Microbiol Biotechnol 76: 625–632.

65. Kohl TA, Baumbach J, Jungwirth B, Pühler A, Tauch A (2008) The GlxR regulon of the amino acid producer *Corynebacterium glutamicum*: in silico and in vitro detection of DNA binding sites of a global transcription regulator. J Biotechnol 135: 340–350.

66. Antonopoulos DA, Aminov RI, Duncan PA, White BA, Mackie RI (2003) Characterization of the gene encoding glutamate dehydrogenase (*gdhA*) from the ruminal bacterium *Ruminococcus flavefaciens* FD-1. Arch Microbiol 179: 184–190.

67. Santero E, Hervás AB, Govantes F, Canosa I (2012) Glutamate dehydrogenases: enzymology, physiological role and biotechnological relevance. In: Canuto RA, editor. Dehydrogenases. Intech. doi:10.5772/47767.

68. Lu CD, Abdelal AT (2001) The *gdhB* gene of *Pseudomonas aeruginosa* encodes an arginine-inducible NAD^+-dependent glutamate dehydrogenase which is subject to allosteric regulation. J Bacteriol 183: 490–499.

69. Belanger AE, Hatfull GF (1999) Exponential-phase glycogen recycling is essential for growth of *Mycobacterium smegmatis*. J Bacteriol 181: 6670–6678.

70. Viljoen AJ, Kirsten CJ, Baker B, van Helden PD, Wiid IJF (2013) The role of glutamine oxoglutarate aminotransferase and glutamate dehydrogenase in nitrogen metabolism in *Mycobacterium bovis* BCG. PLoS ONE 8: e84452. doi:10.1371/journal.pone.0084452.

71. Amon J, Titgemeyer F, Burkovski A (2010) Common patterns - unique features: nitrogen metabolism and regulation in Gram-positive bacteria. FEMS Microbiol Rev 34: 588–605.

72. van den Heuvel RHH, Curti B, Vanoni MA, Mattevi A (2004) Glutamate synthase: a fascinating pathway from L-glutamine to L-glutamate. Cell Mol Life Sci 61: 669–681.

73. Brown JR, Masuchi Y, Robb FT, Doolittle WF (1994) Evolutionary relationships of bacterial and archaeal glutamine synthetase genes. J Mol Evol 38: 566–576.

74. Pesole G, Gissi C, Lanave C, Saccone C (1995) Glutamine synthetase gene evolution in bacteria. Mol Biol Evol 12: 189–197.

75. Yamashita MM, Almassy RJ, Janson CA, Cascio D, Eisenberg D (1989) Refined atomic model of glutamine synthetase at 3.5 Å resolution. J Biol Chem 264: 17681–17690.

76. Mathis R, Gamas P, Meyer Y, Cullimore JV (2000) The presence of GSI-like genes in higher plants: support for the paralogous evolution of GSI and GSII genes. J Mol Evol 50: 116–122.

77. Wyatt K, White HE, Wang L, Bateman OA, Slingsby C, et al. (2006) Lengsin is a survivor of an ancient family of class I glutamine synthetases re-engineered by evolution for a role in the vertebrate lens. Structure 14: 1823–1834.

78. Unno H, Uchida T, Sugawara H, Kurisu G, Sugiyama T, et al. (2006) Atomic structure of plant glutamine synthetase: a key enzyme for plant productivity. J Biol Chem 281: 29287–29296.

79. Krajewski WW, Collins R, Holmberg-Schiavone L, Jones TA, Karlberg T, et al. (2008) Crystal structures of mammalian glutamine synthetases illustrate substrate-induced conformational changes and provide opportunities for drug and herbicide design. J Mol Biol 375: 217–228.

80. Darrow RA, Knotts RR (1977) Two forms of glutamine synthetase in free-living root-nodule bacteria. Biochem Biophys Res Commun 78: 554–559.

81. Edmands J, Noridge NA, Benson DR (1987) The actinorhizal root-nodule symbiont *Frankia* sp. strain CpI1 has two glutamine synthetases. Proc Natl Acad Sci USA 84: 6126–6130.

82. Weisschuh N, Fink D, Vierling S, Bibb MJ, Wohlleben W, et al. (2000) Transcriptional analysis of the gene for glutamine synthetase II and two upstream genes in *Streptomyces coelicolor* A3(2). Mol Gen Genet 264: 461–469.

83. Kumada Y, Benson DR, Hillemann D, Hosted TJ, Rochefort DA, et al. (1993) Evolution of the glutamine synthetase gene, one of the oldest existing and functioning genes. Proc Natl Acad Sci USA 90: 3009–3013.

84. Hill RT, Parker JR, Goodman HJ, Jones DT, Woods DR (1989) Molecular analysis of a novel glutamine synthetase of the anaerobe *Bacteroides fragilis*. J Gen Microbiol 135: 3271–3279.

85. Goodman HJ, Woods DR (1993) Cloning and nucleotide sequence of the *Butyrivibrio fibrisolvens* gene encoding a type III glutamine synthetase. J Gen Microbiol 139: 1487–1493.

86. Amaya KR, Kocherginskaya SA, Mackie RI, Cann IKO (2005) Biochemical and mutational analysis of glutamine synthetase type III from the rumen anaerobe *Ruminococcus albus* 8. J Bacteriol 187: 7481–7491.

87. Eichinger L, Pachebat JA, Glöckner G, Rajandream MA, Sucgang R, et al. (2005) The genome of the social amoeba *Dictyostelium discoideum*. Nature 435: 43–57.

88. van Rooyen JM, Abratt VR, Belrhali H, Sewell T (2011) Crystal structure of Type III glutamine synthetase: surprising reversal of the inter-ring interface. Structure 19: 471–483.

89. Rexer H-U, Schäberle T, Wohlleben W, Engels A (2006) Investigation of the functional properties and regulation of three glutamine synthetase-like genes in *Streptomyces coelicolor* A3(2). Arch Microbiol 186: 447–458.

90. Carroll P, Pashley CA, Parish T (2008) Functional analysis of GlnE, an essential adenylyl transferase in *Mycobacterium tuberculosis*. J Bacteriol 190: 4894–4902.

91. Harth G, Maslesa-Galić S, Tullius MV, Horwitz MA (2005) All four *Mycobacterium tuberculosis glnA* genes encode glutamine synthetase activities but only GlnA1 is abundantly expressed and essential for bacterial homeostasis. Mol Microbiol 58: 1157–1172.

92. Harth G, Clemens DL, Horwitz MA (1994) Glutamine synthetase of *Mycobacterium tuberculosis*: extracellular release and characterization of its enzymatic activity. Proc Natl Acad Sci USA 91: 9342–9346.

93. Tullius MV, Harth G, Horwitz MA (2001) High extracellular levels of *Mycobacterium tuberculosis* glutamine synthetase and superoxide dismutase in actively growing cultures are due to high expression and extracellular stability rather than to a protein-specific export mechanism. Infect Immun 69: 6348–6363.

94. Tian ZX, Li QS, Buck M, Kolb A, Wang YP (2001) The CRP-cAMP complex and downregulation of the *glnA*p2 promoter provides a novel regulatory linkage between carbon metabolism and nitrogen assimilation in *Escherichia coli*. Mol Microbiol 41: 911–924.

95. Almassy RJ, Janson CA, Hamlin R, Xuong NH, Eisenberg D (1986) Novel subunit-subunit interactions in the structure of glutamine synthetase. Nature 323: 304–309.

96. Helling RB (1998) Pathway choice in glutamate synthesis in *Escherichia coli*. J Bacteriol 180: 4571–4575.

97. Pashley CA, Brown AC, Robertson D, Parish T (2006) Identification of the *Mycobacterium tuberculosis* GlnE promoter and its response to nitrogen availability. Microbiology 152: 2727–2734.

98. Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, et al. (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science 273: 1058–1073.

99. Jongsareejit B, Rahman RN, Fujiwara S, Imanaka T (1997) Gene cloning, sequencing and enzymatic properties of glutamate synthase from the hyperthermophilic archaeon *Pyrococcus* sp. KOD1. Mol Gen Genet 254: 635–642.

100. Klenk HP, Clayton RA, Tomb JF, White O, Nelson KE, et al. (1997) The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. Nature 390: 364–370.

101. Nesbo CL, L'Haridon S, Stetter KO, Doolittle WF (2001) Phylogenetic analyses of two “archaeal” genes in *Thermotoga maritima* reveal multiple transfers between archaea and bacteria. Mol Biol Evol 18: 362–375.

102. Beckers G, Nolden L, Burkovski A (2001) Glutamate synthase of *Corynebacterium glutamicum* is not essential for glutamate synthesis and is regulated by the nitrogen status. Microbiology 147: 2961–2970.

103. Schulz AA, Collett HJ, Reid SJ (2001) Nitrogen and carbon regulation of glutamine synthetase and glutamate synthase in *Corynebacterium glutamicum* ATCC 13032. FEMS Microbiol Lett 205: 361–367.

104. Atkinson MR, Kamberov ES, Weiss RL, Ninfa AJ (1994) Reversible uridylylation of the *Escherichia coli* PII signal transduction protein regulates its ability to stimulate the dephosphorylation of the transcription factor nitrogen regulator I (NRI or NtrC). J Biol Chem 269: 28288–28293.

105. Kamberov ES, Atkinson MR, Ninfa AJ (1995) The *Escherichia coli* PII signal transduction protein is activated upon binding 2-ketoglutarate and ATP. J Biol Chem 270: 17797–17807.

106. Arcondeguy T, Jack R, Merrick M (2001) P_{II} signal transduction proteins, pivotal players in microbial nitrogen control. Microbiol Mol Biol Rev 65: 80–105.

107. Forchhammer K (2008) P_{II} signal transducers: novel functional and structural insights. Trends Microbiol 16: 65–72.

108. Shapiro BM, Stadtman ER (1968) Glutamine synthetase deadenylylating enzyme. Biochem Biophys Res Commun 30: 32–37.

109. Carr PD, Cheah E, Suffolk PM, Vasudevan SG, Dixon NE, et al. (1996) X-ray structure of the signal transduction protein from *Escherichia coli* at 1.9 Å. Acta Crystallogr D Biol Crystallogr 52: 93–104.

110. Coutts G, Thomas G, Blakey D, Merrick M (2002) Membrane sequestration of the signal transduction protein GlnK by the ammonium transporter AmtB. EMBO J 21: 536–545.

111. Ninfa AJ, Jiang P (2005) PII signal transduction proteins: sensors of α-ketoglutarate that regulate nitrogen metabolism. Curr Opin Microbiol 8: 168–173.

112. Nichols CE, Sainsbury S, Berrow NS, Alderton D, Saunders NJ, et al. (2006) Structure of the P_{II} signal transduction protein of *Neisseria meningitidis* at 1.85 Å resolution. Acta Crystallogr Sect F Struct Biol Cryst Commun 62: 494–497.

113. Jaggi R, Ybarlucea W, Cheah E, Carr PD, Edwards KJ, et al. (1996) The role of the T-loop of the signal transducing protein P_{II} from *Escherichia coli*. FEBS Lett 391: 223–228.

114. Tsinoremas NF, Castets AM, Harrison MA, Allen JF, Tandeau de Marsac N (1991) Photosynthetic electron transport controls nitrogen assimilation in cyanobacteria by means of posttranslational modification of the *glnB* gene product. Proc Natl Acad Sci USA 88: 4565–4569.

115. Xu Y, Cheah E, Carr PD, van Heeswijk WC, Westerhoff HV, et al. (1998) GlnK, a P_{II}-homologue: structure reveals ATP binding site and indicates how the T-loops may be involved in molecular recognition. J Mol Biol 282: 149–165.

116. Blauwkamp TA, Ninfa AJ (2002) Physiological role of the GlnK signal transduction protein of *Escherichia coli*: survival of nitrogen starvation. Mol Microbiol 46: 203–214.

117. Shetty ND, Reddy MCM, Palaninathan SK, Owen JL, Sacchettini JC (2010) Crystal structures of the apo and ATP bound *Mycobacterium tuberculosis* nitrogen regulatory PII protein. Protein Sci 19: 1513–1524.

118. Nolden L, Ngouoto-Nkili CE, Bendt AK, Krämer R, Burkovski A (2001) Sensing nitrogen limitation in *Corynebacterium glutamicum*: the role of *glnK* and *glnD*. Mol Microbiol 42: 1281–1295.

119. Martínez-Hackert E, Stock AM (1997) Structural relationships in the OmpR family of winged-helix transcription factors. J Mol Biol 269: 301–312.

120. Yoshida T, Qin L, Egger LA, Inouye M (2006) Transcription regulation of *ompF* and *ompC* by a single transcription factor, OmpR. J Biol Chem 281: 17114–17123.

121. Delgado J, Forst S, Harlocker S, Inouye M (1993) Identification of a phosphorylation site and functional analysis of conserved aspartic acid residues of OmpR, a transcriptional activator for *ompF* and *ompC* in *Escherichia coli*. Mol Microbiol 10: 1037–1047.

122. Drake SK, Bourret RB, Luck LA, Simon MI, Falke JJ (1993) Activation of the phosphosignaling protein CheY. I. Analysis of the phosphorylated conformation by ^{19}F NMR and protein engineering. J Biol Chem 268: 13081–13088.

123. Jenkins VA, Robertson BD, Williams KJ (2012) Aspartate D48 is essential for the GlnR-mediated transcriptional response to nitrogen limitation in *Mycobacterium smegmatis*. FEMS Microbiol Lett 330: 38–45.

124. Amon J, Bräu T, Grimrath A, Hänssler E, Hasselt K, et al. (2008) Nitrogen control in *Mycobacterium smegmatis*: nitrogen-dependent expression of ammonium transport and assimilation proteins depends on the OmpR-type regulator GlnR. J Bacteriol 190: 7108–7116.

125. Wang Y, Cen X-F, Zhao G-P, Wang J (2012) Characterization of a new GlnR binding box in the promoter of *amtB* in *Streptomyces coelicolor* inferred a PhoP/GlnR competitive binding mechanism for transcriptional regulation of *amtB*. J Bacteriol 194: 5237–5244.

126. Rodríguez-García A, Barreiro C, Santos-Beneit F, Sola-Landa A, Martín JF (2007) Genome-wide transcriptomic and proteomic analysis of the primary response to phosphate limitation in *Streptomyces coelicolor* M145 and in a Δ*phoP* mutant. Proteomics 7: 2410–2429.

127. Muhl D, Jessberger N, Hasselt K, Jardin C, Sticht H, et al. (2009) DNA binding by *Corynebacterium glutamicum* TetR-type transcription regulator AmtR. BMC Mol Biol 10: 73. doi:10.1186/1471-2199-10-73.

128. Ramos JL, Martínez-Bueno M, Molina-Henares AJ, Terán W, Watanabe K, et al. (2005) The TetR family of transcriptional repressors. Microbiol Mol Biol Rev 69: 326–356.

129. Hasselt K, Rankl S, Worsch S, Burkovski A (2011) Adaptation of AmtR-controlled gene expression by modulation of AmtR binding activity in *Corynebacterium glutamicum*. J Biotechnol 154: 156–162.

130. Chen Y, Zhu H, Zheng G, Jiang W, Lu Y (2013) Functional analysis of TetR-family regulator AmtRsav in *Streptomyces avermitilis*. Microbiology 159: 2571–2583.

131. Hisbergues M, Jeanjean R, Joset F, de Marsac NT, Bédu S (1998) Protein PII regulates both inorganic carbon and nitrate uptake and is modified by a redox signal in *Synechocystis* PCC 6803. FEBS Lett 463: 216–220.

132. Forchhammer K (2004) Global carbon/nitrogen control by P_{II} signal transduction in cyanobacteria: from signals to targets. FEMS Microbiol Rev 28: 319–333.

133. Cho B-K, Barrett CL, Knight EM, Park YS, Palsson BØ (2008) Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. Proc Natl Acad Sci USA 105: 19462–19467.

134. Cho B-K, Federowicz S, Park YS, Zengler K, Palsson BØ (2012) Deciphering the transcriptional regulatory logic of amino acid metabolism. Nat Chem Biol 8: 65–71.

135. Maheswaran M, Forchhammer K (2003) Carbon-source-dependent nitrogen regulation in *Escherichia coli* is mediated through glutamine-dependent GlnB signalling. Microbiology 149: 2163–2172.

136. Mao X-J, Huo Y-X, Buck M, Kolb A, Wang Y-P (2007) Interplay between CRP-cAMP and PII-Ntr systems forms novel regulatory network between carbon metabolism and nitrogen assimilation in *Escherichia coli*. Nucleic Acids Res 35: 1432–1440.

137. Newman EB, Lin R (1994) Leucine-responsive regulatory protein: a global regulator of gene expression in *Escherichia coli*. Annu Rev Microbiol 49: 747–775.

138. Calvo JM, Matthews RG (1994) The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. Microbiol Rev 58: 466–490.

139. Sola-Landa A, Moura RS, Martin JF (2003) The two-component PhoR-PhoP system controls both primary metabolism and secondary metabolite biosynthesis in *Streptomyces lividans*. Proc Natl Acad Sci USA 100: 6133–6138.

140. Rodríguez-García A, Sola-Landa A, Apel K, Santos-Beneit F, Martín JF (2009) Phosphate control over nitrogen metabolism in *Streptomyces coelicolor*: direct and indirect negative control of *glnR*, *glnA*, *glnII* and *amtB* expression by the response regulator PhoP. Nucleic Acids Res 37: 3230–3242.

141. Silberbach M, Hüser A, Kalinowski J, Pühler A, Walter B, et al. (2005) DNA microarray analysis of the nitrogen starvation response of *Corynebacterium glutamicum*. J Biotechnol 119: 357–367.

142. Williams KJ, Bryant WA, Jenkins VA, Barton GR, Witney AA, et al. (2013) Deciphering the response of *Mycobacterium smegmatis* to nitrogen stress using bipartite active modules. BMC Genomics 14: 436. doi:10.1186/1471-2164-14-436.

143. Schmid R, Uhlemann EM, Nolden L, Wersch G, Hecker R, et al. (2000) Response to nitrogen starvation in *Corynebacterium glutamicum*. FEMS Microbiol Lett 187: 83–88.

144. Silberbach M, Schäfer M, Hüser AT, Kalinowski J, Pühler A, et al. (2005) Adaptation of *Corynebacterium glutamicum* to ammonium limitation: a global analysis using transcriptome and proteome techniques. Appl Environ Microbiol 71: 2391–2402.

145. Silberbach M, Burkovski A (2005) Application of global analysis techniques to *Corynebacterium glutamicum*: new insights into nitrogen regulation. J Biotechnol 126: 101–110.

146. Tiffert Y, Franz-Wachtel M, Fladerer C, Nordheim A, Reuther J, et al. (2011) Proteomic analysis of the GlnR-mediated response to nitrogen limitation in *Streptomyces coelicolor* M145. Appl Microbiol Biotechnol 89: 1149–1159.

147. Lewis RA, Shahi SK, Laing E, Bucca G, Efthimiou G, et al. (2011) Genome-wide transcriptomic analysis of the response to nitrogen limitation in *Streptomyces coelicolor* A3(2). BMC Res Notes 4: 78. doi:10.1186/1756-0500-4-78.

148. Elharar Y, Roth Z, Hermelin I, Moon A, Peretz G, et al. (2014) Survival of mycobacteria depends on proteasome-mediated amino acid recycling under nutrient limitation. EMBO J 33: 1802–1814.

149. Behrends V, Williams KJ, Jenkins VA, Robertson BD, Bundy JG (2012) Free glucosylglycerate is a novel marker of nitrogen stress in *Mycobacterium smegmatis*. J Proteome Res 11: 3888–3896.

150. Saier MH, Ballou CE (1968) The 6-O-methylglucose-containing lipopolysaccharide of *Mycobacterium phlei*. Identification of D-glyceric acid and 3-O-methyl-D-glucose in the polysaccharide. J Biol Chem 243: 992–1005.

151. Tuffal G, Albigot R, Rivière M, Puzo G (1998) Newly found 2-N-acetyl-2,6-dideoxy-beta-glucopyranose containing methyl glucose polysaccharides in *Mycobacterium bovis* BCG: revised structure of the mycobacterial methyl glucose lipopolysaccharides. Glycobiology 8: 675–684.

152. Empadinhas N, Albuquerque L, Mendes V, Macedo-Ribeiro S, da Costa MS (2008) Identification of the mycobacterial glucosyl-3-phosphoglycerate synthase. FEMS Microbiol Lett 280: 195–202.

153. Pospísl S, Halada P, Petrícek M, Sedmera P (2007) Glucosylglycerate is an osmotic solute and an extracellular metabolite produced by *Streptomyces caelestis*. Folia Microbiol (Prague, Czech Repub.) 52: 451–456.

152. Tiffert Y, Franz-Wachtel M, Fladerer C, Nordheim A, Reuther J, et al. (2011) Proteomic analysis of the GlnR-mediated response to nitrogen limitation in *Streptomyces coelicolor* M145. Appl Microbiol Biotechnol 89: 1149–1159.

153. Lewis RA, Shahi SK, Laing E, Bucca G, Efthimiou G, et al. (2011) Genome-wide transcriptomic analysis of the response to nitrogen limitation in *Streptomyces coelicolor* A3(2). BMC Res Notes 4: 78. doi:10.1186/1756-0500-4-78.

154. Engleman EG, Francis SH (1978) Cascade control of *Escherichia coli* glutamine synthetase. II. Metabolite regulation of the enzymes in the cascade. Arch Biochem Biophys 191: 602–612.

155. Jiang P, Peliska JA, Ninfa AJ (1998) Enzymological characterization of the signal-transducing uridylyltransferase/uridylyl-removing enzyme (EC 2.7.7.59) of *Escherichia coli* and its interaction with the PII protein. Biochemistry 37: 12782–12794.

156. Keener J, Kustu S (1988) Protein kinase and phosphoprotein phosphatase activities of nitrogen regulatory proteins NTRB and NTRC of enteric bacteria: roles of the conserved amino-terminal domain of NTRC. Proc Natl Acad Sci USA 85: 4976–4980.

157. Burkovski A (2003) Ammonium assimilation and nitrogen control in *Corynebacterium glutamicum* and its relatives: an example for new regulatory mechanisms in *actinomycetes*. FEMS Microbiol Rev 27: 617–628.

158. Wray LV, Atkinson MR, Fisher SH (1991) Identification and cloning of the *glnR* locus, which is required for transcription of the *glnA* gene in *Streptomyces coelicolor* A3(2). J Bacteriol 173: 7351–7360.

159. Zimmer DP, Soupene E, Lee HL, Wendisch VF, Khodursky AB, et al. (2000) Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: scavenging as a defense against nitrogen limitation. Proc Natl Acad Sci USA 97: 14674–14679.

160. Green RM, Seth A, Connell ND (2000) A peptide permease mutant of *Mycobacterium bovis* BCG resistant to the toxic peptides glutathione and S-nitrosoglutathione. Infect Immun 68: 429–436.

161. Loh KD, Gyaneshwar P, Markenscoff Papadimitriou E, Fong R, Kim K-S, et al. (2006) A previously undescribed pathway for pyrimidine catabolism. Proc Natl Acad Sci USA 103: 5114–5119.

162. Vogels GD, Van der Drift C (1976) Degradation of purines and pyrimidines by microorganisms. Bacteriol Rev 40: 403–468.

163. Dusch N, Pühler A, Kalinowski J (1999) Expression of the *Corynebacterium glutamicum panD* gene encoding L-aspartate-alpha-decarboxylase leads to pantothenate overproduction in *Escherichia coli*. Appl Environ Microbiol 65: 1530–1539.

164. Csonka LN (1989) Physiological and genetic responses of bacteria to osmotic stress. Microbiol Rev 53: 121–147.

165. Berney M, Weimar MR, Heikal A, Cook GM (2012) Regulation of proline metabolism in mycobacteria and its role in carbon metabolism under hypoxia. Mol Microbiol 84: 664–681.

166. Griffin JE, Pandey AK, Gilmore SA, Mizrahi V, McKinney JD, et al. (2012) Cholesterol catabolism by *Mycobacterium tuberculosis* requires transcriptional and metabolic adaptations. Chem Biol 19: 218–227.

167. Repoila F, Darfeuille F (2009) Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. Biol Cell 101: 117–131.

168. Haneburger I, Eichinger A, Skerra A, Jung K (2011) New insights into the signaling mechanism of the pH-responsive, membrane-integrated transcriptional activator CadC of *Escherichia coli*. J Biol Chem 286: 10681–10689.

169. Garland PB, Randle PJ (1962) A rapid enzymatic assay for glycerol. Nature 196: 987–988.

170. Weatherburn MW (1967) Phenol-hypochlorite reaction for determination of ammonia. Anal Chem 39: 971–974.

171. Dodt M, Roehr JT, Ahmed R, Dieterich C (2012) FLEXBAR-Flexible barcode and adapter processing for next-generation sequencing platforms. Biology (Basel) 1: 895–905.

172. Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30: 923–930.

173. Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11: R106. doi:10.1186/gb-2010-11-10-r106.

174. Hulsen T, de Vlieg J, Alkema W (2008) BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics 9: 488. doi:10.1186/1471-2164-9-488.

175. NCBI Resource Coordinators (2015) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 43: D6–D17. doi:10.1093/nar/gku1130.

176. UniProt Consortium (2015) UniProt: a hub for protein information. Nucleic Acids Res 43: D204–D212. doi:10.1093/nar/gku989.

177. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, et al. (2014) Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res 42: D199–D205. doi:10.1093/nar/gkt1076.

178. Caspi R, Altman T, Billington R, Dreher K, Foerster H, et al. (2014) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. Nucleic Acids Res 42: D459–D471. doi:10.1093/nar/gkt1103.

179. Fan C, Chou C-Y, Tong L, Xiang S (2012) Crystal structure of urea carboxylase provides insights into the carboxyltransfer reaction. J Biol Chem 287: 9389–9398.

180. Leftley JW, Syrett PJ (1973) Urease and ATP: urea amidolyase activity in unicellular algae. J Gen Microbiol 77: 109–115.

181. Whitney PA, Cooper TG, Magasanik B (1973) The induction of urea carboxylase and allophanate hydrolase in *Saccharomyces cerevisiae*. J Biol Chem 248: 6203–6209.

182. Kanamori T, Kanou N, Atomi H, Imanaka T (2004) Enzymatic characterization of a prokaryotic urea carboxylase. J Bacteriol 186: 2532–2539.

183. Fan C, Li Z, Yin H, Xiang S (2013) Structure and function of allophanate hydrolase. J Biol Chem 288: 21422–21432.

184. Dunn BE, Campbell GP, Perez-Perez GI, Blaser MJ (1990) Purification and characterization of urease from *Helicobacter pylori*. J Biol Chem 265: 9464–9469.

185. Batey RT, Gilbert SD, Montange RK (2004) Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. Nature 432: 411–415.

186. Schumacher MA, Choi KY, Zalkin H, Brennan RG (1994) Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by alpha helices. Science 266: 763–770.

187. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7: 539.

188. Pelicic V, Jackson M, Reyrat JM, Jacobs WR, Gicquel B, et al. (1997) Efficient allelic exchange and transposon mutagenesis in *Mycobacterium tuberculosis*. Proc Natl Acad Sci USA 94: 10955–10960.

189. Blokpoel MCJ, Murphy HN, O'Toole R, Wiles S, Runn ESC, et al. (2005) Tetracycline-inducible gene regulation in mycobacteria. Nucleic Acids Res 33: e22. doi:10.1093/nar/gni023.

190. Robson J, McKenzie JL, Cursons R, Cook GM, Arcus VL (2009) The *vapBC* operon from *Mycobacterium smegmatis* is an autoregulated toxin-antitoxin module that controls growth via inhibition of translation. J Mol Biol 390: 353–367.

191. Niesen FH, Berglund H, Vedadi M (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nat Protoc 2: 2212–2221.

192. Bailey TL, Gribskov M (1998) Combining evidence using *p*-values: application to sequence homology searches. Bioinformatics 14: 48–54.

193. Aung HL, Dixon LL, Smith LJ, Sweeney NP, Robson JR, et al. (2015) Novel regulatory roles of cAMP receptor proteins in fast-growing environmental mycobacteria. Microbiology 161: 648–661.

194. Ofer N, Wishkautzan M, Meijler M, Wang Y, Speer A, et al. (2012) Ectoine biosynthesis in *Mycobacterium smegmatis*. Appl Environ Microbiol 78: 7483–7486.

195. Paget MS, Kang JG, Roe JH, Buttner MJ (1998) σ^R, an RNA polymerase sigma factor that modulates expression of the thioredoxin system in response to oxidative stress in *Streptomyces coelicolor* A3(2). EMBO J 17: 5776–5782.

196. Poomthongdee N, Duangmal K, Pathom-aree W (2015) Acidophilic actinomycetes from rhizosphere soil: diversity and properties beneficial to plants. J Antibiot 68: 106–114.

197. Pflüger K, Baumann S, Gottschalk G, Lin W, Santos H, et al. (2003) Lysine-2,3-aminomutase and β-lysine acetyltransferase genes of methanogenic archaea are salt induced and are essential for the biosynthesis of Nε-acetyl-β-lysine and growth at high salinity. Appl Environ Microbiol 69: 6047–6055.

198. Rauschmeier M, Schüppel V, Tetsch L, Jung K (2014) New insights into the interplay between the lysine transporter LysP and the pH Sensor CadC in *Escherichia coli*. J Mol Biol 426: 215–229.

199. Pavelka MS, Jacobs WR (1996) Biosynthesis of diaminopimelate, the precursor of lysine and a component of peptidoglycan, is an essential function of *Mycobacterium smegmatis*. J Bacteriol 178: 6496–6507.

200. Robertson DE, Noll D, Roberts MF (1992) Free amino acid dynamics in marine methanogens. β-Amino acids as compatible solutes. J Biol Chem 267: 14893–14901.

201. Lee YH, Kim BH, Kim JH, Yoon WS, Bang SH, et al. (2007) CadC has a global translational effect during acid adaptation in *Salmonella enterica* serovar Typhimurium. J Bacteriol 189: 2417–2425.

202. Tetsch L, Koller C, Haneburger I, Jung K (2008) The membrane-integrated transcriptional activator CadC of *Escherichia coli* senses lysine indirectly via the interaction with the lysine permease LysP. Mol Microbiol 67: 570–583.

203. Endo G, Silver S (1995) CadC, the transcriptional regulatory protein of the cadmium resistance system of *Staphylococcus aureus* plasmid pI258. J Bacteriol 177: 4437–4441.

204. Eichinger A, Haneburger I, Koller C, Jung K, Skerra A (2011) Crystal structure of the sensory domain of *Escherichia coli* CadC, a member of the ToxR-like protein family. Protein Sci 20: 656–669.

205. Soksawatmaekhin W, Kuraishi A, Sakata K, Kashiwagi K, Igarashi K (2004) Excretion and uptake of cadaverine by CadB and its physiological functions in *Escherichia coli*. Mol Microbiol 51: 1401–1412.

206. Silhavy TJ, Kahne D, Walker S (2010) The bacterial cell envelope. Cold Spring Harb Perspect Biol 2: doi:10.1101/cshperspect.a000414.

207. Schleifer KH, Kandler O (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol Rev 36: 407–477.

208. Consaul SA, Jacobs WR, Pavelka MS (2003) Extragenic suppression of the requirement for diaminopimelate in diaminopimelate auxotrophs of *Mycobacterium smegmatis*. FEMS Microbiol Lett 225: 131–135.

209. Consaul SA, Wright LF, Mahapatra S, Crick DC, Pavelka MS (2005) An unusual mutation results in the replacement of diaminopimelate with lanthionine in the peptidoglycan of a mutant strain of *Mycobacterium smegmatis*. J Bacteriol 187: 1612–1620.

210. Pavelka MS, Jacobs WR (1999) Comparison of the construction of unmarked deletion mutations in *Mycobacterium smegmatis*, *Mycobacterium bovis* Bacillus Calmette-Guérin, and *Mycobacterium tuberculosis* H37Rv by allelic exchange. J Bacteriol 181: 4780–4789.

211. Pavelka MS, Chen B, Kelley CL, Collins FM, Jacobs WR Jr (2003) Vaccine efficacy of a lysine auxotroph of *Mycobacterium tuberculosis*. Infect Immun 71: 4190–4192.

212. Cremer J, Eggeling L, Sahm H (1991) Control of the lysine biosynthesis sequence in *Corynebacterium glutamicum* as analyzed by overexpression of the individual corresponding genes. Appl Environ Microbiol 57: 1746–1752.

213. Galán JE, Nakayama K, Curtiss R (1990) Cloning and characterization of the *asd* gene of *Salmonella typhimurium*: use in stable maintenance of recombinant plasmids in *Salmonella* vaccine strains. Gene 94: 29–35.

214. Sezonov G, Joseleau-Petit D, D'Ari R (2007) Escherichia coli physiology in Luria-Bertani broth. J Bacteriol 189: 8746–8749.

215. Anders S, Pyl PT, Huber W (2015) HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166–169.

216. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15: 550. doi:10.1186/s13059-014-0550-8.

217. Huang DW, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37: doi:10.1093/nar/gkn923.

218. Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4: 44–57.

219. Parish T, Stoker NG (2000) *glnE* is an essential gene in *Mycobacterium tuberculosis*. J Bacteriol 182: 5715–5720.

220. Tullius MV, Harth G, Horwitz MA (2003) Glutamine synthetase GlnA1 is essential for growth of *Mycobacterium tuberculosis* in human THP-1 macrophages and guinea pigs. Infect Immun 71: 3927–3936.

221. Nilsson MT, Krajewski WW, Yellagunda S, Prabhumurthy S, Chamarahally GN, et al. (2009) Structural basis for the inhibition of *Mycobacterium tuberculosis* glutamine synthetase by novel ATP-competitive inhibitors. J Mol Biol 393: 504–513.

222. Mowbray SL, Kathiravan MK, Pandey AA, Odell LR (2014) Inhibition of glutamine synthetase: a potential drug target in *Mycobacterium tuberculosis*. Molecules 19: 13161–13176.

223. Sohaskey CD, Wayne LG (2003) Role of *narK2X* and *narGHJI* in hypoxic upregulation of nitrate reduction by *Mycobacterium tuberculosis*. J Bacteriol 185: 7247–7256.

224. Sohaskey CD (2008) Nitrate enhances the survival of *Mycobacterium tuberculosis* during inhibition of respiration. J Bacteriol 190: 2981–2986.

225. Cunningham-Bussel A, Zhang T, Nathan CF (2013) Nitrite produced by *Mycobacterium tuberculosis* in human macrophages in physiologic oxygen impacts bacterial ATP consumption and gene expression. Proc Natl Acad Sci USA 110: E4256–E4265. doi:10.1073/pnas.1316894110.

226. Ehrt S, Schnappinger D (2009) Mycobacterial survival strategies in the phagosome: defence against host stresses. Cell Microbiol 11: 1170–1178.

227. Grant SG, Jessee J, Bloom FR, Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proc Natl Acad Sci USA 87: 4645–4649.

228. Snapper SB, Melton RE, Mustafa S, Kieser T, Jacobs WR (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol Microbiol 4: 1911–1919.

229. Brosch R, Gordon SV, Garnier T, Eiglmeier K, Frigui W, et al. (2007) Genome plasticity of BCG and impact on vaccine efficacy. Proc Natl Acad Sci USA 104: 5596–5601.

230. Jain P, Hsu T, Arai M, Biermann K, Thaler DS, et al. (2014) Specialized transduction designed for precise high-throughput unmarked deletions in *Mycobacterium tuberculosis*. mBio 5: e01245–14. doi:10.1128/mBio.01245-14.