Functional and regulatory characterization of the *ntrPR* operon of *Sinorhizobium meliloti*

Ph.D. thesis

Monica Bodogai

Supervisor: Ilona Dusha, D.Sc.

Biological Research Center of the Hungarian Academy of Sciences
Institute of Genetics

University of Szeged
2007

Acknowledgements

It is a pleasure for me to acknowledge people whose contribution in completing this PhD has been of great value.

I thank my supervisor, Dr. Ilona Dusha for giving me the opportunity to work in her laboratory despite the fact that I had no experience in molecular biology work. I thank her for the constant support and trust and also for general criticism concerning scientific writing.

I express my gratitude to Prof. Dr. Istvan Rasko, director of the Institute of Genetics for the financial support that made my work possible in the Nitrogen Fixation Group of the Biological Research Center of HAS.

I thank Dr. Peter Papp for giving me the opportunity to work in the Molecular Genetics Group of the Research Center for Agricultural Biotechnology, Godollo, Hungary and Szilamer Ferenczi for his valuable help and friendly atmosphere that he created in the lab during the primer extension and footprint experiments.

I am grateful to Gyongyi Cinege for revealing me the details of DNA cloning with fantastic teaching abilities. I also thank her for being such a good friend all these years.

I would like to thank to my colleague Sebastian Paul Miclea and our assistant Dr. Judit Borzi for their help and patience.

I owe special thanks to the community of the International Training Course students for their friendship and moral support. The special relations with them contributed essentially to the final outcome of my studies.

I thank to my opponents Dr. Laszlo Szabados and Dr. Erno Kiss for carefully reading the thesis and helping me to present it in the best possible shape.

I would like to thank to all the people from the Biological Research Center who helped me in different ways during my stay in Szeged.

CONTENT

1. Introduction 6

1.1 Biocomputation reveals the diversity of TA loci 6

1.2 Types of TA modules 8

1.2.1 Type I loci 8

1.2.2 Type II loci 10

1.3 Classification of type II TA modules 10

1.3.1 *ccd* (coupled cell division) locus of the F plasmid 11

1.3.2 *parDE* locus of the RK2 plasmid 12

1.3.3 *relBE* loci 12

1.3.4 *pem* and *mazEF* (*chp*) loci 12

1.3.5 *phd/doc* locus of the P1 plasmid 12

1.3.6 *higBA* locus of plasmid Rts1 13

1.3.7 *vapBC* loci 13

1.3.8 *ω*-*ε*-ζ locus of plasmid pSM19035 14

1.4 Modules analogous to the TA systems 14

1.4.1 Restriction-modification systems 14

1.4.2 *hipBA* loci 14

1.5 Mechanisms of action of characterized TA systems 15

1.5.1 Regulation of the TA modules 15

1.5.2 Targets of the toxic components of the TA modules 15

1.5.2.1 DNA gyrase as the target of the TA systems 16

1.5.2.2 mRNA as the target of the TA systems 16

1.5.2.3 Other intracellular targets of the toxin proteins 17

1.6 Physiological role of the toxin-antitoxin modules 18

1.7 Abundance of TA loci in the prokaryotic genomes 19

1.8 A toxin-antitoxin module in nitrogen fixing symbiotic bacterium 21

2. Objectives **24**

3. Materials and Methods **25**

3.1 Bacterial strains and growth conditions **25**

3.2 *β*-galactosidase measurements **27**

3.3. Viability tests **27**

3.4. The effect of antibiotics on viability **27**

3.5. DNA manipulations **27**

3.6. Databases and software **27**

3.7. Strain and plasmid constructions **28**

3.7.1. Construction of *ntrPR* deletion derivative of *S. meliloti* 28

3.7.2. Plasmids containing the promoter region of *ntrPR* operon 29

3.7.3. Plasmids carrying different regions of the ntr operon under the arabinose inducible promoter of pBAD24 30

3.8. Crude extracts of cellular proteins **32**

3.9. Gel mobility shift assays **32**

3.10. Transcription start determination **33**

3.11. DNase I footprinting **33**

4. Results **34**

4.1 Identification of additional putative TA loci encoded by the genome of *Sinorhizobium meliloti* 1021 **34**

4.2 Autoregulation of the *ntrPR* operon **37**

4.2.1 Expression of the *ntrPR* operon 37

4.2.2 Transcription start point determination 38

4.2.3 Binding of NtrP and NtrR to the *ntrPR* promoter 39

4.2.4 Determination of the exact binding site of the NtrPR complex in the *ntrPR* promoter 42

4.3 The function of the *ntrPntrR* module of *S. meliloti* **45**

4.3.1 Toxicity of the NtrR protein 45

4.3.2 Response to amino acid deprivation 46

4.3.3 Effect of antibiotics on the viability of *ΔntrPR* mutant strain 48

5. Discussion **49**

5.1 Autoregulation of the *ntrPR* operon **49**

5.2 Toxicity of NtrR protein **51**

5.3 Physiological role of the *ntrPR* operon in *Sinorhizobium meliloti* **51**

5.3.1 The activation of the *ntrPR* module 51

5.3.2 Target of the NtrR toxin 53

5.3.3 The role of a toxin-antitoxin system in a symbiotic bacterium 54

5.3.4 Diversity of TA modules in *Sinorhizobium meliloti* 55

6. References **57**

List of publications **63**
Abstract **64**
Összefoglalás **67**

1. Introduction

The technological developments of the recent years contributed to our better understanding of the prokaryotic world. Organisms that were believed to have simple organization and function revealed complex regulatory pathways, communication networks and a special type of social behavior.

Recently, the combination of genome sequencing, bioinformatic analysis and techniques of molecular biology, led to the identification and characterization of a novel type of prokaryotic regulatory mechanism exerted by the chromosomal toxin-antitoxin (TA) systems.

Originally, the TA modules were identified on bacterial extrachromosomal elements and their function was generally accepted as to prevent the proliferation of plasmid-free progeny (reviewed by Engelberg-Kulka and Glaser 1999; Gerdes 2000; Gerdes et al. 2005). After their discovery these loci were considered relatively unimportant, until 1993 when two homologues of these plasmid located TA loci were identified on the chromosome of *E. coli* (Masuda et al., 1993).

Today, the generally accepted idea about the function of these chromosomally located systems is that they act as bacterial metabolic stress managers, being associated with the modulation of the global level of translation under conditions of nutrient limitation (Christensen et al. 2001), or under various stress conditions (Sat et al. 2001; Sat et al. 2003; Hazan et al. 2004).

1.1 Biocomputation reveals the diversity of TA loci

The relative ease and rapidity of small genome sequencing and the continuous improvement of DNA sequencing techniques made possible the publication of the completely sequenced genomes at an ever-increasing pace. In 2003, the number of completely sequenced bacterial genomes was 126, while in November 2006, the number of bacterial genome sequencing projects was 1388 (397 complete and 991 ongoing projects), while the number of archaeal genome sequencing projects was 85 (29 complete and 56 ongoing projects) (Liolios et al., 2006).

A first consequence of the fast development was the accumulation of an enormous amount of data that required evaluation. A biocomputational approach contributed successfully to the information analysis. Organization of the databases and the annotation of the nucleotide and amino acid sequences helped their easier manipulation and comparison with already known functions.

Thus, in 2003 the availability of a large number of prokaryotic genome sequences enabled Anantharaman and Aravind to use a variety of computational approaches for finding relations between well characterized proteins and proteins of unknown function. The starting model for their biocomputation investigations was the RelE and ParE proteins of *Escherichia coli*. The choice was justified by the fact that these proteins were among the functionally best characterized toxins of the post-segregational killing system (PSK). This is a widespread mechanism that helps low copy number plasmids to maintain themselves in their bacterial hosts. It is based on the function of two interacting proteins encoded by the same operon. Typically, the first gene encodes a labile antitoxin that is also the transcriptional regulator of the operon and the second gene encodes a stable toxin. The antitoxin forms a complex with the toxin neutralizing its toxic effect by this way. If the plasmid carrying these modules is lost in the progeny, the antidote decays, but the stable toxin persists, killing the cell (Yarmolinsky, M. B., 1995). This should eliminate all plasmid-free cells from the population, independently of the manner by which the plasmid was lost, thus ensuring plasmid maintenance (Figure 1).

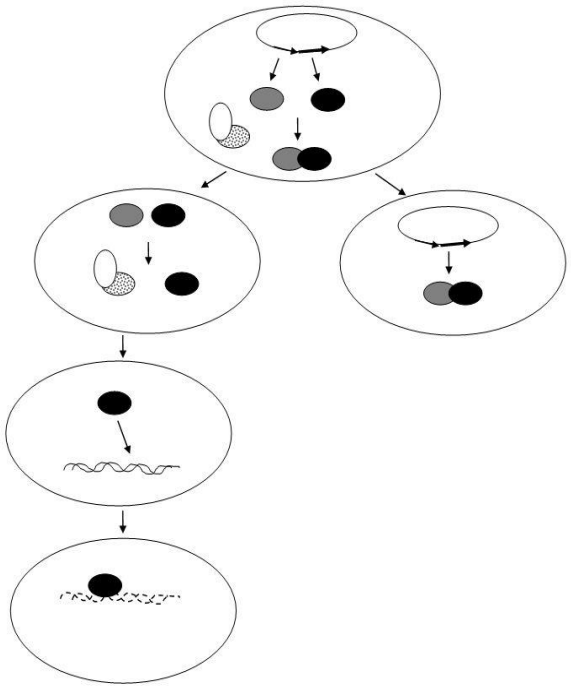

Figure 1 Schematic representation of cell death induced by PSK.

The toxin (black) and antitoxin (grey) proteins form a tight complex that negates the harmful activity of the toxin. The antitoxin is degraded by a protease (white) more rapidly than the toxin, but the latter is continually

sequestered by fresh antitoxin. As long as the plasmid is maintained, the cell tolerates the presence of the TA complex (**right**). If a missegregation event or replication defect produces a plasmid-free cell (**left**), the degraded antitoxin cannot be replenished so that the liberated toxin attacks an intracellular target to cause death or growth restriction of the plasmid-free cell. The targeting of DNA by the toxin is illustrative only (Hayes, F, 2003).

Using sequence profile analysis, Anatharaman and Aravind demonstrated the similarities between the RelE-type of toxins and several families of small, uncharacterized proteins from diverse bacteria and archaea. Gene neighborhood analysis showed that the majority of these proteins were encoded by genes in characteristic neighborhoods, in which genes encoding the toxins always co-occurred with genes encoding transcription factors that are also antitoxins. Although the transcription factors and the toxins of these operons form a comparable functional pair, they can belong to totally unrelated superfamilies of proteins. The authors observed that despite these very diverse protein associations, the strongly coupled operon architecture of the coding genes appears to be a signature of the prokaryotic TA systems.

1.2 Types of TA modules

Two types of TA systems have been identified and analyzed. In one of them (type I TA loci), the regulators of the TA operons are small antisense RNAs, in the other type (type II TA loci), the regulators are protein antitoxins.

1.2.1 Type I loci

The *hok/sok* system of plasmid R1 is the prototype of a family of antisense RNA regulated gene systems. This family consists of 13 homologue members, six of which were isolated from *Escherichia coli* plasmids: *hok/sok* of R1, *flm* and *srnB* of F, *pnd* of R483, R64 and R16 (Gerdes et al., 1990).

The *hok/sok* locus codes for the Hok (*ho*st *k*illing) protein, a membrane-associated toxin consisting of 52 amino acids that causes irreversible damage of the cell membrane and thus it is lethal to host cells. The *sok* (*s*uppression *o*f *k*illing) gene of this locus encodes an antisense RNA of 64 nucleotides (nt), which is complementary to the *hok* mRNA leader region. Sok-RNA is unstable, but is constitutively expressed from a relatively strong promoter. In contrast, *hok* mRNA is very stable and is constitutively expressed from a relatively weak promoter. The target of Sok-RNA in the *hok* mRNA leader region (SokT) overlaps with the translational initiation region (TIR) of the *mok* (*m*odulation *o*f *k*illing) reading frame, which starts upstream of and overlaps *hok*. The single stranded 5' end of Sok-RNA is complementary to the part of *hok* mRNA that encodes the *mok* TIR, including the

mok start codon. Genetic analyses showed that Sok-RNA inhibits translation of the *mok* reading frame and that translation of *hok* is coupled to the translation of *mok*. Consequently, Sok-RNA regulates translation of *hok* indirectly via *mok.* It was also proven that the RNase E is responsible for the decay of Sok-RNA by cleaving it at 6 nt from its 5' end. The RNase E cleavage occurs in the part of Sok-RNA that is responsible for the initial recognition of the target loop in *hok* mRNA and thus leads to functional inactivation of the antisense RNA (Gerdes et al, 1997, Figure 2). This causes the synthesis of the Hok protein and the killing of the plasmid-free cells.

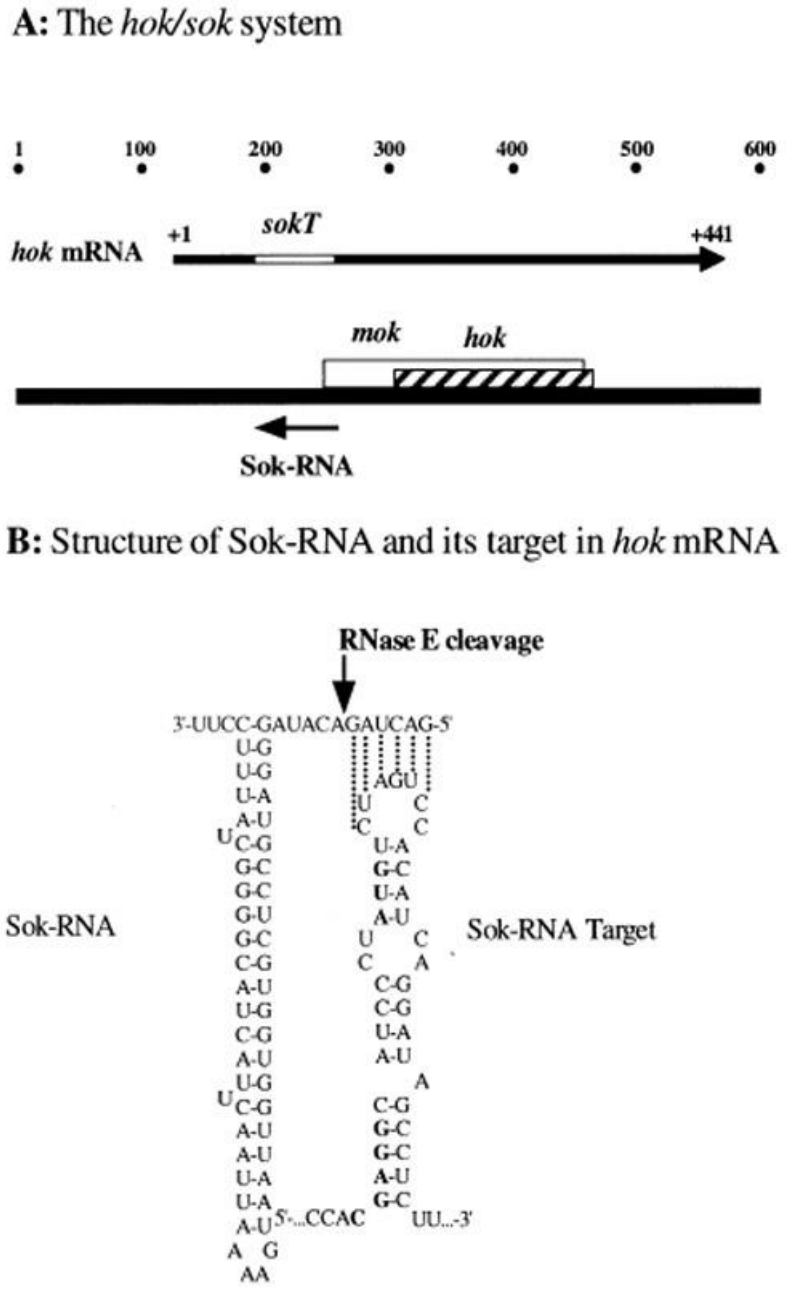

Figure 2 Overview of the *hok/sok* system of plasmid R1 (Mikkelsen and Gerdes, 1997)

A: Structural organization of the *hok-sok* locus. The *hok* gene is shown as a cross-hatched box, and the *mok* gene as an open box. Arrows pointing right and left indicate the lengths and polarities of the full-length *hok* mRNA

and the Sok antisense RNA respectively. The open region in the 5' end of the *hok* mRNA indicates the Sok target denoted sokT.
B. Primary sequence and secondary structure of the 64nt Sok antisense RNA. The arrow indicates the RNase E cleavage site, 6nt from the 5' end of Sok antisense RNA. The structure of the *sok* RNA target in *hok* mRNA is also shown. The nucleotides from the very 5' end of Sok RNA that are complementary to the loop in the target stem loop are indicated with dots.

1.2.2. Type II loci

The type II TA loci are also described as: 'poison-antidote systems', 'plasmid addiction systems' or, simply 'toxin-antitoxin loci'.

The genetic organization of a type II locus is represented in Figure 3. Two adjacent genes form an operon. They partly overlap at the stop/start codons. The encoded proteins (antitoxin and toxin) form a complex and by this way the antitoxin prevents the lethal or bacteriostatic effect of the toxin. The complex negatively autoregulates the transcription of the module, by binding to its own promoter. Under nutritional stress (e.g. amino acid starvation), transcription of TA modules decreases. Specific proteases, such as Lon or ClpA activated under these conditions degrade antitoxin molecules faster than the stable toxins (Figure 3).

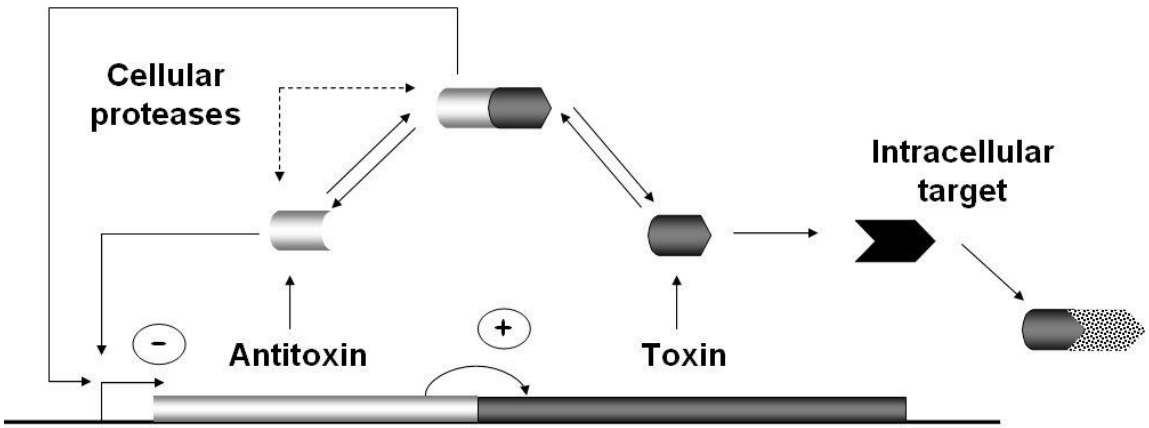

Figure 3 Genetic organization and components of a typical protein TA locus.

Toxin gene and protein are shown in dark grey, the antitoxin is represented in light grey. The interrupted arrows indicate that cellular proteases degrade the antitoxin, either free in solution or in complex with the toxin. The arrow pointing rightward indicates the presence of a promoter upstream of the TA operon. Arrows pointing to the promoter show that both the antitoxin and the TA complex are able to bind to the promoter and repress transcription. The curved arrow marks a translational coupling between the toxin and antitoxin gene. The intracellular target is represented with black or patterned symbol in accordance to its intact or toxin-degraded state (Gerdes et al., 2005)

The present work is focusing on the type II toxin-antitoxin systems. For brevity and consistency in the presentation, when referring to these systems, I will use the term TA systems.

1.3 Classification of type II TA modules

Gerdes and his coworkers arranged the toxin-antitoxin loci in eight families, seven of which are composed of two genes while one of them consists of three components (Gerdes et al., 2005, Table 1).

Table 1. The seven typical toxin-antitoxin gene families and ω-ε-ζ (Gerdes et al., 2005)

TA family (locus)	Toxin	Antitoxin	Target of toxin	Number of Loci*	Phyletic distribution
Ccd	CcdB	CcdA	DNA gyrase	5	Gram-negative bacteria
parDE	ParE	ParD	DNA gyrase	59	Gram-negative and Gram-positive bacteria
relBE	RelE	RelB	mRNA	156	Gram-negative and Gram-positive bacteria, Archaea
mazEF	MazF/PemK	MazE/PemI	mRNA	67	Gram-negative and Gram-positive bacteria
phd/doc	Doc	Phd	Translation	25	Gram-negative and Gram-positive bacteria, Archaea
higBA	HigB	HigA	Unknown	74	Gram-negative and Gram-positive bacteria
vapBC/vag	VapC	VapB	Unknown	285	Gram-negative and Gram-positive bacteria, Archaea
ω-ε-ζ	ζ	ε	Unknown	16	Gram-positive bacteria

* Number of loci found by exhaustive BLAST search in the genomes of 126 prokaryotic organisms.

1.3.1 *ccd* (coupled cell division) locus of the F plasmid

This locus was the first characterized TA module. Originally, the name of the locus derives from ‘coupled cell division’ because the coupling of cell division with plasmid proliferation, but after the detailed analysis of operon function the name of ‘control cell death’ was proposed. This operon consists of two genes: *ccdA* and *ccdB* that code for two proteins of 72 and 101 amino acids (AA), respectively. The CcdB toxin is targeting the essential DNA gyrase of *E. coli* when it is an antitoxin-free state.

The detailed analysis of the *ccdB* toxin mechanism resulted in the finding of new applications for it. The Gateway system vectors apply the *ccdB* gene as a marker for negative selection. Moreover, it was shown that the CcdB protein has the same target as quinolones, the other gyrase inhibitors, but the fact that the mode of interaction is different opens the possibility of developing a novel class of therapeutic agents (Van Melderen, 2002).

Blast search of the prokaryotic genomes showed that these loci are rare, being restricted only to a small group of Gram-negative bacteria (Pandey, 2005).

1.3.2 *parDE* locus of the RK2 plasmid

The *parDE* locus was identified on the broad-host-range IncPα plasmid RK2. Among the hosts of this plasmids are *Escherichia coli* and five other gram-negative hosts *Agrobacterium tumefaciens*, *Azotobacter vinelandii*, *Acinetobacter calcoaceticus*, *Caulobacter crescentus*, and *Pseudomonas aeruginosa*. However, the relative importance of the *parDE* operon in plasmid maintenance varied from host to host. Deletion of *parDE* had no effect on the maintenance in *A. calcoaceticus* but severely reduced it in *A. vinelandii*, and resulted in significant instability in the other hosts (Sia et al., 1995).

Although ParE toxins have weak but significant similarity with RelE and HigB toxins (Anantharaman et al., 2003), their cellular targets are different. As it was mentioned above (Table 1), RelE cleaves mRNA, while experimental data indicate that ParE targets DNA gyrase (Jiang et al., 2002).

1.3.3 *relBE* loci

The chromosome of *E. coli* K12 contains three *relBE* operons (*relBE*, *relBE-2* or *dinJ yafQ* and *relBE-3* or *yefM yoeB*), but some plasmids also harbor *relBE* homologues (Smith et al., 1997, Gronlund et al., 1999). The toxin components are activated by nutritional stresses, such as amino acid or glucose starvation, via Lon protease cleavage of their cognate antitoxins. The toxicity is achieved by inhibition of translation. In vivo and in vitro studies showed that RelE cleaves translated mRNA positioned at the ribosomal A-site (Pedersen et al., 2003, Christensen et al, 2003).

1.3.4 *pem* and *mazEF* (*chp*) loci

The *pem* (plasmid emergency maintenance) locus was identified on plasmid R100. The *pem* locus encode PemK toxin and PemI antitoxins and it is organized in a similar way as other TA loci. Two chromosomal homologues of the plasmid encoded *pemIK* have been identified in *E. coli*, *chpA* (also known as *mazEF* – 'ma-ze' in Hebrew means 'what is it?') and *chpB*. The toxins of these operons also cleave mRNA, but the specificity of cleavage is different compared to RelE toxin (Pedersen et al, 2002, Christensen et al., 2003, Zhang et al., 2003a, Zhang et al, 2005, Ruiz-Echevarria et al., 1995, Zhang et al., 2004).

1.3.5 *phd/doc* locus of the P1 plasmid

The P1 prophage is inherited as a genetically stable, extrachromosomal plasmid, partly due to the *phd/doc* (*prevent host death/death on curing*) locus of P1. It is a family whose members were identified both in bacteria and archaea. The cellular target of the Doc protein has not been identified yet, but indirect evidence indicates that the toxin inhibits translation (Hazan et al., 2001).

1.3.6 *higBA* locus of plasmid Rts1

This family of TA loci was identified in many Gram negative and Gram positive bacteria. The characteristic feature of this family is that the operon organization is different compared to other TA loci: the gene encoding the antitoxin is located downstream of the toxin gene. Phylogenetic analysis showed that HigB toxin have weak similarity to RelE, whereas conserved domain analysis showed that HigA antitoxin belongs to a different family of DNA-binding proteins, than RelB (Gerdes et al., 2005, Table 2).

Table 2. DNA-binding domains in various antitoxin proteins (Gerdes et al., 2005)

Toxin family	Antitoxin (DNA binding protein)	DNA binding motif in antitoxin
RelE	RelB; YefM	metJ/Arc/CopG; Phd/YefM
HigB	HigA	cHTH
ParE	ParD	MetJ/Arc/CopG
MazF	MazE	AbrB/MazE
VapC PIN domain	VapB	metJ/Arc/CopG; Phd/YefM; cHTH; AbrB
Doc	Phd	Phd/YefM
CcdB	CcdA	metJ/Arc/CopG
ζ	ω	metJ/Arc/CopG

1.3.7 *vapBC* loci

The *vapBC* (virulence associated protein) loci were identified on *Salmonella dublin* virulence plasmid (Pullinger et al., 1992), *Shigella flexneri* plasmid pMYS6000 (the TA locus was designated as *mvpAT* – maintenance of virulence plasmid) (Sayeed et al., 2000), in *Dichelobacter nodosus* (Katz et al., 1992) and *Leptospira interrogans* (Zhang et al., 2004). The *vapBC* family is surprisingly abundant in archaea (species like *Archaeoglobus fulgidus* and *Sulfolobus tokodaii*, contain more than 20 such loci), and more than 250 loci were identified in 126 prokaryotic organisms (Pandey et al., 2005). The cellular target of the VapC toxins was not identified, but a ribonuclease function was proposed, based on bioinformatic analysis of PIN domains, that are the conserved domains of the VapC proteins. Clissold and Pontig (2000) connected the PIN domains to eukaryotic proteins involved in nonsense-mediated mRNA decay and RNA interference suggesting a possible link between quality control gene expression in prokaryotic and eukaryotic cells.

1.3.8 *ω-ε-ζ* locus of plasmid pSM19035

This locus was identified on the pSM19035 low-copy-number, broad-host-range plasmid of *Streptococcus pyogenes*, but ε and ζ genes are also present in several Gram-positive bacteria. The operon encodes three components (an exception among the TA families): the ω repressor autoregulates the transcription of the operon, ε econdes and antitoxin and ζ encodes a toxin (Ceglowski et al., 1993, de la Hoz et al., 2000, Camacho et al., 2002, Meinhart et al., 2003).

1.4 Modules analogous to the TA systems

Other systems are also known, that present architectural analogy to the toxin-antitoxin module, but they can differ in size, in regulation and in their function in the prokaryotic cell.

1.4.1 Restriction-modification systems

Restriction-modification enzyme pairs, either plasmid- or chromosomally-located can be considered as multifunctional TA systems that promote segregational stability, as well as providing protection against invading DNAs and directing genome rearrangements. The restriction enzyme is analogous to the toxin; the modification methylase is equivalent to the antitoxin (Kobayashi, 2001).

1.4.2 *hipBA* loci

The multidrug resistance is a mechanism that protects bacterial populations from the lethal effect of antibacterial agents by producing the so called persister cells. Since the discovery of these cells in 1944, the mechanism of their appearance has remained unknown. Eventually, a mutation in the *hipA* gene resulted in a more efficient resistance to penicillin treatment by increasing the persister cell formation from 1:1.000.000 to 1: 100. Subsequent experiments led to the discovery of the *hipBA* operon (its designation derives from "high persistence"). The operon resembles in many aspects to TA modules: it is autoregulated by HipB, the complex formed between HipA and HipB inhibits the toxic effect, and the two genes are translationally coupled. The antibiotic resistance determined by the persisters is especially significant in the survival of bacterial biofilms which are responsible for nearly 65% of the human infections (Costeron et al., 2003).

1.5 Mechanisms of action of characterized TA systems

1.5.1 Regulation of the TA modules

The TA modules are regulated at different levels. Antitoxins which are DNA binding proteins are involved in the transcriptional regulation of these operons. The DNA binding motifs present in their N-termini are able to interact with operators located in the promoter region of the TA operons, repressing in this way their transcription. This was demonstrated for various TA loci such as *relBE* (Gotfredsen et al., 1998), *mazEF* (Marianovsky et al., 2001, Zhang et al., 2003b) or *phd/doc* (Magnuson et al., 1998). In most of the cases, the toxin proteins act as co-repressors by introducing cooperativity into the system because the toxin-antitoxin complexes form a much stable protein-DNA complex than the antitoxin alone (Gotfredsen et al., 1998, Zhang et al., 2003b, Marianovsky et al., 2001, Magnuson et al., 1998, Dao-Thi et al., 2002). The ParD protein encoded by the RK2 plasmid represents an exception, since this protein is sufficient for the full repression of the *parDE* operon (Davis et al, 1992, Eberl et al., 1992, Roberts et al., 1993).

The difference between the half life of the toxin and the antitoxin provides a different type of regulation for the TA operons. It is known that proteins that have a low thermal stability (T_m ~ 25ºC) have a predominant random coil conformation at physiological temperatures (Gazit et al., 1999). Due to the unfolded state they are constantly degraded by the quality control proteases that identify them as damaged proteins, thus ensuring the short half life that is critical for their physiological activity (Lehnherr et al., 1995).

The antitoxin components of the TA modules are also sensitive to the degradation by specific proteases. Studies on CcdA protein showed that under physiological conditions this protein can adopt a partially unfolded conformation which is the preferred substrate for the Lon protease (Van Melderen et al., 1996, Dao-Thi et al., 2000). This protease is involved in the selective recognition and proteolytic removal of damaged proteins, and in protein turnover. A similar C-terminal domain that can be present both in a folded or unfolded conformation was observed also in the ParD protein of plasmid RK2 (Oberer et al., 2002), in MazE (Loris et al., 2003) and YefM (Cherny et al., 2005) of *E.coli*. After the degradation of the antitoxin proteins, the toxins become free to attack an intracellular target.

1.5.2 Targets of the toxic components of the TA modules

To understand the physiological role of TA loci it is important to identify their intracellular targets. Up to now, two types of targets were identified: DNA gyrase and mRNA.

1.5.2.1 DNA gyrase as the target of the TA systems

DNA gyrase is a class II topoisomerase that introduces negative supercoils into the DNA by double-strand breaking and rejoining activity. It has been demonstrated that both CcdB and ParE proteins act as gyrase inhibitors. When the DNA gyrase is covalently linked to the DNA during one of the breaking reaction steps, CcdB can trap them into a complex that stalls the *E. coli* and T7 RNA polymerases in *in vitro* transcription reactions. This blockage was shown to be inactivated by the addition of the CcdA antitoxin. CcdB also inhibits DNA gyrase directly by trapping the free GyrA subunits. Experimental evidence indicates that the ParE toxin traps the DNA gyrase in the same way as CcdB.

1.5.2.2 mRNA as the target of the TA systems

Studies on the RelE and MazF toxins showed that both proteins inhibit translation via mRNA degradation.

RelE cleaves mRNA positioned at the ribosomal A-site both *in vitro* and *in vivo* (Pedersen et al., 2003, Christensen et al., 2003). Cleavage occurs between the second and third bases of A-site codon. It was also shown that UAG codon is cleaved 800-fold and UAA codon 100-fold more efficiently, than the UGA stop codon. By the addition of the release factor 1 that firmly binds to the ribosomal A-site (Pedersen et al., 2003, Zavialov et al., 2002) the RelE induced cleavage was reduced in vitro. In this way it was proved that this toxin must have access to the A-site in order to degrade the mRNA.

MazF also exerts endoribonuclease activity, but the experimental data concerning the recognition site are contradictory: Zhang and his colleagues (Zhang et al., 2003a, Zhang et al., 2005) demonstrated that purified MazF specifically cleaves mRNA at ACA sites independently of ribosomes, and during ectopic expression this protein cleaves mRNAs in their non-coding region. *In vivo* experiments performed by Christensen et al. (2003b) showed that both MazF and its homologue, ChpBK cleave *lpp* mRNA and SsrA within the coding region.

A common characteristic for both RelE and MazF toxins is that mRNA degradation induced by these proteins is rescued by SsrA (Christensen et al., 2003a, Christensen et al., 2003b). SsrA also known as tmRNA or 10 S RNA is a small stable RNA that is found in all eubacteria as well as in some chloroplasts and mitochondria. Its 3'- and 5'-ends are folded into a tRNA-like structure with an amino acid acceptor stem that possesses identity elements of $tRNA^{Ala}$ and enables specific aminoacylation of the SsrA by alanyl-tRNA synthetase. SsrA has a short open reading frame in the middle of the molecule surrounded by pseudoknots (the mRNA module) that encodes a degradation signal (tag peptide) for certain cellular proteases (ClpXP, ClpAP). The combination of the properties of tRNA and mRNA results in an unusual

translational mechanism of this molecule known as "*trans*-translation," switching translation from cellular mRNA to the coding part of SsrA, thus adding the tag peptide to the truncated polypeptide chain (Shpanchenko et al., 2005) (Figure 4).

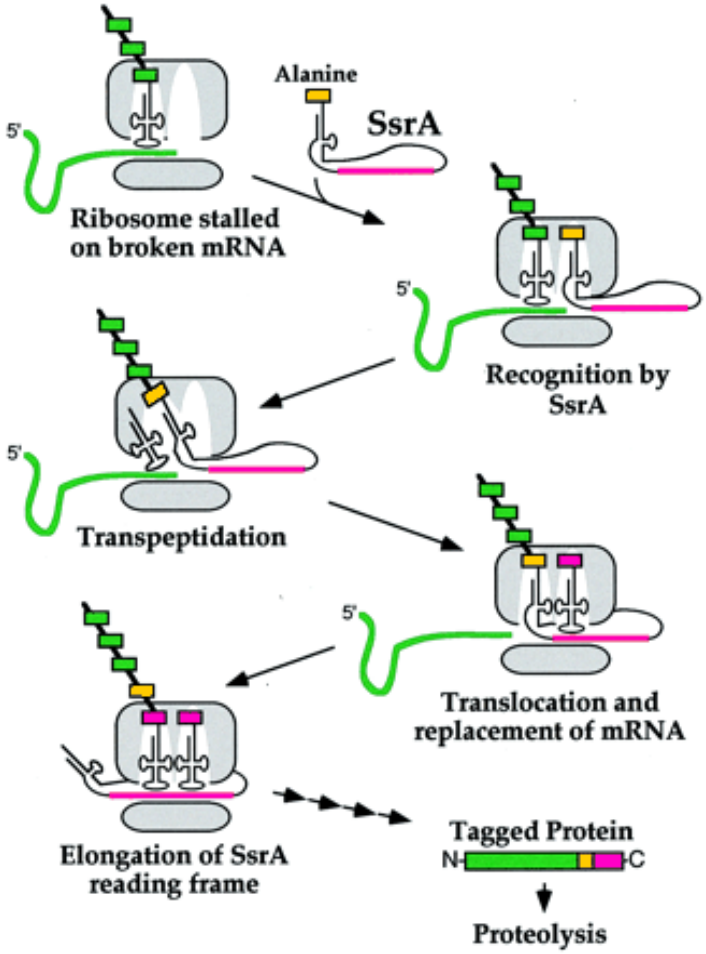

Figure 4 The tmRNA model for SsrA mediated tagging of proteins

A ribosome stalls on an incomplete or untranslatable message, leading to the recruitment of aminoacylated-SsrA RNA to the ribosomal A site and transfer of the nascent chain to the alanine-charged tRNA domain of SsrA. A message switching event then replaces the faulty mRNA by an open reading frame (shown in dark gray) within SsrA, which is translated until a stop codon is reached and the tagged protein is released to be degraded by proteases (Karzai et al., 2000).

A homologue of the MazF protein, PemK of plasmid R100 was shown to act also as a site-specific endoribonuclease, cleaving preferentially at the 5' or 3' side of the A residue in UAH sequences (where H is C, A or U).

1.5.2.3 Other intracellular targets of the toxin proteins

Very few experimental data are available about the toxins of the members of other TA gene families such as: higBA, vapBC or ω-ε-ζ.

The predicted homology of the VapC proteins to the PIN domains suggests that these proteins also have RNase activity (Clissold et al., 2000). The PIN domain containing proteins were shown to have a role in the non-sense mediated RNA decay, a quality control

mechanism that removes aberrant mRNA in eukaryotes (Cali et al., 1999, Maderazo et al., 2000). The crystal structure of an archaeal PIN domain protein has shown that this protein has a fold closely related to those of RNase H of phage T4 and the exonuclease domain of *Thermophilus aquaticus* polymerase (Arcus et al., 2004).

1.6 Physiological role of the toxin-antitoxin modules

Based on the analogy with the plasmid-located toxin-antitoxin modules whose function is to prevent the proliferation of the plasmid free progenies, TA loci identified on the bacterial chromosomes were considered to act as selfish operons that ensure their own inheritance and that of the neighboring genes (Pandey et al., 2005, Cooper et al., 2000). This hypothesis seemed to be supported by the observation that in *Vibrio cholerae* all 13 chromosomal TA systems are clustered in a mega-integron, a mobile DNA fragment that can move from one chromosomal site to another (Pandey et al., 2005, Rowe-Magnus et al., 2003). However, according to experimental data, it is relatively easy to obtain deletions in the chromosomal TA operons. This suggests that the above mentioned hypothesis may not be valid, since selfish DNA fragments that 'parasite' their hosts are usually difficult to delete from the genomes.

One of the most controversial hypotheses attributes a role of altruistic killing to TA modules (Aizenman et al., 1996, Engelberg-Kulka et al., 2003, Gonzales-Pastor et al., 2003). It is supposed that these systems can be devices that are activated under various stressful conditions such that a subpopulation of cells in a bacterial culture dies to permit the survival of the bacterial population as a whole (Aizenman, et al., 1996, Engelberg-Kulka, et al., 2004). Experimental data are available about the activation of these systems in various stressful conditions such as amino acid starvation (Hazan et al., 2004, Amitai et al., 2004), thymine starvation, DNA damage (Sat et al., 2003), the presence of antibiotics (Sat et al., 2001) or infecting phages (Hazan et al., 2004). An interesting opinion of Loris and his colleagues (2005) presents the altruistic killing hypothesis as contradiction with the Darwinian evolution theory. A mutation that inactivates TA systems will generate defector cells that always choose survival over altruistic cell death. After a single episode of stress condition, the cells that lost the TA systems will survive, eliminating altruism from the population (Buts et al., 2005).

Another explanation for the role of chromosomal TA modules is that they do not kill cells, but induce a reversible stasis that enables cells to cope with periods of extreme stress (Pedersen et al., 2002, Christensen et al. 2001, Sat et al. 2001; Sat et al. 2003; Hazan et al. 2004). When conditions improve, at least part of the population is capable of recovery and

resumes normal stress physiology. The abundance of the TA loci in the genomes of bacteria that are confronted with periodic changes in the environment strongly supports this hypothesis.

Despite 25 years of active research and the accumulation of increasing information concerning the TA modules, their importance remains enigmatic. The individual hypotheses do not provide an unambiguous answer, but there is still the possibility that the various suggestions are not fully incompatible.

1.7 Abundance of TA loci in the prokaryotic genomes

The exhaustive search of 126 completely sequenced prokaryotic genomes available in 2003 (110 bacteria and 16 archaea), led to the identification of 671 TA loci belonging to the 7 known two-component TA gene families.

Table 3 Phyletic distribution of TA loci in 126 prokaryotic organisms (Pandey et al., 2005)

Gene family	relBE	parDE	higBA	vapBC	mazEF	phd/doc	ccdAB	Total
Total in bacteria	129	59	74	139	67	22	5	495
Total in archaea	27	0	0	146	0	3	0	176
Total in 126 organisms	156	59	74	285	67	25	5	671

Table 3 summarizes the phyletic distribution of the 7 two-component TA families. The largest TA gene families were represented by *vapBC* and *relBE* families, constituting 42% and 23% of the 671 analyzed TA loci, respectively.

It was shown that prokaryotic chromosomes often encode for multiple copies of TA loci (Figure 5), while others contain no such loci. *Nitrosomonas europaea* and *Mycobacterium tuberculosis* H37Rv have 43 and 38 intact TA loci, respectively, but the genomes of obligate intracellular organisms lack TA loci. Thus, bacteria living in constant environments do not encode toxin-antitoxin modules. *Rickettsia conorii* and *Coxiella burnetii* two obligate intracellular pathogens represent exceptions, since they encode a few TA loci. An exaplanation for these exceptional cases could be that intracellular pathogens such as *Rickettsia* evolved from extracellular organisms, and the genomes of *R. conorii* and *C. burnetii* are still undergoing reductive evolution (Andersson et al., 1998, Tamas et al., 2002). Further support for this observation came from the pattern of TA loci in *Mycobacteria* (*Mycobacterium tuberculosis* H37Rv has 38 loci, whereas *Mycobacterium leprae* presents no intact TA loci) and spirochetes (free-living *Leptospira interrogans* has 5 TA loci, while

obligate parasitic spirochetes *Treponema pallidum* and *Borrelia burgdorferi* have no TA loci) (Pandey et al., 2005).

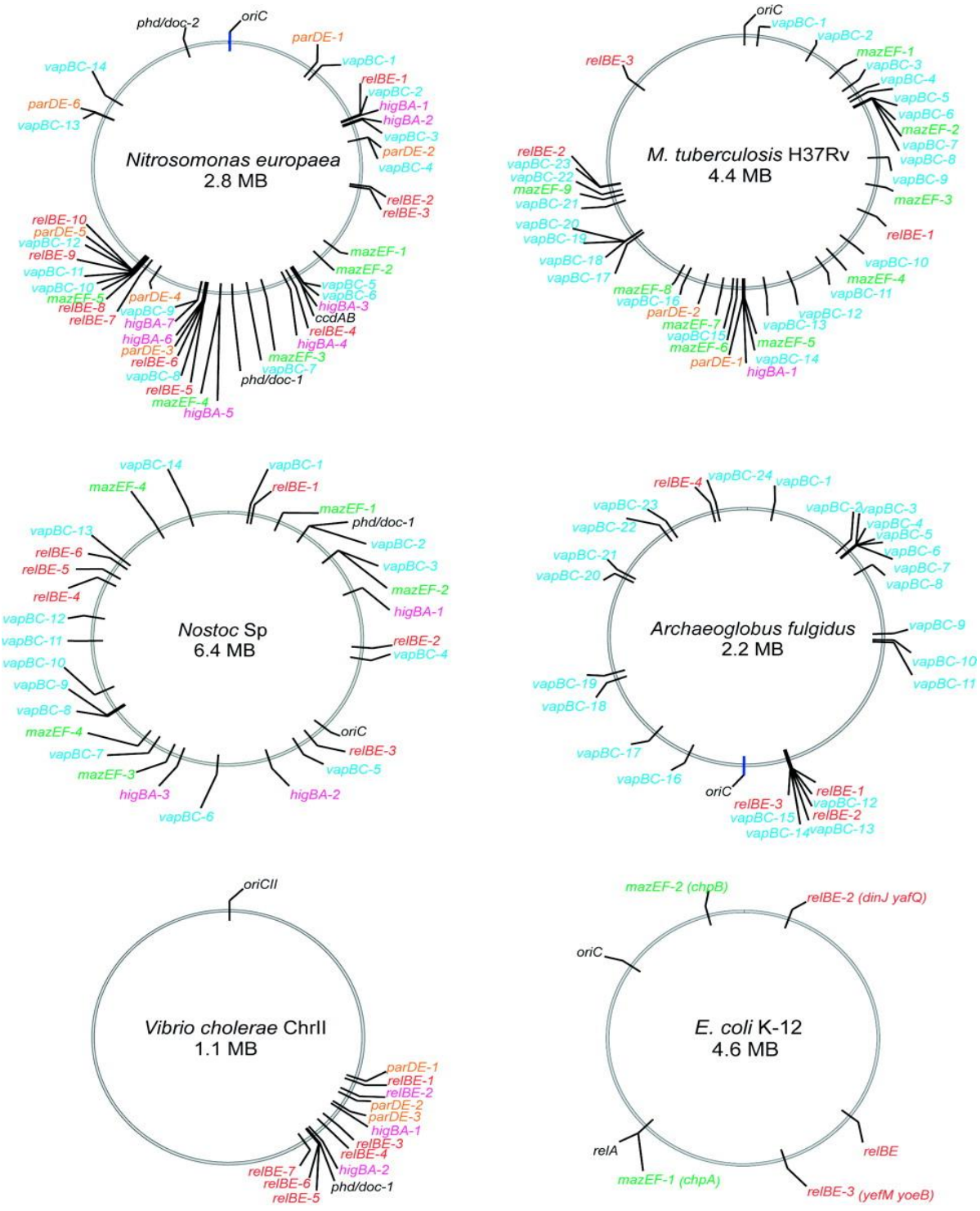

Figure 5 Chromosomal maps of TA loci of different species.

oriC denotes the origin of replication. Solitary toxin genes are also shown here as TA loci (Pandey et al., 2005).

The common feature of prokaryotic organisms possessing multiple TA loci was that they all live in nutrient-limited environments or are chemilithoautotrophs. These organisms grow very slowly and the optimization of quality control of gene expression seems to be

highly important for such organism. This observation supports the above mentioned hypothesis that TA systems are stress-response elements and/or are quality control elements that increase the fitness of free-living prokaryotes (Pandey et al., 2005).

1.8 A toxin-antitoxin module in nitrogen fixing symbiotic bacterium

The process of symbiotic nitrogen fixation is restricted to a limited number of bacterial groups, including genera *Rhizobium, Mesorhizobium, Sinorhizobium, Bradyrhizobium, Azorhizobium* and *Frankia*. The members of these genera have the unique ability to induce nitrogen-fixing nodules on the roots or stems of leguminous plants in a host specific manner: i.e. *Sinorhizobium meliloti* nodulates *Medicago, Melilotus* and *Trigonella* plants.

To initiate a productive symbiosis, bacteria must recognize and then respond to the presence of the host plant roots. At the very early steps of symbiosis, the bacterial and plant genes are activated consecutively by signal exchanges between the two partners. First, flavonoid signal molecules exuded by the host plant root induce the expression of nodulation genes (*nod, nol*). Proteins encoded by the *nod* genes synthesize bacterial signal molecules, the Nod factors. These are lipooligosaccharide molecules with various host specific structural modifications, which induce various plant reactions, such as root hair deformation, membrane depolarization, intracellular calcium oscillation, and the initiation of cell division in the root cortex, which establishes a meristem and nodule primordium.

The next step is the invasion of the bacteria from the root surface to the inner root tissue where they populate cells in the incipient nodule. This is achieved through a special tubule, called the infection thread. Once inside the nodule, bacteria differentiate to bacteroids and synthesize proteins required for nitrogen fixation. For this purpose, a symbiotic organelle-like structure is formed termed the symbiosome, which consists of the plant-derived peribacteroid membrane, the bacteroids (differentiated bacteria capable of fixing nitrogen), and the intervening peribacteroid space. In this structure all the requirements for the nitrogen fixation are fulfilled: nearly anoxic conditions for the oxygen sensitive nitrogenase enzyme and carbon supplies provided by the host plant for the extremely energy-intensive process of atmospheric nitrogen reduction.

One of the factors that were demonstrated to control the efficiency of this special plant-bacterium interaction is the available nitrogen source. It was shown that nodulation genes (*nod*) involved in the production of the first bacterial signal molecules, the Nod factors, are repressed in the presence of ammonium (Dusha et al., 1989). At a later stage of nodule

development, nitrogen fixation efficiency is also reduced, due to the decreased expression of *nif* genes encoding the nitrogenase enzyme (Olah et al., 2001).

Previous work showed that a Tn5 insertion into the chromosomal gene *ntrR* of *Sinorhizobium meliloti* resulted in increased transcription of *nod* and *nif* genes as compared to that of the wild type strain, and this effect was more pronounced in the presence of an external ammonium source (Dusha et al. 1989; Oláh et al. 2001). As a result of this pleiotropic effect, alfalfa plants inoculated by the mutant strain had increased nitrogen content and biomass production (Oláh et al. 2001). Based on these observations, *ntrR* and the upstream *ntrP* gene were first proposed to function as nitrogen regulatory genes in *S. meliloti*. However, when the gene expression patterns of the entire genomes of wild type and *ntrR* mutant strains were compared under oxic and microoxic conditions, an unexpectedly large number of genes exhibited altered expression in the mutant strain. Since a relatively wide range of unrelated functions was modulated, we suggested that *ntrR* is not a nitrogen regulatory gene, and a more general function for NtrR was proposed (Puskás et al. 2004).

Earlier the comparison of DNA sequence (and the deduced amino acid sequence) of the *ntrPR* operon with databases revealed homology to the *vapBC* genes of *Dichelobacter nodosus* (Katz et al. 1992) and to related genes of various microorganisms (Puskás et al. 2004). Since the functions of the homologous systems were either unknown or apparently unrelated, the molecular mechanism of NtrPR remained a question. As a further attempt, NtrP and NtrR were searched for conserved domains to elucidate their function. The NtrP protein was shown to carry a SpoVT/AbrB-like domain between amino acids 15-65 (Puskás et al. 2004). The AbrB domain, first described in a protein of *Bacillus subtilis* (Strauch et al. 1989) regulates a set of unrelated genes, which contribute to the adaptive capacity for survival under suboptimal conditions and represents a novel class of DNA recognition domain (Vaughn et al. 2000). The NtrR protein had high homology to the PIN domain, which was described as the N-terminal domain of the PilT protein (Wall and Kaiser 1999; Wolfgang et al. 2000), and later was identified in many bacteria, archaea and eukaryota (Makarova et al. 1999). Apart from suggesting a possible role for the PIN domain in signalling processes (Noguchi et al. 1996), no information was available about its functional role in different proteins.

Sequence analysis, motif-based studies and secondary structure predictions, helped Clissold and Ponting (2000) to identify PIN domains in proteins participating in nonsense-mediated mRNA decay and RNAi in eukaryotic cells. They found significant similarities between the 5’ nuclease domain families and the PIN domain, which suggested a phosphodiesterase, more specifically RNase activity for PIN domain proteins.

By using bioinformatic approaches Anantharaman and Aravind (2003) examined the evolutionary relationship between the PIN domains and bacterial toxin-antitoxin (TA) modules.

Based on the homology, we proposed that the NtrPR proteins belong to the most highly abundant class of TA systems, the *vapBC* family (Gerdes et al. 2005). The organization of the *ntrPR* operon exhibits remarkable similarities to bacterial TA systems. The two partly overlapping genes encode two small proteins (90 and 134 amino acids in NtrP and NtrR, respectively) (Puskás et al. 2004). Preliminary data also indicated a negative autoregulatory function for NtrR (Oláh et al. 2001). The domain organization resembled one family of TA systems: the presence of the AbrB domain in the first, and a PIN domain in the second protein.

2. Objectives

Our aim was to examine whether or not, the NtrP and NtrR proteins represent an active TA system in *Sinorhizobium meliloti,* based on their autoregulatory properties, complex formation and function.

If the toxin-antitoxin modules are considered to function as metabolic stress managers, *Sinorhizobium meliloti* can be a valuable test organism for studying such systems.

The metabolism of this symbiotic nitrogen fixing bacterium is influenced by a wide-range of conditions under which it is able to survive: it can be found in the soil as a free-living bacterium, or it occupies the nodules developed on the roots of leguminous plant, it is able to live under aerobic conditions or function under microoxic conditions as a bacteroid, the differentiated form of the bacterium. Another distinct feature of *S. meliloti* is the special metabolism in the symbiotic state: the carbon source utilized by the bacteroids in the form of dicarboxylic acids is supplied by the plant metabolism to fuel nitrogen fixation, while in exchange, fixed nitrogen is transferred from the bacteroid to the plant cell.

Toxin-antitoxin systems present in the genome of the bacterium may have an important role during the transition from one way of life to the other, and may determine the adaptation to the varying environmental conditions.

The aim of our work was to examine:

- the different aspects of the autoregulation of the *ntrPR* operon by:
    - the examination of the role of the two encoded proteins in the autoregulation;
    - determining the role of the NtrP and NtrR proteins in DNA binding and NtrPR complex formation;
    - localization of the binding site of the NtrPR protein complex in the *ntrPR* promoter region.
- the possible function of this TA system in *Sinorhizobium meliloti* by determining:
    - the toxicity of the NtrR protein;
    - the response of the *ntrPR* operon to different stress conditions.

3. Materials and Methods

3.1 Bacterial strains and growth conditions

Bacterial strains and plasmids used in this work are described in Table 4. *E. coli* strains were grown at 37ºC in LB medium (Maniatis et al. 1982) supplemented with the appropriate antibiotics (ampicillin, 100 μg/ml; kanamycin, 50 μg/ml, tetracycline, 10 μg/ml). L-arabinose was used at a final concentration of 1%. *S. meliloti* strains were grown at 31ºC in TA enriched medium (Kondorosi et al. 1984) supplemented with the appropriate antibiotics: kanamycin (200 μg/ml), streptomycin (200 μg/ml), and tetracycline (15 μg/ml).

Table 4 Bacterial strains and plasmids used in this study

Strain – Plasmid	Relevant characteristics	References
Strains		
E. coli		
MG1655	wild type *E. coli* K-12 *F⁻ lambda⁻ ilvG⁻ rfb⁻⁵⁰ rph⁻¹*	Blattner, et al. (1997)
DH5α	*supE44 D lacU169 hsdR17 recA1 endA1gyrA96 thi-1 relA1*	Maniatis et al. (1982)
Sinorhizobium meliloti		
1021	wild type, Sm^r, derivative of SU47	Ausubel, F.M.
BM175	*S.m.* 1021 carrying the pBM174, Sm^r, Km^r	Bodogai et al. (2006)
BM111	*S.m.* 1021 carrying a deletion in the *ntrPR* operon	Bodogai et al. (2006)
399	*S. m.*1021 *ntrR::Tn5*	Dusha et al. (1989)
866	*S.m.*1021 *ntrP*::Km	Olah et al. (2004)
Plasmids		
pBluescriptII SK+/-	Amp^r, f1 (+) *ori*	Alting-Mees and Short (1989)
pSEM211	Km^r, pCU1 replicon, RP4*mob*, promoterless *lacZ*	Ferenczi et. al. (2004)
pEP82	Tc^r. broad-host-range, low copy number plasmid, pRK290 derivative	Elo et al., (1998)
pBM174	Km^r, derivative of pSEM211 carrying the PPF-PPR PCR product cloned in the EcoRI site	Bodogai et al. (2006)
pBM88	Amp^r, derivative of pBluescriptII(SK+/-) lacking the EcoRI-SalI region	Bodogai et al. (2006)
pBAD24	Amp^r, ColE1 replication origin, pBAD promoter	Guzman et. al. (1995)
pBM120	Amp^r, derivative of pBM88 carrying the 108-109 PCR fragment cloned in the filled XhoI site (*ntrR* gene)	Bodogai et al. (2006)
pBM121	Amp^r, derivative of pBM88 carrying the 105-106 PCR fragment cloned in the filled XhoI site (*ntrP* gene)	Bodogai et al. (2006)
pBM122	Amp^r, derivative of pBM88 carrying the 111-109 PCR fragment cloned in the filled XhoI site (second part of the *ntrR* gene)	Bodogai et al. (2006)
pBM123	Amp^r, derivative of pBM121 and pBM122 carrying the 105-106 and 111-109 PCR fragments (*ntrP* and the second part of *ntrR* fused *in frame* through a *BglII* site)	Bodogai et al. (2006)
pBM124	Amp^r, derivative of pBM88 carrying the 105-	Bodogai et al. (2006)

	109 PCR fragment cloned in the filled XhoI site (*ntrPR* gene)	
pBM125	Amp^r, derivative of pBM88 carrying the 105-112 PCR fragment cloned in the filled XhoI site (first part of *ntrP* gene)	Bodogai et al. (2006)
pBM126	Amp^r, derivative of pBM88 carrying the 107-109 PCR fragment cloned in the filled XhoI site (*ntrR* gene)	Bodogai et al. (2006)
pBM127	Amp^r, derivative of pBM88 carrying the 105-112 and 107-109 PCR fragments (first part of *ntrP* and the *ntrR* gene fused in frame through anEcoRV site)	Bodogai et al. (2006)
pBM128	Amp^r, derivative of pBM88 carrying the 110-109 PCR fragment cloned in the filled XhoI site (second part of *ntrP* and the *ntrR* gene)	Bodogai et al. (2006)
pBM130	Amp^r, derivative of pBM88 carrying the 105-113 PCR fragment cloned in the filled XhoI site (*ntrP* and the first part of the *ntrR* gene)	Bodogai et al. (2006)
pBM135	Amp^r, derivative of pBAD24 carrying the *NcoI-PstI* fragment of pBM120 (*ntrR* gene)	Bodogai et al. (2006)
pBM136	Amp^r, derivative of pBAD24 carrying the *NcoI-PstI* fragment of pBM121 (*ntrP* gene)	Bodogai et al. (2006)
pBM137	Amp^r, derivative of pBAD24 carrying the *NcoI-PstI* fragment of pBM123 (*ntrP* gene and the second part of *ntrR*)	Bodogai et al. (2006)
pBM138	Amp^r, derivative of pBAD24 carrying the *NcoI-PstI* fragment of pBM124 (*ntrPR* gene)	Bodogai et al. (2006)
pBM139	Amp^r, derivative of pBAD24 carrying the *NcoI-PstI* fragment of pBM127 (first part of *ntrP* and the *ntrR* gene)	Bodogai et al. (2006)
pBM140	Amp^r, derivative of pBAD24 carrying the *NcoI-PstI* fragment of pBM128 (second part of *ntrP* and the *ntrR* gene)	Bodogai et al. (2006)
pBM141	Amp^r, derivative of pBAD24 carrying the *NcoI -KpnI* (*KpnI* filled in) fragment of pBM130 (*ntrP* and the first part of the *ntrR* gene)	Bodogai et al. (2006)
pBM24	Amp^r, derivative of pBluescriptII(SK+/-) carrying the PPF-PPR PCR fragment in the EcoRV site (397 bp *ntrP* promoter region)	Bodogai et al. (2006)
pBM170	Amp^r, derivative of pBluescriptII(SK+/-) carrying the 116-PPR PCR fragment in the EcoRV site (312 bp *ntrP* promoter region)	Bodogai et al. (2006)
pBM171	Amp^r, derivative of pBluescriptII(SK+/-) carrying the 117-PPR PCR fragment in the EcoRV site (224 bp *ntrP* promoter region)	Bodogai et al. (2006)
pBM172	Amp^r, derivative of pBluescriptII(SK+/-) carrying the 118-PPR PCR fragment in the EcoRV site (141 bp *ntrP* promoter region)	Bodogai et al. (2006)
pBM40	Amp^r, derivative of pBluescriptII SK+/- carrying the 4384 SalI-EcoRI *S.m.* genomic region in the SalI-EcoRI site	Not published
pBM220	Amp^r, Km^r, derivative of pBM40, carrying a Km resistance cassette between the MfeI-PaeI sites	Not published
pBM90	Tc^r, derivative of pEP82, carrying the PPF-PPR PCR fragment in the BglII site	Not published
pBM92	Tc^r, Km^r, derivative of pRK290 carrying the SalI-EcoRI fragment of pBM220 in the EcoRI site	Not published
pRK2013	Km^r, helper plasmid for mobilization of pRK290 derivatives, tra (RK2)	Ditta et al. (1980)
pRK290	Tc^r, broad-host-range vector, IncP	Ditta et al. (1980)

3.2 *β*-galactosidase measurements

The *S. meliloti* strains grown at 31 °C in complete TA enriched medium containing the appropriate antibiotics, were diluted to OD_{600}= 0,1 and further cultivated. The *β*-galactosidase activity was measured when these cultures reached OD_{600}=0,5 (Miller et al., 1972). The enzyme activity is calculated as 1000 x OD_{420} / min x OD_{600} x ml.

3.3. Viability tests

E. coli DH5α cells with plasmids pBM135, pBM136, pBM138 or the vector pBAD24 were grown overnight in LB medium supplemented with ampicillin. The cultures were diluted in the same medium to obtain equal optical densities (OD_{600}=0.1), and were induced with 1% L-arabinose. Cultures were grown at 37^{o}C and samples were removed to determine optical density and viable cell number.

3.4. The effect of antibiotics on viability

S. meliloti 1021 wild type and BM111 deletion derivative (Δ*ntrPR*) strains were grown in complete TA medium to OD_{600}=0.4, centrifuged and washed with M9 minimal medium (Dusha et al. 1989), then suspended in M9 containing antibiotics as indicated (rifampicin: 10-25 μg/ml; chloramphenicol: 50-75 μg/ml; spectinomycin: 100-200 μg/ml; claforan: 100-200 μg/ml). Samples were incubated for 10 min at 31^{o}C, then diluted and plated on TA medium to determine CFU/ml.

3.5. DNA manipulations

Plasmids were transferred into *S. meliloti* strains by conjugation using the triparental mating system (Ditta et al. 1980). Preparation of plasmid DNA, digestion with restriction enzymes, agarose gel electrophoresis, fragment isolation, cloning procedures and transformation of *E. coli* cells were performed according to Maniatis et al. (1982). Enzymes used for restriction and modification of DNA samples were obtained from Fermentas or Promega.

3.6. Databases and software

The complete genomic sequence of *Sinorhizobium meliloti* is available at http://www.kazusa.or.jp/rhizobase/. All other sequence analyses were done using the Bioedit program.

For the identification of all presumptive TA modules present in the genome of *Sinorhizobium meliloti*, a search for TA protein homologues was initiated in the Pfam database. After the identification of the ID of the homologous proteins, the properties of the given proteins were compared with those of a TA module: gene neighbourhood, protein size and conserved domains. If two genes presented a length and topology similar to the toxin-antitoxin modules, based on the domain content of the encoded proteins, we classified them according to the nomenclature of TA gene families proposed by Gerdes in 2005.

3.7. Strain and plasmid constructions

3.7.1. Construction of *ntrPR* deletion derivative of *S. meliloti*

A 4384 bp EcoRI-SalI genomic fragment carrying the *ntrPR* operon was cloned in pBluescriptIISK. The resulting pBM40 plasmid was used to generate a construct carrying deletion in the *ntrPR* operon (Figure 6).

By MfeI-PaeI digestion a fragment was removed, which covered 184 bps upstream of the translation start of *ntrP*, the complete *ntrP* gene, and a large part of *ntrR* gene up to the PaeI restriction site (located 89 bp upstream of the stop codon of *ntrR*). In this way a 771 bp deletion was created. The MfeI and PaeI sites were filled up by Klenow polymerase and T_4 DNA polymerase, respectively, and the kanamycin resistance gene was inserted between them giving rise to plasmid pBM220. The EcoRI-SalI fragment with the deletion derivative carrying the Km^r marker was cut from pBM220 and recloned in the vector pRK290 at the EcoRI site (also filled up by Klenow polymerase), generating plasmid pBM92. The schematic representation of these plasmid constructs is shown in Figure 4.

Finally, the recombinant plasmid (pBM92) was introduced into the wild type *S. meliloti* 1021 by triparental mating, and the deletion was introduced into the chromosome by double recombination at the homologous regions flanking the Km^r marker. The resulting Km^r and Tc^s colonies were checked by Southern hybridizations and the strain BM111 was selected for further manipulations.

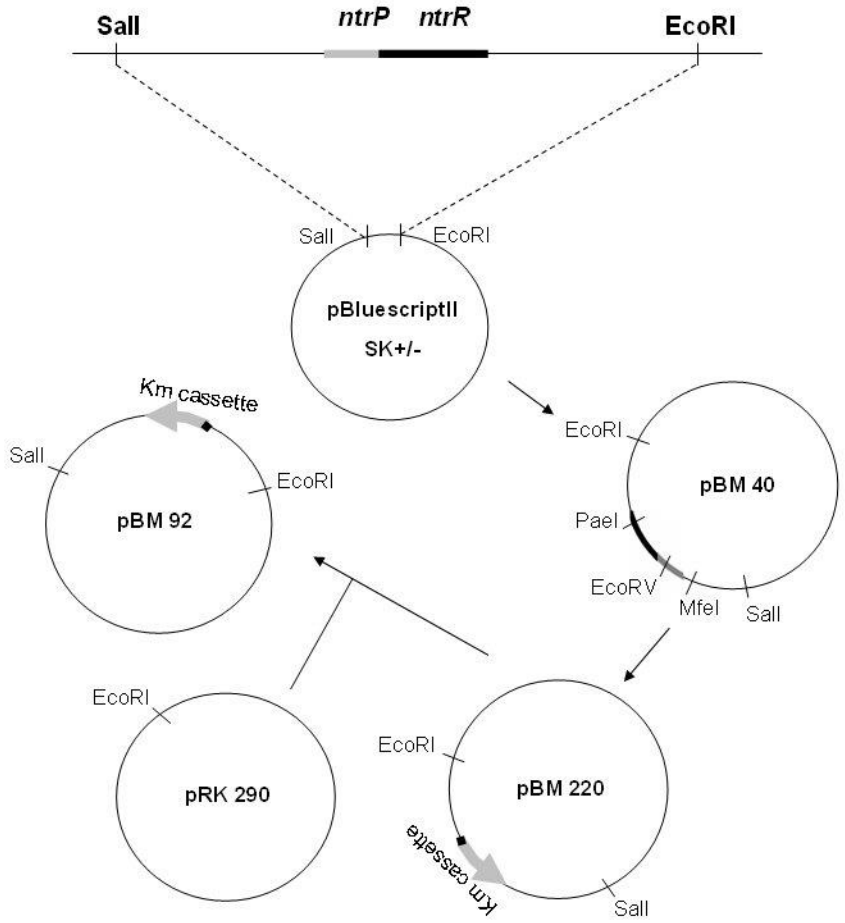

Figure 6 Schematic representation of plasmid construction for generating the deletion mutant strain *S. meliloti* BM111

The black thick lines represent the *ntrR* gene or gene fragments, the gray lines show the *ntrP* gene, while the gray arrow corresponds to the Km resistance cassette.

3.7.2. Plasmids containing the promoter region of *ntrPR* operon

For determining the β-galactosidase levels expressed from an *ntrPR* promoter-*lacZ* fusion in different mutant backgrounds, pBM90 transcriptional fusion vector was constructed. The pEP82 vector contains a promoterless *lacZ*-gene. By the insertion of a promoter region upstream of the start codon of the *lacZ* gene, a transcriptional vector can be generated that allows the detection of the activity of the inserted promoter. A DNA fragment carrying the ntrPR promoter was generated by DNA-polymerase chain reaction (PCR) by using the primers PPF-101 and inserted in the the BglII site of plasmid pEP82, filled up by Klenow polymerase. In this way, we obtained the transcriptional fusion vector, pBM90. The fragment amplified with the PPF-101 primers, covered a 339 bp region upstream of the translational start point of *ntrP* gene. pBM90 was introduced by triparental mating in *S. meliloti* 1021, *S. meliloti 866*, *S. meliloti* 399 and BM111 to determine the *ntrPR* promoter activity in S. meliloti wild type, *ntrR*, *ntr*P and *ntrPR* mutant derivatives.

For transcriptional start determination, a DNA fragment was amplified by PCR using the primers PPF-PPR (Table 5). The fragment was filled up by Klenow polymerase and was cloned in the EcoRI site (also filled up) of the transcriptional fusion vector, pSEM211 (Ferenczi et al. 2004), resulting in the plasmid pBM174.

To localize the binding site of the autoregulatory NtrPR complex, plasmids carrying the promoter region on fragments of various lengths were constructed (Figure 12). DNA fragments of 397 bp, 312 bp, 224 bp and 141 bp fragments were amplified using the primers shown in Figure 12 filled up by Klenow polymerase and cloned in the EcoRV site of BluescriptIISK vector resulting in the plasmids pBM24, pBM170, pBM171 and pBM172, respectively. The XhoI-BamHI restriction fragments of these pBluescriptIISK derivatives were used for electrophoretic mobility shift assays.

3.7.3. Plasmids carrying different regions of the ntr operon under the arabinose inducible promoter of pBAD24

Plasmids carrying intact or truncated genes of the *ntrPR* operon were constructed by cloning the appropriate PCR fragments first in pBM88, and then the inserts were recloned in the vector pBAD24 (Figure 10). pBM88 is a derivative of pBluescriptII lacking the EcoRI-SalI region from the multicloning site, generated for obtaining a plasmid that do not contain the EcoRV restriction site. We chose for our work a vector that lacks an EcoRV recognition site because primers 107 and 112 contain an EcoRV site in their sequence that was used in order to obtain *in frame* fusion proteins. The initial cloning of the amplified fragments in pBM88 was carried out to generate by digestion the appropriate sticky-ends necessary for the following *in frame* cloning in pBAD24. NcoI site was used for the fusion to the arabinose inducible promoter of the pBAD24 vector, while EcoRV and BglII sites were used for the generation of in frame truncated versions of the *ntrPR* operon. The following plasmids were obtained: pBM121 with the 105-106 PCR fragment, pBM120 with the 108-109 PCR fragment, pBM124 with the 105-109 PCR fragment, pBM125 with the 105-112 PCR fragment, pBM126 with the 107-109 PCR fragment, pBM128 with the 110-109 PCR fragment, pBM122 with the 111-109 PCR fragment, pBM130 with the 105-113 PCR fragment. Two additional plasmids were created in consecutive steps: pBM123 carrying 105-106 and 111-109 fragments fused in frame at a BglII restriction site, and pBM127 containing 105-112 and 107-109 fragments fused in frame at an EcoRV restriction site. For the creation of plasmid BM123, we used plasmids pBM121 and pBM122 carrying the PCR fragments mentioned above. pBM121 was digested by BglII (restriction site present on the 5' end of primer 106) and BamHI (present on the polylinker), and subsequently used as a vector for

cloning the BglII-BamHI fragment isolated from pBM122. In this way, the *ntrP* gene was fused in frame to the second part of the *ntrR* gene at a BglII site. In the case of plasmid pBM127, plasmids pBM125 and pBM126 were used. pBM125 was digested with EcoRV and BamHI, and used as a vector for cloning the EcoRV- BamHI fragment isolated from pBM126, obtaining in this way an in-frame fusion of the first part of the *ntrP* gene with the *ntrR* gene.

The oligonucleotide primers used to obtain the various constructions are listed in Table 5. The pBM88 derivatives carrying the cloned fragments were digested with NcoI (present on the fragment) and BamHI (present on the multicloning site of the vector), and the inserts were recloned in the NcoI-BamHI digested pBAD24 vector resulting in the plasmids listed in Table 4. In the case of pBM141 the cloned fragment was cut with NcoI-KpnI enzymes, and after filling up the KpnI site with Klenow polymerase, the fragment was cloned in the NcoI-BamHI digested pBAD24 vector (also filled up at the BamHI site). Thus, the NcoI restriction site of pBAD24 and the translational start codon ATG were restored in plasmids pBM135, pBM136, pBM137, pBM138, pBM139, pBM140 and pBM141, due to the presence of an NcoI site at the 5' end of the corresponding primers.

Table 5. Oligonucleotide sequences

Primer	Sequence 5'- 3'	Orientation and localization *
101^0	GATATCGGCATTCGTCTCCTTTGT	-(+5 - +22)
105^1	CCATGGTGCCCGTCCCGTTGCCGTCA	+(+24 - +44)
106^2	AGATCTAAGATAGCCGTTCAAAAG	-(+289 - +306)
107^0	GATATCTTTTGAACGGCTATC	+(+287 - +304)
108^1	CCATGGACGGCTATCTTCTCGATACGAA	+(+296 - +317)
109	AAAACGTGTCGCCTATCGCTC	-(+688 - +708)
110^1	CCATGGGACTCATTATCGAACC	+ (+156 - +171)
111^2	AGATCTCTGGAAATCGTCCCGGTGCTG	+(+481 - +501)
112	GCTCCACGATATCGGCTG	-(+183 - +200)
113^3	CAGCACCGGGACGATTTACAG	-(+481 - +501)
PPF	AGGTCATCGAAGAGGGCGT	+(-317 - -299)
PPR	GGTTATTGCGGAACAGCTTT	-(+61 - +80)
116	TGTTTGTCGCGCCGAGGCAGATCA	+(-61 - +38)
117	TTGCTTCTTCCTGGAGGGGGCGCT	+(-144 - +121)
118	ATACGGCGTCGAGGGTCA	+ (-232 - -215)
7	GGCGCGCCCCCGGGTACC	

123[4]	TCTAGAATGGTTCGGCATATACATTTAGGCA TATACAAAGGAGACGAAGTCGAC	+(-24 - +18)
124[5]	GTCGACTTCGTCTCCTTTGTATATGCCTAAAT GTATATGCCGAACCATTCTAGA	-(-24 - +18)
125	TTTAGGCATATACAAAGGAGACGAATGCCA	+(-7 - +23)
126	TGGCATTCGTCTCCTTTGTATATGCCTAAA	-(-7 - +23)

[0] The primer contains a EcoRV site at the 5` end

[1] The primer contains a NcoI site at the 5' end

[2] The primer contains a BglII site at the 5' end

[3] The 4th nucleotide from the 3' end of the primer was changed from C to A to obtain a stop codon

[4] The oligo has a XbaI site at the 5' end and a SalI site at the 3' end

[5] The oligo has a SalI site at the 5' end and a XbaI site at the 3' end

*positions are shown relative to the transcription start site. + and - before the brackets indicate lower or upper DNA strand, respectively.

3.8. Crude extracts of cellular proteins

Cellular extracts were made from the *E. coli* strain DH5α(pBAD24) as control, and *E. coli* DH5α derivatives carrying pBM135, pBM136, pBM137, pBM138, pBM139, pBM140, and pBM141 plasmids described in Table 3. The strains were grown to OD_{600}=0.8 at 37ºC in LB medium, supplemented with ampicillin. The arabinose promoter of pBAD24 vector was induced by adding 1% L-arabinose to the cultures and the strains were grown for additional 30 minutes. Cells were harvested, sonicated and centrifuged to remove cell debris. After addition of glycerol to obtain 10% final concentration, supernatants were stored at -80ºC. Protein concentration of the samples was measured by the Bradford assay (Bio-Rad).

3.9. Gel mobility shift assays

Crude extracts were diluted to obtain various protein concentrations (as indicated in Results) by using the binding buffer (0.1 M Tris HCl, pH 7.4, 2 mM EDTA, 1 mM dithiothreitol, 5 mM $MgCl_2$, 5% glycerol). To inactivate nucleases, the samples were heated at 65°C for 3 min. Mixtures were cooled to room temperature, then 2 µg of herring sperm DNA and 3 µl labelled DNA fragment were added to 10 µl crude extract (Marianovsky et al. 2001). DNA fragments for gel mobility shift assays were obtained from plasmids pBM24, pBM170, pBM171 and pBM172 by XhoI-BamHI digestions and the fragments were labelled with α[^{32}P]-dCTP using Klenow polymerase. When oligonucleotides (123-126) were applied as target DNA, the samples were end-labelled with γ[^{32}P]-ATP using T_4 polynucleotide kinase.

Binding reactions were carried out at room temperature for 10 min, then the samples were loaded onto 4% or 8% native polyacrylamide gels and run in TBE buffer (pH=8) at 150 V at 4ºC.

3.10. Transcription start determination

Total RNA was extracted from *S. meliloti* strain BM175 by using the RNeasy Mini kit (Qiagen). Primer extension experiments were carried out with Tth DNA polymerase (Promega). Primer 7, which corresponds to the sequence located within the 5' end of the *lacZ* gene on the vector plasmid pSEM211, was end-labelled with $\gamma[^{32}P]$-ATP by using T_4 polynucleotide kinase. The reaction mixture (50 µl) contained 2 pmoles of end-labelled primer, 46 µg total RNA from *S. meliloti* strain BM175, 0.2 mM of each dNTP, 2 M betaine, 1mM $MnCl_2$ and 15 units of Tth DNA polymerase. The reaction was performed at 75ºC for 30 min and terminated by adding EGTA (final concentration of 20mM). The extended products were resolved on a denaturing 6% polyacrylamide gel. Sequencing reactions as controls were performed on plasmid pBM174 with primer 7 and run on the gel parallel to the primer extension reaction. The gel was analyzed using Storm 840 Phosphorimager (Molecular Dynamics).

3.11. DNase I footprinting

Binding reactions were carried out as described for gel mobility shift assays. Various amounts of cellular extracts, calculated according to the binding efficiencies determined previously were added to the reaction mixtures. After incubation for 10 min, reaction mixtures were treated with 0,2 U of DNase I at room temperature for 3 minutes, and reactions were inactivated at -80ºC. For the lower strand, a 224 bp DNA fragment (obtained by PCR amplification with primers 117 and PPR), and for the upper strand, a 141 bp DNA fragment (obtained by PCR amplification with primers 116 and PPR) were used and labelled with $\gamma[^{32}P]$-ATP by using T_4 polynucleotide kinase. Markers were generated by carrying out a Sanger AG-only sequencing reaction. The products were separated by electrophoresis on a 6% sequencing gel and visualized by autoradiography.

4. Results

4.1 Identification of additional putative TA loci encoded by the genome of *Sinorhizobium meliloti* 1021

Based on protein homologies, domain architectures and gene neighborhood analysis we could identify in the genome of *S. meliloti* strain 1021 17 TA modules belonging to different gene families (Figure 7). The *S. meliloti* 1021 genome consists of three replicons: one large replicon of 3.65 Mb and two smaller replicons, pSymA and pSymB, of 1.35 and 1.68 Mb, respectively. The complete sequencing of the genome of *S. meliloti* 1021 followed by the interpretation of the structural and functional features of the three replicons, revealed that the the genes located on pSymB considerably extend the metabolic capabilities of the microbe by allowing it to metabolize a large variety of small compounds encountered in the soil or in the plant rhizosphere and they increase the capacity of synthesizing surface polysaccharides, including exopolysaccharides, lipopolysaccharides, capsular polysaccharides and cyclic β-glucans (Galibert et al., 2001). The sequencing of pSymB revealed that 12% of the genes encoded by this plasmid may be involved in polysaccharides biosynthesis. These compounds are crucial for successful plant infection, possibly by suppressing plant defense responses (Viprey et al., 2000). Moreover, the acquisition of pSymA led to the emergence of nodulation, as well as the bacterium's capacity to colonize the low-oxygen environment of the nodule. pSymA also expanded the capacity to metabolize nitrogen compounds of various chemical forms, including molecular dinitrogen.

As it shown in Figure 7 and Table 6, members of the TA gene families are present on each of the three replicons of *Sinorhizobium meliloti* 1021.

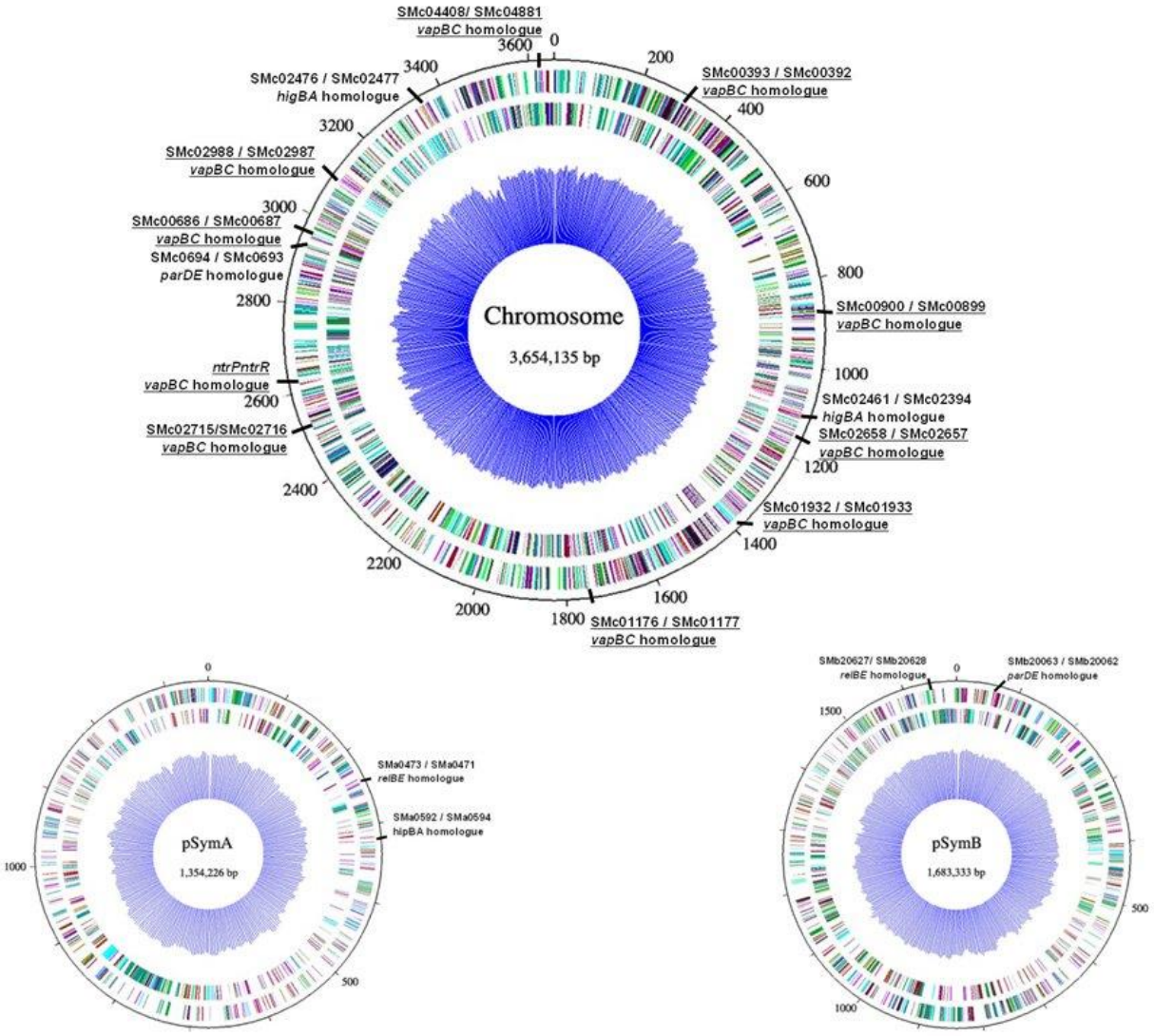

Figure 7 Localization of the complete TA modules identified in the genome of *Sinorhizobium meliloti* 1021

In Table 6 the primary accession numbers, the conserved domains of the toxin and antitoxin proteins, the names and the localization of the ORFs encoding these proteins are listed for all 17 putative TA modules identified in our genome search. The chromosome of the *Sinorhizobium meliloti* bacterium encodes 13 of them, far more than its two megaplasmids, pSymA and pSymB that encode only four.

Of all 7 gene families described by Gerdes and his coworkers (2005), only 4 are represented in this organism. Modules belonging to the *ccdAB, mazEF* or *Phd/Doc* are not encoded by the genome of *Sinorhizobium meliloti* 1021.

It is interesting to note that 10 of the 17 modules are members of the *vapBC* gene family, and all of them are encoded by the bacterial chromosome. In our search we identified five additional proteins that contained a PIN domain, characteristic for the VapC toxins. All

the genes encoding these putative proteins are located either on pSymA or pSymB megaplasmids.

Table 6. Complete TA loci identified in the genome of *Sinorhizobium meliloti* 1021

	Toxin			Antitoxin		
Chromosomal						
TA gene family	**Protein***	**Domain/** Protein family[#]	**ORF name**[^]	**Protein***	**Domain/** Protein family[#]	**ORF name**[^]
vapBC	Q9L380	PIN	SMc01521	Q9L381	SpoVT_AbrB	SMc03949
	Q92PV8	PIN	SMc01177	Q92PV7	SpoVT_AbrB	SMc01176
	Q92KH9	PIN	SMc02657	Q92KI0	SpoVT_AbrB	SMc02658
	Q92KY5	PIN	SMc04881	Q92KY4	SpoVT_AbrB	SMc04408
	Q92ME0	PIN	SMc00687	Q92MD9	SpoVT_AbrB	SMc00686
	Q92LY3	PIN	SMc02987	Q92LY2	PSK TF	SMc02988
	Q92RQ5	PIN	SMc00900	Q92RQ4	PSK TF	SMc00899
	Q92QN4	PIN	SMc01932	Q92QN3	RHH	SMc01933
	Q9X7L4	PIN	SMc02715	Q9X7L3	Wing-helix PSK TF	SMc02716
	Q92SQ4	PIN	SMc00393	Q92SQ5	Phd/YefM	SMc00392
higBA	Q92R61	Plasmid killer	SMc02461	Q92R62	HTH	SMc02394
	Q92LJ2	Plasmid killer	SMc02476	Q92LJ3	HTH	SMc02477
parDE	Q92ME5	Plasmid stability	SMc00694	Q92ME4	RHH	SMc00693
Plasmid located						
relBE	Q930F0	RelE/StbE	SMa0473	Q930F1	RelB/RHH	SMa0471
	Q92TH9	Plasmid stability	SMb20627	Q92TH8	CopG	SMb20628
hipBA	Q92ZY8	HipA	SMa0592	Q92ZY7	HTH	SMa0594
parDE	Q92X99	Plasmid stability	SMb20063	Q92XA0	RHH	SMb20062

* Protein primary accession number

[#] In the Pfam database we could not identify a conserved domain for all the listed proteins, but we could classify them in either Plasmid killer or Plasmid stability family of proteins which are members of the Plasmid toxin-antitoxin system clan.

[^] The letters a, b or c in the name of the ORF, indicate the genomic localization of the given ORF: on pSymA, pSymB or chromosome, respectively

Table 7. Solitary PIN domain containing proteins encoded by the two megaplasmids of *Sinorhizobium meliloti* 1021

ORF name	Locus name	Protein primary accession number
SMa0453	RA0238	Q930F9
SMa2253	RA1214	Q92XN0
SMa0981	RA0529	Q92ZF7
SMa2231	RA1203	Q92XP1
SMb21651	RB1571	Q92TF3

Whether all these TA modules are functional or silent, what is there function or why they are present in such a high number in the genome of this bacterium are still open questions.

During our work we focused on the characterization of the *ntrPR* operon, in order to prove that it represents an active toxin-antitoxin system in *Sinorhizobium meliloti*.

4.2 Autoregulation of the *ntrPR* operon

4.2.1 Expression of the *ntrPR* operon

By comparing the β-galactosidase activities expressed from an *ntrR-lacZ* fusion in wild type strain *S. meliloti* 1021 and its *ntrR* mutant derivative (*S. meliloti* 399), we have shown previously that the NtrR protein may repress its own transcription (Oláh et al. 2001).

If the *ntrPR* operon functions as a TA module, this observation may be due to the negative autoregulatory properties of NtrP and NtrR acting on their promoter as a protein complex.

To determine if NtrP is also involved in the autoregulation, we compared the β-galactosidase activities expressed from the *ntrP-lacZ* transcriptional fusion vector, pBM90 in *S. meliloti* 1021 and its *ntrP* (*S. meliloti* 866), *ntrR* (*S. meliloti* 399) and *ntrPR* (BM111) mutant derivatives (Figure 8).

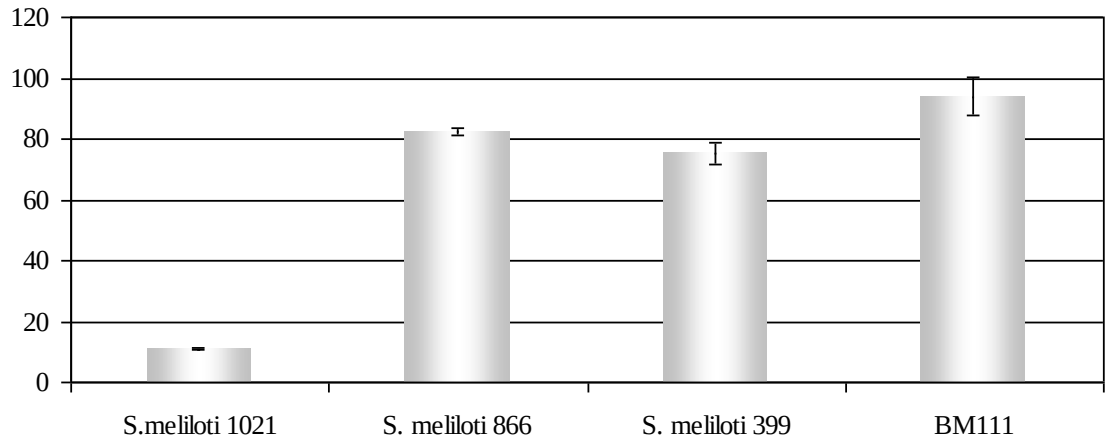

Figure 8 Regulation of the expression of *ntrPR* operon

β-galactosidase activities expressed from the *ntrP-lacZ* fusion (pBM90) were determined in the wild type strain *Sinorhizobium meliloti* 1021, and in mutant strains: *S. meliloti* 866, *S. meliloti* 399 and BM111. Enzyme activities were calculated as Miller units from parallel samples of 3 independent measurements.

If the ntrPR autoregulation is similar to that of a toxin-antitoxin module, where the antitoxin protein is responsible for DNA binding, one could expect that the NtrP protein alone inhibits the translation of the operon by binding to its own promoter. In this case the *S.*

meliloti 399 strain that encodes an intact NtrP protein should present a lower promoter activity than *S. meliloti* 866 which was obtained by a Tn5 insertion in the *ntrP* gene, or than BM111 from which the *ntrPR* operon was eliminated.

However, the promoter activity observed in the mutant derivatives did not meet these expectations. In all three *ntrPR* mutant strains the promoter activity was higher than in the wild type strain.

The strain *S. meliloti* 866 could not provide further insights in the autoregulation of the *ntrPR* operon. Since the promoter activity in this strain is comparable to that observed in the *ntrPR* deletion mutant, we concluded that most probably the Tn5 insertion in the *ntrP* gene affected also the transcription of *ntrR*.

To further investigate this regulation, we performed electrophoretic mobility shift assays (EMSA).

4.2.2 Transcription start point determination

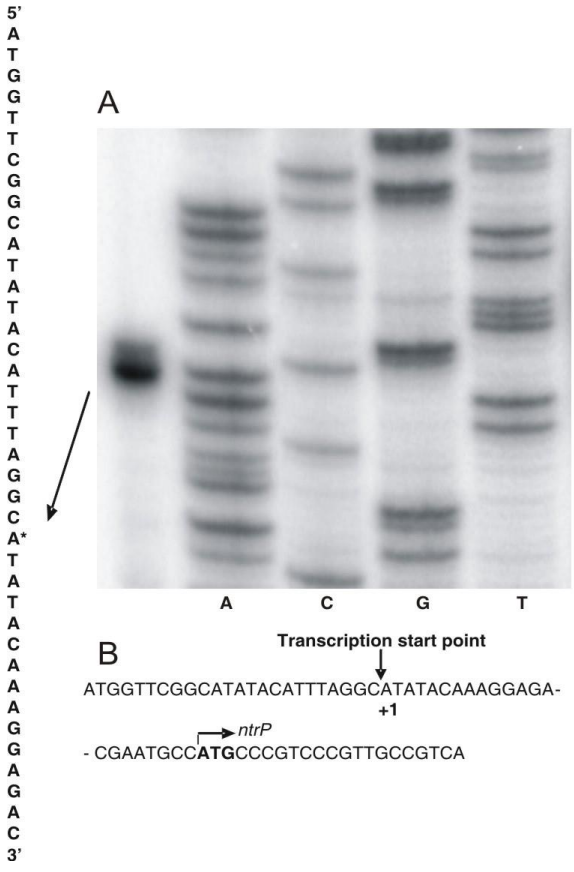

Figure 9 Determination of the transcriptional start point of the *ntrPR* operon.

A: lane 1: primer extension reaction using the total RNA of the strain BM175; lanes 2-5: sequencing reactions. The position of the band corresponding to the 5' end of *ntrP* transcript is indicated by an arrow and an asterisk in the DNA sequence (left side).
B: nucleotide sequence of the *ntrPR* promoter region. The arrows show the transcriptional starting site and the start codon of *ntrP*.

We first determined the transcriptional starting point of the *ntrPR* operon by primer extension analysis. A 397 bp fragment of the *ntrPR* promoter region (obtained by PCR amplification using the primers PPF and PPR; Table 5 and Figure 12) was cloned in the pSEM211 vector and the resulting plasmid (pBM174) was mobilized into *S. meliloti* 1021 wild type strain by triparental conjugation. Total RNA was isolated from the resulting strain BM175, which was subjected to primer extension using primer 7 (Table 5). The signal obtained indicates that the transcriptional start point is the A nucleotide located 22 bp upstream of the translational start codon of the *ntrP* gene (Figure 9).

4.2.3 Binding of NtrP and NtrR to the *ntrPR* promoter

We examined the binding abilities of NtrP and NtrR proteins to the 397 bp segment of the *ntrPR* promoter region in EMSA experiments. The promoter fragment cloned in plasmid pBM24 was recovered by XhoI-BamHI digestion and labelled by Klenow polymerase using $\alpha[^{32}P]$-dCTP, then exposed to crude extracts containing NtrP, NtrR or both.

According to experimental data presented in the literature, the complex formed between the toxin and the antitoxin acts as the repressor of the transcription of TA operon. Only the antitoxin interacts directly with DNA, but the toxin introduces cooperativity in the system (Dao-Thi et al., 2002, Gotfredsen et al., 1998, Zhang et al., 2003b, Marianovsky et al., 2001, Magnuson et al., 1998).

The sequence analysis of the NtrP protein revealed that this protein shows homology to the SpoVTAbrB DNA binding domain between aminoacids 15 and 65. To detect the promoter binding ability of this protein, crude extracts from *S. meliloti* 1021 and its *ntrP, ntrR* and *ntrPR* mutant derivatives were incubated with the DNA fragment carrying the promoter region. In the *S. meliloti* 399 strain the *ntrP* gene is intact, since the Tn5 is inserted near to the 3' end of the *ntrR* gene. Therefore, we assumed that the reaction performed with this crude extract will result in the formation of an NtrP protein - promoter DNA complex.

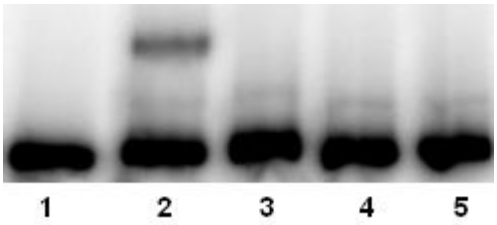

Figure 10 DNA binding ability of NtrP, NtrR and their complex with the *ntrPR* promoter

The *ntrPR* promoter fragment (397 bp), recovered by XhoI-BamHI digestion from pBM24 was incubated with the following crude cell extracts: lane 1: free DNA fragment; lane 2: 10µg extract of *S. meliloti* wild type *1021*; lane 3: 10µg extract of *S. meliloti* 866 insertional *ntrP* mutant; lane 4: 10µg extract of *S. meliloti* 399 insertional *ntrR* mutant; lane 5: 10µg extract of *S. meliloti* BM111 *ntrPR* deletion mutant.

In contrast to this expectation, we were able to detect a difference in gel mobility only in the case when the promoter DNA fragment was exposed to crude extracts containing both NtrP and NtrR proteins (Figure 10). These results suggest that both proteins are required for the formation of a stable protein-DNA complex.

To confirm this hypothesis and to overcome the problem of low expression of the *ntrPR* operon in *Sinorhizobium meliloti*, recombinant plasmids were constructed that carried the *ntrP* gene (pBM136), or the *ntrR* gene (pBM135), or the whole *ntrPR* operon (pBM138) under the control of the arabinose inducible promoter of the vector pBAD24 (Guzman et al. 1995). These constructs were introduced into *E. coli* DH5α, and the resulting strains (BM136, BM135 and BM138, respectively) were used to obtain crude cell extracts containing the corresponding proteins.

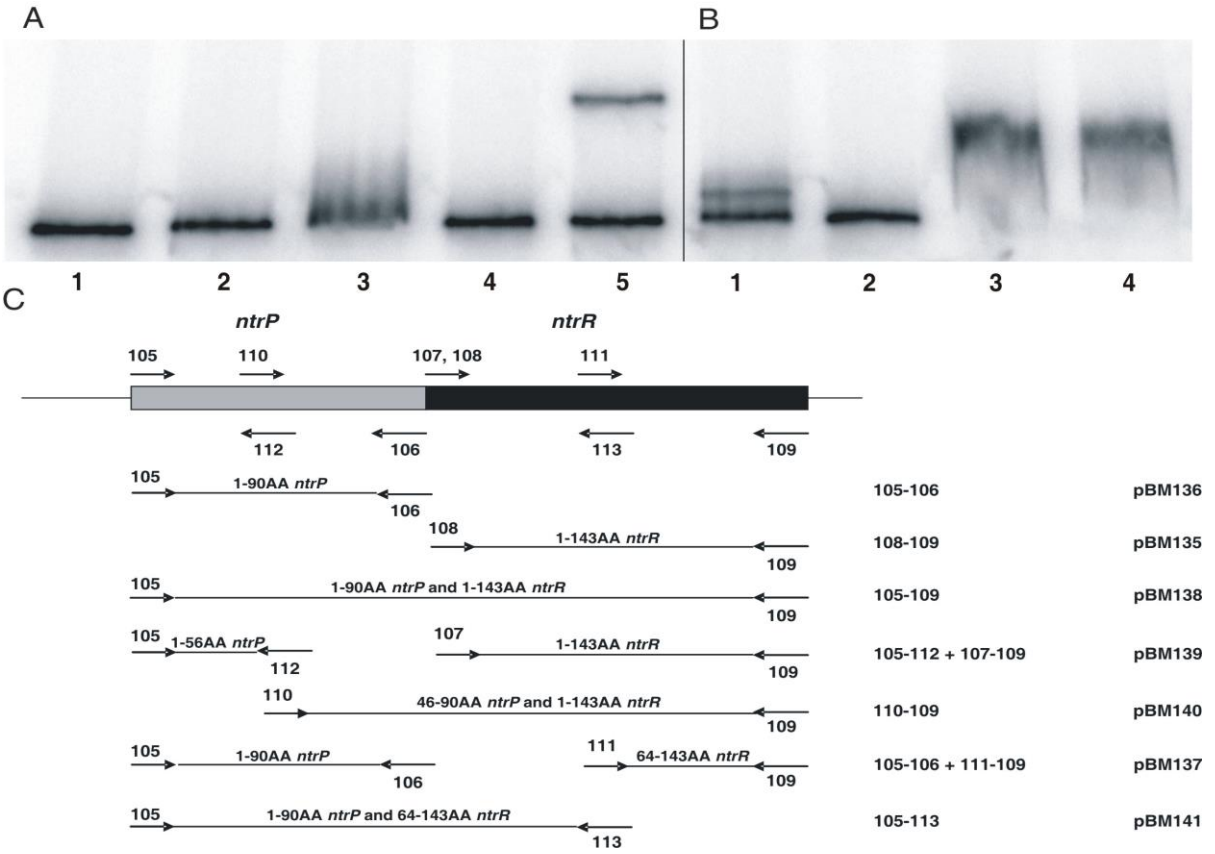

Figure 11 The binding ability of NtrP and NtrR (full length and truncated proteins) to the *ntrPR* promoter region.

The *ntrPR* promoter fragment (397 bp), amplified by PPF-PPR primers was exposed to crude cell extracts from *E.coli* DH5α carrying the constructs as indicated in C. A: lane 1: free DNA fragment; lane 2: 1µg extract of *E.coli* DH5α(pBAD24); lane 3: 0.5 µg extract of *E.coli* DH5α(pBM136); lane 4: 0.5 µg extract of *E.coli* DH5α(pBM135); lane 5: 0.05 µg extract of *E.coli* DH5α(pBM138). B: The 397 bp fragment was incubated with the following extracts: lane 1: 0.5 µg extracts of *E.coli* DH5α(pBM139); lane 2: 0.5 µg extract of *E.coli* DH5α(pBM140); lane 3: 0.5 µg extract of *E.coli* DH5α(pBM141); lane 4: 0.5 µg extract of *E.coli* DH5α(pBM137). C: schematic representation of the *ntrPR* operon. Grey and black bars represent the *ntrP* and *ntrR* genes, respectively. Arrows with numbers demonstrate the site and orientation of primers used for amplification. The lines between the arrows show the corresponding parts of *ntrP* or *ntrR* genes; above the lines the length of the proteins is indicated (AA: amino acid). The resulting recombinant plasmid constructs are listed on the right.

To confirm that other proteins present in *E. coli* cells are not able to recognize the *ntrPR* promoter, a control binding reaction with crude extracts of *E. coli* containing the empty pBAD24 vector was performed (Figure 11A, lane 2). As expected, a retarded band was obtained only when the promoter DNA was incubated with the extract containing both the NtrP and NtrR proteins (Figure 11A, lane 5). When only the NtrP protein was present in the crude extract, a protein-DNA complex was also obtained, but the stability of the complex was very weak (Figure 11A, lane 3). Although the concentration of the extract was 10-fold higher than in lane 5, the protein-DNA complex remained unstable. Finally, when the crude extract containing the NtrR protein was incubated with the promoter fragment, no retarded complex appeared in the gel (Figure 11A, lane 4).

The requirement of a 10 times higher concentration of crude extract (containing the overexpressed NtrP protein) for the detection of the labile protein-DNA complex than for detecting an NtrPR-promoter complex (Figure 11), can give an explanation for failing to observe any shift with the *Sinorhizobium meliloti* 399 crude extract, that expressed the NtrP at normal physiological concentrations (Figure 10).

Data obtained on various TA systems proved that the N- and C-terminal parts of antitoxin and toxin proteins play different roles in DNA-protein and protein-protein complex formation and in toxicity (Bernard and Couturier 1991; Ruiz-Echevarria et al. 1995; Zhang et al. 2003b). The presence of an AbrB-like domain in the N-terminal region of NtrP suggested that this part of the protein may be responsible for binding to the promoter DNA, whereas the second half of the molecule may be involved in the interactions with the NtrR protein.

In order to examine this, truncated versions of the *ntrP* and *ntrR* genes encoding the N- or C-terminal regions of the corresponding proteins were constructed (Figure 11C). The truncated proteins, expressed together with the other member of the protein pair were placed under the control of the arabinose-inducible promoter in vector pBAD24. The constructs were introduced in *E. coli* DH5α and crude cell extracts were prepared for EMSA experiments. The 397 bp fragment carrying the *ntrPR* promoter region, prepared and labelled as in the experiment in Figure 10A was used for binding assays. When the N-terminal part of NtrP (amplified region between primers 105 and 112, representing the residues 1-56) was expressed together with the full-length NtrR (amplified region between primers 107 and 109) from plasmid pBM139, a complex was formed with the promoter fragment (Figure 11B, lane 1). The migration distance differed from that of the complex obtained with full-length NtrP and NtrR (Figure 11A, lane 5), which can be explained by the reduced size of the NtrP protein. Another construct on plasmid pBM140 carried the C-terminal part of NtrP (residues 46-90, plus the full-length NtrR, amplified region between primers 110 and 109). The lack of

a retarded band when using the cell extract of *E. coli* with pBM140 also supports the major role of the N-terminal part of the NtrP protein in the interaction with promoter DNA (Figure 11B, lane 2). Two further constructs contained the full-length NtrP and the C- or N-terminal part of NtrR in plasmids pBM137 and pBM141, respectively. pBM137 was obtained by cloning NtrP (amplified region between primers 105 and 106) and a 79 amino acid long C-terminal region of NtrR (amplified region between primers 111 and 109). pBM141 was constructed by using the primers 105 and 113, therefore the N-terminal 64 residues of NtrR were present on the plasmid. When *E. coli* cell extracts carrying proteins expressed from either of these plasmids were incubated with the promoter DNA fragment, the resulting complexes were less stable than those obtained with the full-length NtrR protein (Figure 11B, lanes 3 and 4, compared to Figure 11A, lane 5). This suggests an important role of both parts of NtrR in stabilizing the DNA-NtrP-NtrR complex, or in maintaining the correct structure of the protein required for a probable oligomerization.

4.2.4 Determination of the exact binding site of the NtrPR complex in the *ntrPR* promoter

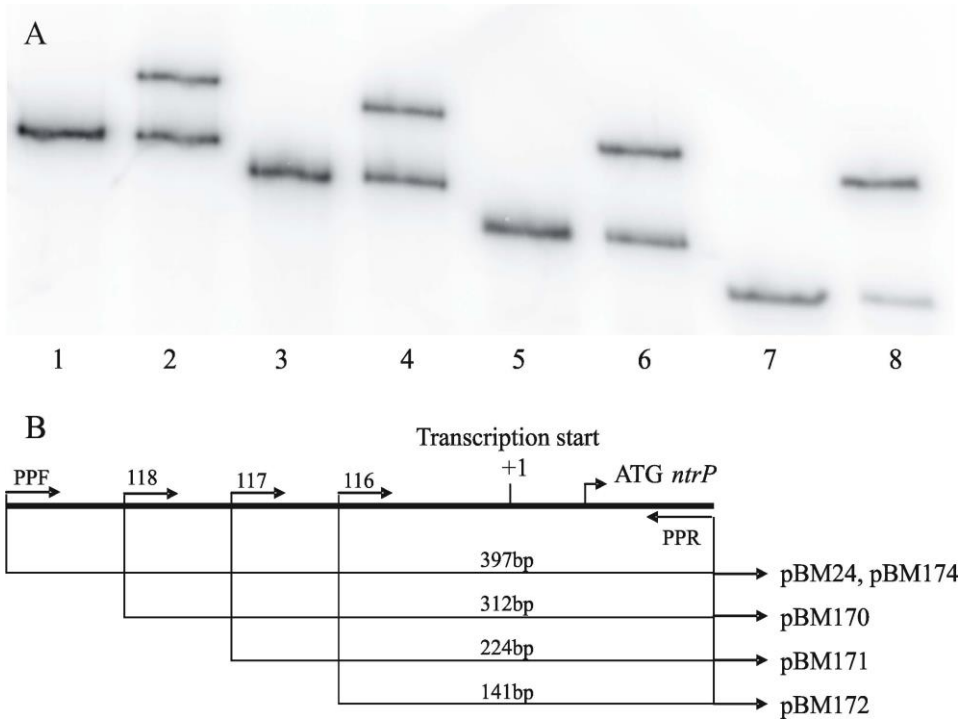

Figure 12 Binding of NtrPNtrR to the *ntrPR* promoter region carried by DNA fragments of different length.

A: lanes 1, 3, 5 and 7: free DNA fragments of 397 bp, 312 bp, 224 bp, and 141 bp, respectively. Lanes 2, 4, 6, and 8: the same DNA fragments incubated with 0.5µg E. coli DH5α(pBM138) cell extract containing NtrP and NtrR proteins.
B: schematic representation of the *ntrPR* promoter region. The thick line represents the promoter region of the *ntrPR* operon, with the transcription and translation start points shown above the line. The arrows indicate primers used to obtain promoter fragments of various lengths. Thin lines represent the shortened fragments, and on the right, the plasmids carrying these fragments are listed.

To localize the binding sites for the autoregulatory proteins in the *ntrPR* promoter region, first we determined a shorter region necessary for the binding of the complex. In addition to the 397 bp fragment (cloned in plasmid pBM24) the following regions were amplified and cloned (Figure 12B): a 312 bp fragment obtained with primers 118 and PPR (plasmid pBM170), a 224 bp fragment between primers 117 and PPR (plasmid pBM171), and a 141 bp fragment between primers 116 and PPR (plasmid pBM172). The cloned fragments were recovered from these plasmids by XhoI-BamHI digestion, labelled by Klenow polymerase with $\alpha[^{32}P]$-dCTP and used in EMSA experiments. All of these DNA regions were fully active in complex formation with the *E. coli* extract containing the NtrP and NtrR proteins (Figure 12A). Therefore, the binding site for the NtrPR complex is located within the 141 bp region. This region covers an 83 bp region upstream and a 58bp region downstream of the transcription initiation site of the *ntrP* gene.

To define the exact binding site, DNase I footprint experiments were carried out using 141 bp and 224 bp DNA fragments for the lower and upper strand reactions, respectively. When, prior to DNase I digestion, increasing amounts of *E. coli*(pBM138) crude extracts (carrying both NtrP and NtrR under the control of the arabinose inducible promoter) were added to the DNA fragments, a 31 bp sequence of the upper strand and a 34 bp sequence of the lower strand were protected by the protein complex (Figure 13). This segment of DNA covers the transcription start site and is located close to the putative ribosome binding site GGAGA. The nucleotide sequence of this region revealed a direct repeat sequence 5'-GGCATATACA-(N4)-GGCATATACA-3'.

The data presented in Figure 13 indicate that these exact direct repeats are involved in the binding of the repressor complex. Interestingly, a palindrome sequence located 9 bp further upstream of the direct repeats is not protected by NtrP and NtrR under the conditions tested.

In order to investigate whether both direct repeats are required for protein binding, two oligonucleotide pairs were synthesized. One pair, 123 and 124 (35 bp) covered the complete repeated sequence (plus 7 nucleotides upstream and 11 nucleotides downstream, Table 5).

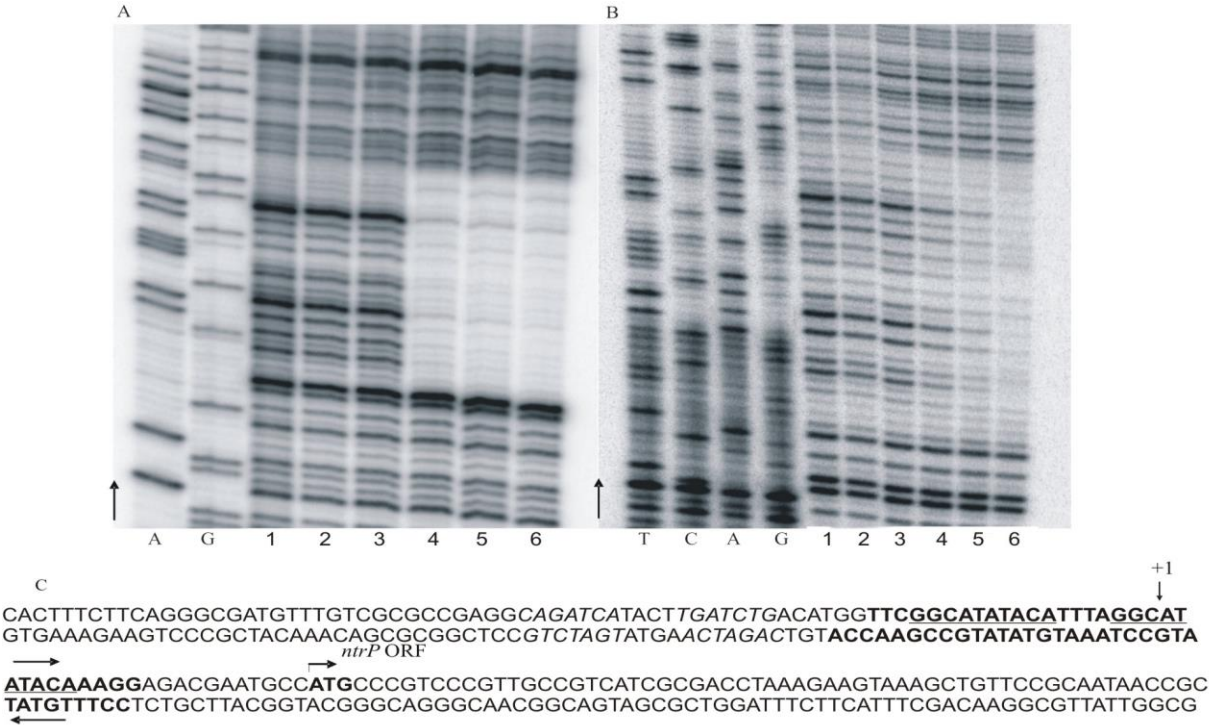

Figure 13 Binding sites for the NtrPR complex in the promoter region of the *ntrPR* operon.

A and **B**: protection of the lower and the upper DNA strand, respectively, by increasing amounts of crude extract of strain *E. coli* DH5α(pBM138). Lanes labelled with A, G, T, C show sequencing reactions. Lane 1: no extract was added prior to DNase I digestion; lanes 2-6: 0.1, 0.2, 0.5, 0.7, and 1.2μg protein extracts were added to protect DNA. A negative control reaction performed by adding 1 μg protein extract of the strain *E. coli* DH5α(pBAD24) prior to DNaseI digestion, resulted in the same pattern as shown in lane 1 of **A** and **B** (data not shown).

C: the nucleotide sequence of the *ntrPR* promoter. Arrows indicate the transcription start site (+1) and the start codon of *ntrP*. Regions protected by protein binding are marked with bold letters. The two direct repeats are underlined. The palindrome sequence upstream of the protected region is shown by italic letters.

* The arrows from the footprint image and those upstream of the ATG indicate the nucleotide sequence reading direction

The other oligonucleotide pair, 125 and 126 (30 bp) carried only the second half of the direct repeat (plus 4 nucleotides upstream and 16 bp downstream). The oligonucleotide fragments were labelled by T_4 polynucleotide kinase using $\gamma[^{32}P]$-ATP. After incubating the labelled oligonucleotide fragments with extracts of *E. coli* (pBM138), complex formation was observed in both cases (Figure 14), indicating that one of the direct repeated sequences (10 bp) is sufficient to accommodate the repressor complex.

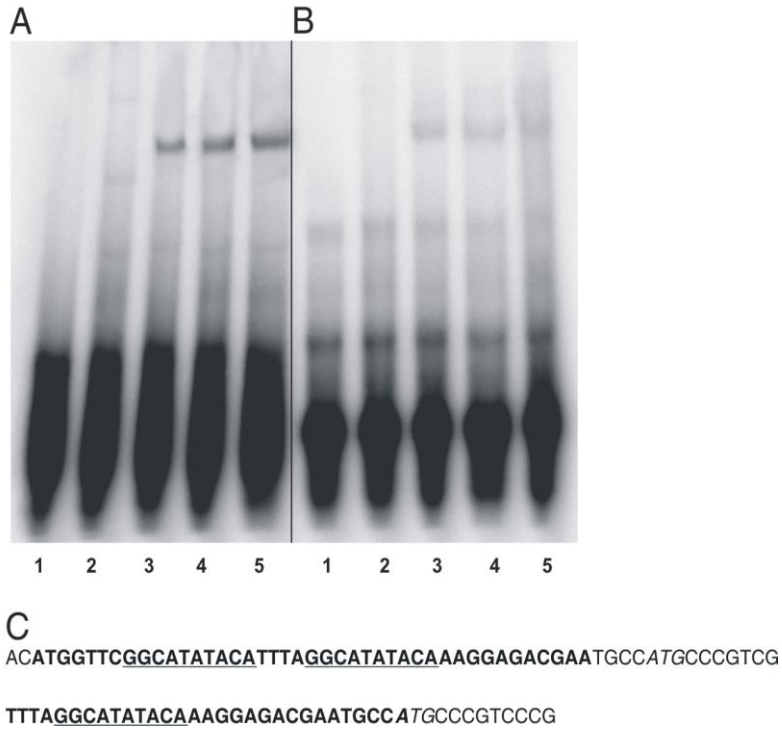

Figure 14 The role of direct repeats in the binding of NtrPR complex.

A: oligonucleotides 123-124 covering both direct repeats were used as DNA target incubated with extracts of *E.coli* DH5α(pBM138). Lane 1: no extract added, lane 2: 3.0 µg extract of *E. coli* DH5α(pBAD24), lanes 3-5: 1.0, 3.0, and 5.0µg extract of *E.coli* DH5α(pBM138) was used in binding reactions.
B: oligonucleotides 125-126 representing only one of the repeated sequences were incubated with cell extracts. Lane 1: free DNA; lane 2: 3.0 µg extract of *E. coli* DH5α(pBAD24), lanes 3-5: 1.0, 3.0, and 5.0µg extract of *E.coli* DH5α(pBM138) was applied.
C: bold letters in the nucleotide sequences represent the oligonucleotides of the upper strand used in the experiments in panels **A** and **B**; underlined letters show the direct repeats involved in DNA-protein interaction.

4.3 The function of the *ntrPntrR* module of *S. meliloti*

4.3.1 Toxicity of the NtrR protein

In order to examine the function of NtrP and NtrR, we tested the growth and viability of *E. coli* derivatives carrying plasmids with *ntrR*, *ntrP*, or both genes (plasmids pBM135, pBM136 and pBM138, respectively) controlled by the arabinose inducible promoter (Figure 15A). The expression of NtrP, or NtrP and NtrR together did not affect the growth of *E. coli* cells, since identical optical densities of these cultures and the control cells carrying the pBAD24 vector were observed. About two hours after the induction of NtrR protein expression, the optical density of the culture was lower than those of the other cultures. By the end of the experiment (after 23 hours) an optical density more than 2,5-fold lower than the controls was detected for the sample expressing NtrR. To determine the number of viable cells, samples from the cultures were plated and colonies were counted. The difference in living cell numbers resulted in a final decrease of about 4 orders of magnitude (Figure 15B).

These results demonstrate that the NtrR protein profoundly affects cell viability, and this toxic effect can be prevented by the presence of the NtrP protein.

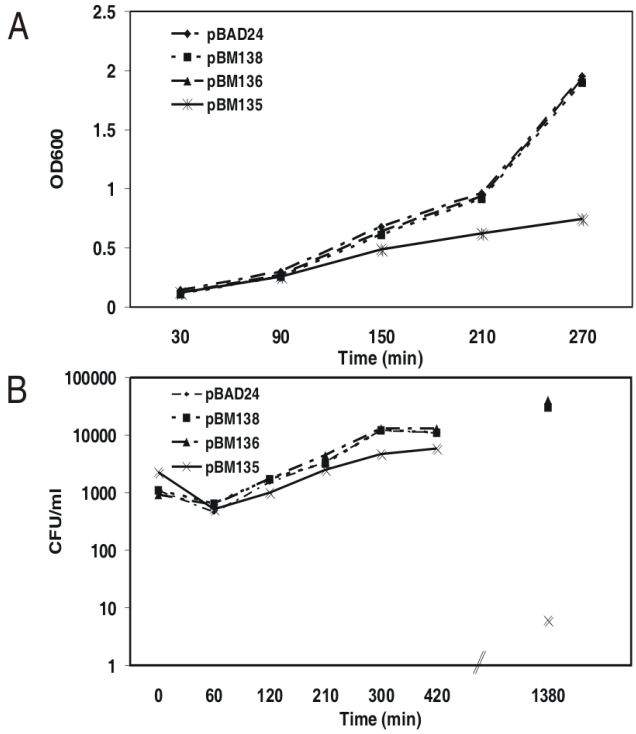

Figure 15 Cell growth and viability of *Escherichia coli* MG1655 expressing NtrP, NtrR or both proteins

Cell growth (A) and viability (B) of *E. coli* derivatives with plasmids encoding NtrP (pBM136), NtrR (pBM135), and both NtrP and NtrR (pBM138) proteins or none (pBAD24). The expression of these proteins was induced by adding 1% L-arabinose to the cultures. In panel B the time scale is interrupted after 420 min, since the last samples were tested after 23 hours. The experiments were repeated twice.

4.3.2 Response to amino acid deprivation

During stringent response elicited by amino acid starvation, guanosine 3',5'-bispyrophosphate (ppGpp) is synthesized by the RelA protein. Under such conditions the expression of *mazEF* in *E. coli* was shown to be negatively regulated by ppGpp (Aizenman et al. 1996). In contrast, the transcription of another TA system, *relBE* was strongly induced during nutritional stress, and this activation was independent of ppGpp or *relA* (Christensen et al. 2001).

The accumulation of ppGpp following amino acid starvation was also documented in the strain *S. meliloti* 1021 (Howorth and England, 1999). We examined whether the expression from the *ntrPR* promoter is dependent on the level of ppGpp. Plasmid pBM90 carrying the *lacZ* gene under the control of the *ntrPR* promoter was introduced into the wild type strain *S. meliloti 1021*, the Tn5 insertion mutant *S. meliloti* 399 and the *ntrPR* deletion derivative BM111.

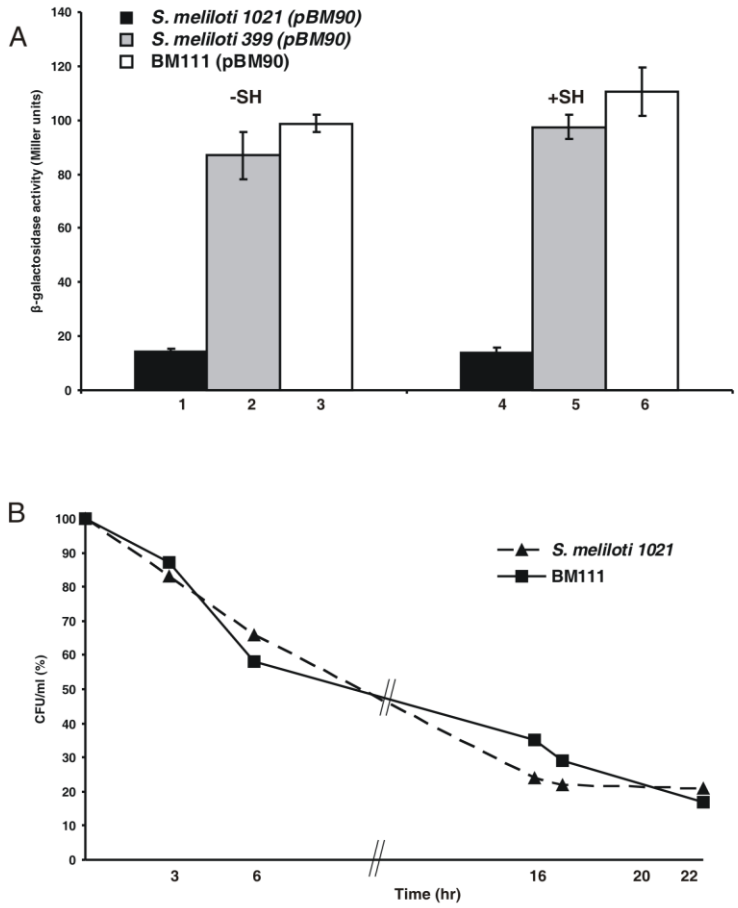

Figure 16 Response to amino acid starvation of the *S. meliloti* strains

A. Expression of the *ntrPR* operon in amino acid starved *S. meliloti* strains

β-galactosidase activity expressed under the control of the *ntrP* promoter was determined in wild type *S. meliloti* 1021 (black column), in the Tn5 insertion mutant *S. meliloti* 399 (grey column), and in the deletion derivative BM111 (empty column). Control samples (1-3) were grown without serine hydroxamate (SH), samples 4-6 were cultured with 1 mg/ml SH.

B. Viability of *S. meliloti* strains under amino acid starvation

CFU/ml values were calculated as the % of viable cell numbers determined prior to SH addition. Triangles and filled squares represent the values of wild type *S. meliloti* 1021 and the mutant strain BM111, respectively.

As expected, β-galactosidase activity was low in the wild type strain and showed a similar low level also in the amino acid starved samples (Figure 16A, samples 1 and 4, respectively). Due to the lack of negative autoregulation, the activity in the mutant strains increased, but the activities in samples incubated in the presence of serine hydroxamate (SH) (Figure 16A, samples 5 and 6) were only slightly higher than the values in the control cultures (Figure 16A, samples 2 and 3). The viability of strains *S. meliloti* 1021 and BM111 also showed a similar pattern when starved for amino acids for 22 hours (Figure 16B). Therefore, the expression of the *ntrPR* genes in free-living bacteria seems to be independent of the starvation-induced ppGpp levels.

4.3.3 Effect of antibiotics on the viability of *ΔntrPR* mutant strain

Previously it was described that antibiotics known to inhibit transcription/translation can trigger *mazEF*-mediated cell death by reducing the MazE antitoxin level (Sat et al. 2001). Three of these antibiotics, rifampicin, chloramphenicol and spectinomycin were tested on the viability of wild type *S. meliloti* 1021, and the BM111 strain carrying a deletion in the *ntrPR* operon (Figure 17). The effect of these antibiotics on the survival of wild type cells was less dramatic than that observed with *E. coli* cells (Sat et al. 2001); however, after a 10 min incubation 15-30% higher CFU/ml values were obtained with the BM111 strain than with the wild type *S. meliloti* 1021. As a control, a similar experiment was carried out with claforan, an antibiotic inhibiting cell-wall synthesis. The number of colony-forming units decreased equally in the cultures of wild type and *ΔntrPR* mutant strains (Figure 17).

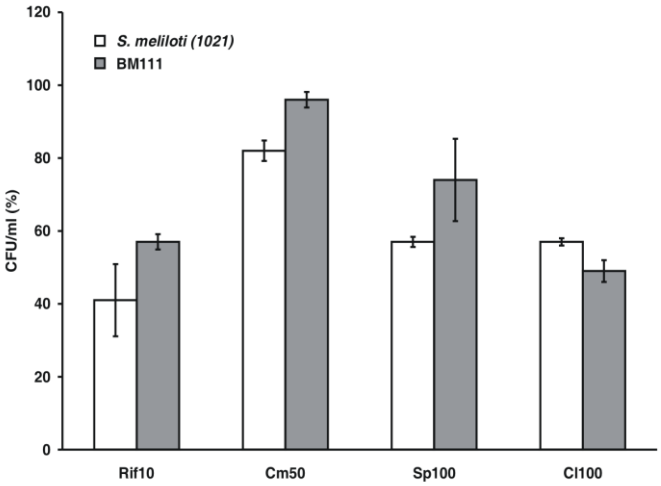

Figure 17 The effect of antibiotics on cell viability.

Wild type *S. meliloti* 1021 (empty bars) and the *ΔntrPR* mutant strain BM111 (filled bars) were treated for 10 min with rifampicin (10 μg/ml, Rif10), chloramphenicol (50 μg/ml, Cm50), spectinomycin (100 μg/ml, Sp100) and claforan (100 μg/ml, Cl100). Viable cell numbers (CFU/ml) are shown as % of the corresponding control samples incubated without antibiotics.

5. Discussion

The present work demonstrates that the *ntrPR* operon of *S. meliloti*, which is structurally similar to bacterial toxin-antitoxin modules, also exhibits the characteristic autoregulatory circuit of such modules. In addition, the experiments revealed new information concerning the *vapBC* gene family and provided data about the physiological role of the TA loci belonging to this family. On the other hand, the *ntrPR* operon as a TA module represents a novelty in the field of symbiotic nitrogen fixation since it is the first description of a functional toxin-antitoxin system in *Rhizobiaceae*.

5.1 Autoregulation of the *ntrPR* operon

The autoregulatory functions of NtrPR have been investigated by using different methods: measurements of the *ntrPR* promoter activity, electrophoretic mobility shift assays and DNase I footprinting.

The mechanism of autoregulation seems to be general for TA modules, located either chromosomally or on plasmids: the antitoxin protein carries a DNA-binding motif and interacts with the promoter region with variable affinities (reviewed by Engelberg-Kulka and Glaser 1999; Gerdes et al. 2005). In the case of the *ntrPR* operon, we observed that the *ntrPR* promoter activity is inhibited in the wild type strain where both NtrP and NtrR proteins are expressed, but the inhibition is released in the *ntrPR* mutant derivatives (Figure 8). Our results obtained from electrophoretic mobility shift assays showed that NtrP protein that presents homology to the SpoVT/AbrB family of DNA binding proteins, is able to recognize a DNA segment in the promoter region, but the binding is weak and results in an unstable DNA-protein complex (Figure 11A).

In the case of nearly all studied TA modules, the toxin, unable to bind DNA independently, enhances the stability of the antidote-DNA complex. Therefore, for the full transcriptional repression of the operon both proteins are required. As an exception, ParD of the RK2 plasmid was found to be sufficient for full transcriptional regulation of the *parDE* operon (Eberl et al. 1992; Roberts et al. 1993).

The N-terminal and C-terminal parts of toxins and antidotes play different roles in protein-protein and protein-DNA interactions. In EMSA experiments we examined various

truncated proteins containing the N- or C-terminal half of the toxin or antitoxin proteins expressed together with the other full-length component of the module (Figure 11B). Our results demonstrate that the DNA-binding domain resides on the N-terminal region of antitoxin.

The binding site for the NtrP protein in the promoter region was delimited. Even the shortest fragment of 141 bp exhibited the same protein binding ability as was determined for the 397 bp fragment (Fig. 12A). The 141 bp fragment comprises a 61 bp region upsteam of the transcription start site, a 22 bp sequence between the transcription start and the translation start codon, and an additional 58 bp region downstream of ATG (Fig. 12B). Within this region a palindrome sequence, 5'-CAGATCA-TATC-TGATCTG-3' and a direct repeat, 5'-GGCATATACA-TTTA-GGCATATACA-3' were identified as potential binding sites. DNase I footprint experiments revealed that the direct repeat sequences are involved in the interaction with the autoregulatory NtrPR complex (Fig. 13 and 14). Since the second repeat overlaps the transcription start site, the repressor activity of the complex is very probably due to interference with RNA polymerase binding.

The role of two palindrome sequences in the binding of antitoxin-toxin complex has also been demonstrated for other TA systems, such as PemIK (Masuda et al. 1993; Tsuchimoto and Ohtsubo 1993) and PhdDoc (Magnuson and Yarmolinsky 1998). In the chromosomal *mazEF* promoter region a unique structure with several palindromes located between –34 to +6 was identified. These sequences may exist in two possible configurations, and the MazEF complex can bind to either of these structures. Interestingly, a nearby perfect palindrome was not involved in protein binding (Marianovsky et al. 2001). In the promoter region of *ntrPR* the role of the direct repeat sequences in complex formation was demonstrated, but the perfect palindrome sequence was not protected by the NtrPR complex in DNase I footprint experiment (Fig. 13). Since in the binding experiments *E. coli* extracts containing NtrR and NtrP were used, we cannot exclude that this palindrome sequence may be involved in binding a protein present in *S. meliloti* cells. Interestingly, the binding of the factor for inversion stimulation (FIS) was also observed in the promoter region of *mazEF*, upstream of the alternating palindrome sequences; and the binding of FIS activated the *mazEF* promoter (Marianovsky et al. 2001).

Preliminary data indicate that SyrM protein, a positive regulatory factor of *nod* gene expression, also participates in the control of *ntrPR*. In the absence of SyrM, a lower derepressed level of *ntrPR* was observed in *S. meliloti* 399 (Olah et al., 2001). Further experiments are required to clarify the exact role of SyrM.

5.2 Toxicity of NtrR protein

Expression of the NtrR protein resulted in the inhibition of *E.coli* cell growth and colony formation. The effect of NtrR was less drastic than that of RelE or MazF (Aizenman et al. 1996; Gotfredsen and Gerdes 1998; Pedersen et al. 2002; Zhang et al. 2003a) when expressed in *E. coli* cells. While the number of colony forming cells decreased by 3-4 orders of magnitude within 30-60 min when RelE or MazF was induced (Aizenman et al. 1996; Pedersen et al. 2002; Christensen et al. 2001), a similar decrease was detected only after 23 hours in the presence of NtrR. The weaker inhibition could be due to a difference between the action mechanisms of NtrR and *E. coli* toxins. The metabolic changes as a consequence of unfavorable environmental conditions are rather different in enteric and in soil bacteria. *S. meliloti,* as a soil bacterium, is adapted to survive in oligotrophic environments (Rozsak and Colwell, 1987) in which low oxygen tension and nutrient deprivation may limit growth in the free-living state. Soil bacteria may therefore tolerate longer periods of starvation by developing distinct mechanisms for their persistence under unfavorable conditions (Davey and de Bruijn, 2000).

5.3 Physiological role of the *ntrPR* operon in *Sinorhizobium meliloti*

The determination of the mode of activation and the mechanism of toxicity are both important to understand the possible role of TA modules in bacterial cells.

5.3.1 The activation of the *ntrPR* module

Two models were proposed for explaining the activation of a TA module in *E. coli*: a simple passive model and the active control (Gerdes et al., 2005). The passive model suggests that the induction of the TA locus occurs without the involvement of an external signal. For example, under physiological conditions that cause a reduced rate of translation, the antitoxin is synthesized at low level. On the other hand, a lower antitoxin level leads to the derepression of the promoter of TA operon, resulting in an increased transcription of the operon. If the conditions that block the reinitiation of translation at normal rates last long enough, the degradation of the labile antitoxin allows the stable toxin to attack the intracellular target. In contrast, the active control model suggests that a signal (low-molecular weight compound, a protein or a stalled ribosome) triggers the degradation of the antitoxin (either free in solution or in complex with the toxin) by a cellular protease (Figure 1).

It is also possible that TA modules belonging to different TA gene families are activated by different ways.

Experiments that could support both the active and passive activation model were performed on the *mazEF* system. It was also demonstrated that the stringent response alarmone ppGpp directly activates *mazEF* that supports the signal-initiated activation suggested by the active model.

During amino acid and carbon-source limitation, free-living bacteria downregulate the synthesis of stable RNAs by increasing the cellular concentration of the ppGpp alarmone. This small molecule directs RNA polymerase away from the synthesis of rRNA and tRNA towards the synthesis of biosynthetic operons (Chatterji, et al., 2001). The increase in the ppGpp concentration induces also the activation of Lon (DNA-binding ATP-dependent protease La) to degrade idling ribosomal proteins during starvation, generating in this way an endogenous source of amino acids for protein synthesis (Kuroda et al., 2001). The chromosomal TA modules of *E. coli* MazEF and RelBE were also shown to play a role in the regulation of macromolecular synthesis rates (Gerdes 2000, Christensen et al., 2001), but for MazEF the direct involvement of ppGpp in the regulation of the TA module was shown (Aizenman et al. 1996). This is an observation that strongly supports the active control model, where the antitoxin degradation is triggered by the signal, in this case by ppGpp, a low-molecular weight compound.

On the other hand, the *mazEF* system was shown to be activated by compounds that inhibit transcription or translation, supporting the passive activation hypothesis. Experiments performed on the *mazEF* system have shown that antibiotics known to inhibit transcription or/and translation can trigger *mazEF*-mediated cell death by reducing the MazE antitoxin level (Sat et al., 2001).

The same compounds were tested on the *S. meliloti* 1021 wild type and *ntrPR* mutant strains. Similarly to the *relBE* system, a ppGpp independent expression of the *ntrPR* operon was observed. However, these data do not exclude completely the possibility of an active control model for the *ntrPR* system. Antibiotics described as activators of the *mazEF* system, were tested also on the *ntrPR* module. The impact of these compounds on the survival of the *S. meliloti* 1021 strains was less drastic than that observed with the *mazEF* system, but can still support the passive activation theory.

Up to date, no other mechanisms that result in the induction of the TA loci have been reported.

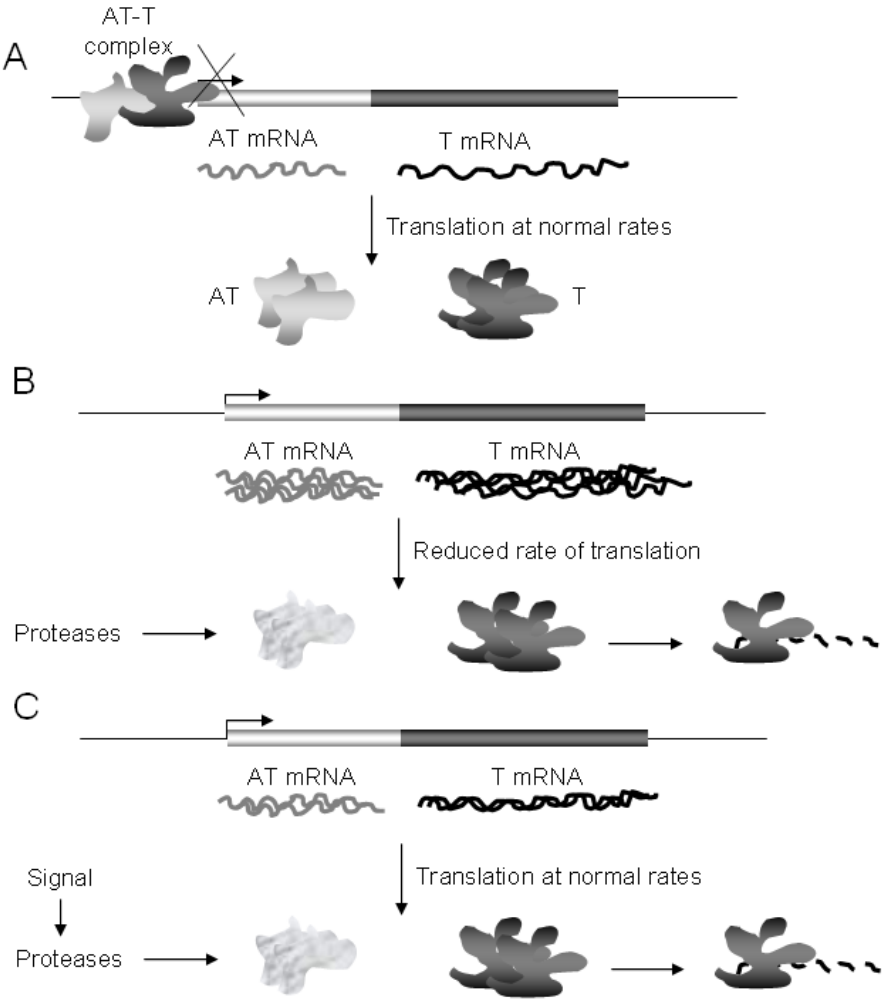

Figure 18 Schematic representation of the activation of TA modules

A. During normal physiological conditions, both transcription and translation of the TA operon proceeds at a normal rate. The antitoxin forms a complex with the toxin, blocking in this way both the toxic effect of the toxin and the transcription of the TA operon.
B. The passive activation model. Under certain physiological conditions that determine a reduced rate of translation *de novo* synthesis of the antitoxin is blocked. The labile antitoxin is degraded by cellular proteases and the intracellular target is attacked by the toxin. The antitoxin degradation results also in the derepression of the TA operon but the low rate of translation hinders the accumulation of a sufficient amount of antitoxin.
C. The active control model. The activation of the TA module is initiated by a signal that triggers specific proteases to degrade the antitoxin, either free in solution or in complex with the toxin.

5.3.2 Target of the NtrR toxin

Very few data are available concerning the toxicity mechanism of the PIN domain containing toxins. The existing bioinformatic and experimental data support the idea that the PIN domain containing proteins are exoribonucleases. Clissold and Ponting reported the homology of PIN domains to proteins involved in the non-sense mediated RNA decay in eukaryotes. In addition, data obtained from the structural analysis of PAE2754 of *Pyrobaculum aerophilum* and *in vitro* exonuclease assays performed with the same protein (Arcus et al., 2004) suggest that the PIN domain containing toxins are exoribonucleases.

Among the characterized toxin components of the TA modules only the VapC type of toxins were shown to have an exonuclease activity. The CcdB and ParE are DNA gyrase inhibitors, while many RelE and MazF homologous were shown to have endoribonuclease activity. These diferences in the type of targets of the toxin may also explain the differences in the toxicity level of NtrR compared to all other characterized toxins.

5.3.3 The role of a toxin-antitoxin system in a symbiotic bacterium

Sinorhizobium meliloti, as a symbiotic nitrogen fixing bacterium, can reveal new perspectives in understanding the functions of the TA modules. This prokaryotic organism develops symbiotic interactions with the leguminous plants, such as *Medicago, Melilotus and Trigonella.* The bacterium in its free-living state may exist under poor environmental conditions, but when enters into the plant roots, a differentiation of the bacteria takes place leading to the formation of the nitrogen fixing forms, the bacteroids. The metabolic processes of the bacteroids show drastic differences compared to that of the free living bacteria both in respiration and acquisition and utilization of the carbon supply. The enzyme responsible for the reduction of the atmospheric nitrogen is irreversibly damaged by oxygen. Due to this fact, bacteroids function in a microaerobic environment that can meet the requirements of the oxygen-sensitive nitrogenase.

The highly energy-consuming nitrogen fixation process requires a balanced nutrient supply of the bacteroid by the host plant. Carbon is provided to the differentiated bacterial cells in the form of dicarboxylic acids to fuel nitrogen fixation. In exchange, fixed nitrogen is transferred to the plant. Both the bacteroid and the plant-derived peribacteroid membrane tightly control the exchange of metabolites. In accordance with the oxygen sensitivity of the nitrogenase, oxidation of dicarboxylic acids occurs in an oxygen limited environment in bacteroids.

Since the transition from free-living to symbiotic state involves drastic metabolic changes, the TA modules present in the genome of *Sinorhizobium meliloti* can help this organism to cope with the management of the metabolic stress induced by these transitions.

The *ntrPR* operon apart from sharing remarkable similarities with TA modules of other families exhibits specific features that can be due either to its membership of the *vapBC* gene family or to its special role as a TA operon of a symbiotic nitrogen fixing bacterium. Previous experiments performed in the nitrogen fixation group seem to support the latter hypothesis.

Earlier data demonstrated higher expression levels of plant factor-induced *nod* genes under free-living conditions, and increased expression of *nif* genes of *S. meliloti* 399 in

symbiosis with the host plant when compared to wild type strain *S. meliloti* 1021 (Dusha et al. 1989; Oláh et al. 2001). Investigations using the methods of functional genomics demonstrated considerable changes in the general metabolic pattern of bacteria under symbiotic conditions. Under these conditions, the expression of a large set of transcription-translation related genes decreased to a low level. However, down-regulation was less effective in the *ntrR* mutant, resulting in a higher relative level of expression for these genes (Puskás et al. 2004).

5.3.4 Diversity of TA modules in *Sinorhizobium meliloti*

A puzzling subject concerning the TA modules is their presence in high number in certain prokaryotic genomes. Speculative explanations raise the possibility that the number of the TA loci is correlated with cell growth rate. In this way, free-living, slowly growing organisms characterized by low-translation rates benefit from having many TA loci, allowing them the optimization of quality control of gene expression (Pandey et al., 2005).

But what could be the explanation for the presence of high number of TA loci of the same family in a given organism? For example, the genome of *Mycobacterium tuberculosis* contains 38 TA loci of which 23 belong to the *vapBC* gene family, while the genome of *Sulfolobus solfataricus* contains 22 TA loci all of them belonging to the *vapBC* family.

An interesting transcriptomic study on the extremely thermoacidophilic archaeon, *Sulfolobus solfataricus* (Tachdjian and Kelly, 2006) showed that the 22 *vapBC* modules of this microorganism are involved in the heat shock response. It was a surprising observation that even if the TA loci belong to the same family, they do not react in the same way under certain conditions. This observation leads to various hypotheses:

- different loci belonging to the same family play different roles in a given organism;
- activation mechanism and toxin potency may differ from one pair to another;
- cells may use TA systems to modulate their metabolic activity even in the absence of perturbations.

Searching for TA modules in the genome of the symbiotic bacterium, *Sinorhizobium meliloti*, showed that 17 such systems are present on its chromosome and on the two symbiotically important megaplasmids. As it was presented in Figure 7 and Table 6, the genome of *S. meliloti* encodes representatives of four of the seven TA gene families. The regulatory processes in which the members of *vapBC, higBA, parDE* and *hipBA* present in this organism are involved should be investigated. Another subject of speculation is the presence of 10 copies of complete *vapBC* TA loci encoded by the genome of *S. meliloti.*

What is the reason for encoding such a high number of *vapBC* modules and why only the chromosome encodes complete modules, whereas the plasmids encode only solitary and probably inactive toxins?

Taking into account the possible physiological role of these modules in stress management, their presence in high number in the genome of *S. meliloti* may not be surprising. Rhizobia, as symbiotic bacteria have to face various stresses: from different environmental conditions (in the soil and inside the root nodule) to different nutrition circumstances in free-living and symbiotic state. The TA systems may be involved in helping these bacteria to cope with these transitions.

Summarizing our results, in these experiments we characterized the first TA module in a symbiotic bacterium. In addition, we presented new information on the *vapBC* gene family, by determining the DNA binding ability of proteins containing an AbrB/SpoVTdomain. Data on the involvement of the *ntrPR* module in the adaptation to symbiotic state opens new perspectives on how Rhizobia manage to adjust their metabolic processes in accordance with the variable environmental requirements.

6. References

Aizenman, E., Engelberg-Kulka, H., and Glaser, G. 1996. An *Escherichia coli* chromosomal "addiction module" regulated by 3',5'-bispyrophosphate: A model for programmed bacterial cell death. Proc. Natl. Acad. Sci. USA, 93:6059-6063.

Alting-Mees, M. A., and Short, J. M. 1989. pBluescript II: gene mapping vectors. Nucleic Acids Res. 17:9494-9494.

Amitai, S., Yassin, Y., and Engelberg-Kulka, H. 2004. MazF-mediated cell death in *Escherichia coli*: a point of no return. J. Bacteriol. 186:8295-8300.

Ampe, F., Kiss, E., Sabourdy, F., and Batut, J. 2003. Transcriptome analysis of *Sinorhizobium meliloti* during symbiosis. Genome Biol. 4:R15.

Anantharaman, V., and Aravind, L. 2003. New connections in the prokaryotic toxin-antitoxin network: relationship with the eukaryotic nonsense-mediated RNA decay system. Genome Biology, 4:R81.1-81.15.

Andersson, S.G., Zomorodipour, A., Andersson, J.O., Sicheritz-Ponten, T., Alsmark, U.C., Podowski, R.M., Naslund, A.K., Eriksson, A.S., Winkler, H.H. and Kurland, C.G. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396:133-140.

Arcus, V.L., Backbro, K., Roos, A., Daniel, E.L. and Baker, E.N. 2004. Distant structural homology leads to the functional characterization o fan archeal PIN domain as an exonuclease. J. Biol. Chem. 279:16471-16478.

Becker, A., Berges, H., Krol, E., Bruand, C., Rüberg, S., Capela, D., Lauber, E., Meilhoc, E., Ampe, F., de Bruijn, F. J., Fourment, J., Francez-Charlot, A., Kahn, D., Küster, H., Liebe, C., Pühler, A., Weidner, S., and Batut, J. 2003. Global changes in gene expression in *Sinorhizobium meliloti* 1021 under microoxic and symbiotic conditions. Mol. Plant-Microbe Interact. 17:292-203.

Bernard, P., and Couturier, M. 1991. The 41 carboxy-terminal residues of the miniF plasmid CcdA protein are sufficient to antagonize the killer activity of the CcdB protein. Mol. Gen. Genet. 226:297-304.

Blattnet, F.R., et al. 1997. The complete genome sequence of Escherichia coli K-12. Science 277:1453-74.

Bloomfield, G.A., Whittle, G., McDonagh, M.B., Katz, M.E. and Cheetham, B.F., 1997: Analysis of sequences flanking the *vap* regions of *Dichelobacter nodosus*: evidence for multiple integration events, a killer system, and a new genetic element, Microbiology, 143: 553-562.

Bodogai, M., Ferenczi, Sz., Bashtovyy, D., Miclea, P., Papp, P. and Dusha, I. 2006. The *ntrPR* operon of *Sinorhizobium meliloti* is organized and functions as a toxin-antitoxin module. MPMI 19:811-822.

Bugge-Jensen, R., and K. Gerdes. 1995. Programmed cell death in bacteria: proteic plasmid stabilization systems. Mol. Microbiol. 17:205-210.

Buts, L., Lah, J., Dao-Thi, M.H., Wyns, L. and Loris, R. 2005. Toxin-antitoxin modules as bacterial metabolic stress managers. Trends Biol. Sci. 30, 12:672-769.

Calderón-Flores, A., Du Pont, G., Huerta-Saguero, A., Merchant-Larios, H., Servín-González, L., and Durán, S. 2005. The stringent response is required for amino acid and nitrate utilization, Nod factor regulation, nodulation, and nitrogen fixation in *Rhizobium etli*. J. Bacteriol. 187:5075-5083.

Cali, B.M., Kuchma, S.L., Latham, J. and Anderson, P. 1999. *smg-7* is required for mRNA surveillance in *Caenorhabditis elegans*. Genetics 151:605-616.

Camacho, A.G., Misselwitz, R., Behlke, J., Ayora, S., Welfle, K., Meinhart, A., Lara, B., Saenger, W., Welfle, H. and Alonso, J.C., 2002: *In vitro* and *in vivo* stability of the ε2ζ2 protein complex of the broad-host-range *Streptococcus pyogenes* pMS19035 addiction system, Biol. Chem., 383: 1701-1713.

Ceglowski, P., Boitsov, A., Karamyan, N., Chai, S. and Alonso, J.C. 1993. Characterization of the effectors required for stable inheritance of Streptococcus pyogenes pSM19035-derived plasmids in Bacillus subtilis. Mol. Gen. Genet. 241: 579-585.

Chatterji, D. and Kumar, O.A. 2001. Revisiting the stringent response, ppGpp and starvation signalling. Cur. Opin. Microbiol. 4:160-165.

Cherny, I. et al. 2005. The YoeB toxin is a folded protein that forms a physical complex with the unfolded protein YefM antitoxin: implications for a structural-based differential stability of toxin-antitoxin systems. J. Biol. Chem. 280:30063-30072.

Christensen, S. K., Mikkelsen, M., Pedersen, K., and Gerdes, K. 2001. RelE, a global inhibitor of translation, is activated during nutritional stress. Proc. Natl. Acad. Sci. USA, 98:14328-14333.

Christensen, S.K. and Gerdes, K. 2003a. RelE toxins from bacteria and Arcaea cleave mRNAs on translating ribosomes, which are rescued by tmRNA. Mol. Microbiol. 48: 1389-1400.

Christensen, S.K., Pedersen, K., Hansen, F.G. and Gerdes, K. 2003b. Toxin-antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA. J. Mol. Biol. 332: 809-819

Clissold, P. M., and Ponting, C. P. 2000. PIN domains in nonsense-mediated mRNA decay and RNAi. Current Biology, 10:R888-890.

Costeron, W., Veeh, R., Shirtliff, M., Pasmore, M., Post, C., Ehrlich, G. 2003. The application of biofilm science to the study and control of chronic bacterial infections. J. Clin. Investig.112: 1466-1477

Cooper, T.F. and Heinemann, J.A. 2000. Postsegregational killing does not increase plasmid stability but acts to mediate the exclusion of competing plasmids. Proc. Natl. Acad. Sci. USA. 97:12643-12648.

de la Hoz, A.B., Ayora, S., Sitkiewicz, I., Fernández, S., Pankiewicz, R., Alonso, J.C. and Ceglowski, P. 2000. Plasmid copy-number control and better-than-random segregation genes of pSM19035 share a common regulator. Proc. Natl. Acad. Sci. USA. 97: 728-733.

Dao-Thi, M.H. et al. 2000. The thermodynamic stability of the proteins of the *ccd* plasmid addiction system. J. Mol. Biol. 299:1373-1386.

Dao-Thi, M.H. 2002. Intricate interactions within the *ccd* plasmid addiction system. J. Biol. Chem. 277:3733-3742.

Davey, M. E., and de Bruijn, F. J. 2000. A homologue of the tryptophanrich sensory protein TspO and FixL regulate a novel nutrient deprivation-induced *Sinorhizobium meliloti* locus. Appl. Environ. Microbiol. 66:5353-5359.

Davis, T.L., Helsinki, D.R. and Roberts, R.C. 1992. Transcription and autoregulation of the stabilizing functions of the broad-host-range plasmid RK2 in *Escherichia coli, Agrobacterium tumefaciens* and *Pseudomonas aeruginosa*. Mol. Microbiol. 6:1981-1994.

Ditta, G., Stanfield, S., Corbin, D., and Helinski, D. R. 1980. Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA. 77:7347–7351.

Dusha, I., Bakos, Á., Kondorosi, Á., de Bruijn, F. J., and Schell, J. 1989. The *Rhizobium meliloti* early nodulation genes (*nodABC*) are nitrogen-regulated: isolation of a mutant strain with efficient nodulation capacity on alfalfa in the presence of ammonium. Mol. Gen. Genet. 219:89-96.

Eberl, L., Givskov, M., and Schwab, H. 1992. The divergent promoters mediating transcription of the par locus of plasmid RP4 are subject to autoregulation. Mol. Microbiol. 6:1969-1979.

Élő, P., Semsey, S., Kereszt, A., Nagy, T., Papp, P. and Orosz, L. 1998. Integrative promoter cloning plasmid vectors for *Rhizobium meliloti.* FEMS Microbiol. Lett. 159:7–13.

Engelberg-Kulka, H., and Glaser, G. 1999. Addiction modules and programmed cell death and antideath in bacterial cultures. Annu. Rev. Microbiol. 53:43-70.

Engelberg-Kulka, H. and Hazan, R. 2003. Cannibals defy starvation and avois sporulation. Science 301:467-468.

Engelberg-Kulka, H. et al. 2004. Bacterial programmed cell-death systems as targets for antibiotics. Trends Microbiol. 12:66-71.

Ferenczi, S., Ganyu, A., Blaha, B., Semsey, S., Nagy, T., Csiszovszki, Z., Orosz, L., Papp, P. 2004. Integrative plasmid vector for constructing single-copy reporter systems to study gene regulation in *Rhizobium meliloti* and related species. Plasmid 52:57-62.

Galibert, F. et al., 2001: The Composite Genome of the Legume Symbiont *Sinorhizobium meliloti.* Science 293:668-672.

Gazit, E. and Sauer, R.T. 1999. Stability and DNA binding of the Phd protein of the phage P1 plasmid addiction system. J. Biol. Chem. 274:2652-2657.

Gerdes, K., Poulsen, L.K., Thisted, T., Nielsen, A.K., Martinussen, J. and Andreasen, P,H. 1990. The hok killer gene family in gram-negative bacteria. New Biol. 2, 11:946-956.

Gerdes, K. 2000. Toxin-antitoxin modules may regulate synthesis of macromolecules during nutritional stress. J. Bacteriol. 182:561-572.

Gerdes, K., Christensen, S. K., and Løbner-Olesen, A. 2005. Prokaryotic toxin-antitoxin stress response loci. Nature Rev. Microbiol. 3:371-382.

Gerdes, K., Gultyaev, A. P., Franch, T. Pedersen, K. and Mikkelsen, N. D. 1997. Antisense RNA regulated programmed cell death. Annu. Rev. Genet. 31:1–31.

Gerdes, K., Rasmussen, P. B. and Molin, S. 1986. Unique type of plasmid maintenance function: Postsegregational killing of plasmid free cells. Proc. Natl. Acad. Sci. U.S.A. 83:3116-3120.

Gonzales-Pastor, J.E. et al. 2003. Cannibalism by sporulating bacteria. Science 301:510-513.

Gotfredsen, M., and Gerdes, K. 1998. The *Escherichia coli relBE* genes belong to a new toxin-antitoxin gene family. Mol. Microbiol. 29:1065-1076.

Gronlund, H. and Gerdes, K. 1999. Toxin-antitoxin systems homologous with relBE of Escherichia coli plasmid P307 are ubiquitous in prokaryotes. J. Mol. Biol. 285:1401-1415.

Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P(BAD) promoter. J. Bacteriol. 177:4121–4130.

Hayes, F. 2003. Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. Science 301:1496-1499.

Hazan, R., Sat, B., and Engelberg-Kulka, H. 2004. *Escherichia coli mazEF*-mediated cell death is triggered by various stressful conditions. J. Bacteriol. 186:3663-3669.

Hazan, R., Sat, B., Reches, M. and Engelberg-Kulka, H. 2001. Postsegregational killing mediated by the P1 phage ‘addiction module’ *phd-doc* requires the *Escherichia coli* programmed cell death system *mazEF*. J. Bacteriol. 183:2046-2050.

Howorth, S. M., and England, R. R. 1999. Accumulation of ppGpp in symbiotic and free-living nitrogen-fixing bacteria following amino acid starvation. Arch. Microbiol. 171:131-134.

Jaffe, A., Ogura, T. and Hiraga, S. 1985. Effects of the *ccd* Function of the F Plasmid on Bacterial Growth. J. Bacteriol. 163, 841-849.

Jensen, R.B., Grohmann, E., Schwab, H., Diaz-Orejas, R. and Gerdes, K. 1995. Comparison of ccd of F, parDE of RP4 and parD of R1 using a novel conditional replication control system of plasmid R1. Mol. Microbiol. 17:211-220.

Jiang, Y., Pogliano, J., Helsinki, D.R. and Konieczny, I. 2002. ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of Escherichia coli gyrase. Mol. Microbiol. 44:971-979.

Karzai, A. V., Riche, E.D. and Sauer R.T. 2000. The SsrA−SmpB system for protein tagging, directed degradation and ribosome rescue. Nat. Struct. Biol. 7:449-455.

Katz, M. E., Strugnell, R. A., and Rood, J. I. 1992. Molecular characterization of a genomic region associated with virulence in *Dichelobacter nodosus*. Infect. Immun. 60:4586-4592.

Kobayashi, I. 2001. Behaviour of restriction–modification systems as selfish mobile elements and their impact on genome evolution. Nucleic Acids Res. 29, 18:3742-3756.

Kondorosi, E., Banfalvi, Z., and Kondorosi, A. 1984. Physical and genetic analysis of a symbiotic region of *Rhizobium meliloti*: identification of nodulation genes. Mol. Gen. Genet. 193:445–452.

Kuroda, A. et al. 2001. Role of inorganic polyphosphate in promoting ribosomal protein degradation by the Lon protease in *E. coli.* Science 293:705-708.

Lehnherr, H. and Yarmolinsky, M.B. 1995. Addiction protein Phd pf plasmid prophage P1 is a substrate of the ClpXP serine protease of *Escherichia coli*. Proc. Natl. Acad. Sci. USA. 92:3274-3277.

Liolios, K., Tavernarakis, N., Hugenholtz, P. and Kyrpides, N.C. 2006 The Genomes On Line Database (GOLD) v.2: a monitor of genome projects worldwide. Nucl. Acids Res. 34 Database Issue: D332-334.

Loris, R. et al. 2003. Crystal structure of the intrinsically flexible addiction antidote MazE. J. Biol. Chem. 278:28252-28257.

Magnuson, R., and Yarmolinsky, M. B. 1998. Corepression of the P1 addiction operon by Phd and Doc. J. Bacteriol. 180:6342-6351.

Makarova, K. S., Aravind, L., Galperin, M. Y., Grishin, N. V., Tatusov, R. L., Wolf, Y. I., and Koonin E. V. 1999. Comparative genomics of the Archaea (Euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. Genome Res. 9:608-628.

Maniatis, T., Fritsch, E. F., and Sambrook, J. E. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Marianovsky, I., E. Aizenman, H., Engelberg-Kulka, and Glaser, G. 2001. The regulation of the *Escherichia coli mazEF* promoter involves an unusual alternating palindrome. J. Biol. Chem. 276:5975-5984.

Masuda, Y., Miyakawa, K., Nishimura, Y., and Ohtsubo, E. 1993. *chpA* and *chpB*, *Escherichia coli* chromosomal homologs of the *pem* locus responsible for stable maintenance of plasmid R100. J. Bacteriol. 175:6850-6856.

Maderazo, A.B., He, F., Mangus, D.A. and Jacobson, A. 2000. Upf1p control of nonsense mRNA translation is regulated by Nmd2p. Mol. Cell Biol. 20:4591-4603.

Meinhart, A., Alonso, J.C., Strater, N. and Saenger, W. 2003. Crystal structure of the plasmid maintenance system ε/ζ: functional mechanism of toxin ζ and inactivation by ε2ζ2 complex formation. Proc. Natl. Acad. Sci. USA. 100:1661-1666.

Mikkelsen, ND, and Gerdes, K. 1997. Sok antisense RNA from plasmid R1 is functionally inactivated by RNase E cleavage and polyadenylated by poly(A) polymerase I. Mol. Microbiol. 26:311-320.

Miller, J.H. 1972. *Experiments in Molecular Genetics.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory

Moris, M., Braeken, K., Schoeters, E., Verreth, C., Beullens, S., Vanderleyden, J., and Michiels, J. 2005. Effective symbiosis between *Rhizobium etli* and *Phaseolus vulgaris* requires the alarmone ppGpp. J. Bacteriol. 187:5460-5469.

Muñoz-Gómez, A. J., Santos-Sierra, S., Berzal-Herranz, A., Lemonnier, M., and Díaz-Orejas, R. 2004. Insights into the specificity of RNA cleavage by the *Escherichia coli* MazF toxin. FEBS Letters, 567:316-320.

Noguchi, E., Hayashi, N., Azuma, Y., Seki, T., Nakamura, M., Nakashima, N., Yanagida, M. et al. 1996. Dis3, implicated in mitotic control, binds directly to Ran and enhances the GEF activity of RCC1. EMBO J. 15:5595-5605.

Nordstrom, K., and S. J. Austin. 1989. Mechanisms that contribute to the stable segregation of plasmids. Annu. Rev. Genet. 23:37-69.

Oláh, B., Kiss, E., Györgypál, Z., Borzi, J., Cinege, G., Csanádi, G., Batut, J., Kondorosi, A., Dusha, I. 2001. Mutation in the *ntrR* gene, a member of the *vap* gene family, increases the symbiotic efficiency of *Sinorhizobium meliloti*. Mol. Plant-Microbe Interact. 14:887-894.

Oberer, M., Zangger, K., Prytulla, S. and Keller, W. 2002. The antitoxin ParD of plasmid RK2 consists of two structurally distinct moieties and belongs to the ribbon-helix-helix family of DNA-binding proteins. Biochem. J. 361:41-47.

Pandey, D.P. and Gerdes, K. 2005. Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. Nucl. Acid Res. 33, 3:966-976.

Pecota, D. C., Osapay, G., Selsted, M. E. and Wood, T. K. 2003. Antimicrobial properties of the *Escherichia coli* R1 plasmid host killing peptide. J. Biotechnol. 100:1–12.

Pedersen, K., Christensen, S. K., and Gerdes, K. 2002. Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. Mol. Microbiol. 45:501-510.

Pedersen, K., Zavialov, A. V., Pavlov, N. Y., Elf, J., Gerdes, K., and Ehrenberg, M. 2003. The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. Cell, 112:131-140.

Pogliano, J. 2002. Dynamic cellular location of bacterial plasmids. Curr. Opin. Microbiol. 5:586-590.

Pullinger, G.D. and Lax, A.J. 1992. A *Salmonella dublin* virulence locus that affects bacterial growth under nutrient-limited conditions. Mol. Microbiol. 6:1631-1643.

Puskás, L. G., Nagy, Z. B., Kelemen, J. Z., Rüberg, S., Bodogai, M., Becker, A., and Dusha, I. 2004. Wide-range transcriptional modulating effect of *ntrR* under microaerobiosis in *Sinorhizobium meliloti*. Mol. Gen. Genomics, 272:275-289.

Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T., Nelson K.E., et al. 2003. The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. Nature 423:81-86.

Roberts, R.C., Spangler, C., and Helinski, D.R. 1993. Characteristics and significance of DNA binding activityo of plasmid stabilization protein ParD from the broad host-range plasmid RK2. J. Biol. Chem. 268:27109-27117.

Rowe-Magnus, D.A., Guerout, A.M., Biskri, L., Bouige, P. and Mazel, D. 2003. Comparative analysis of superintegrons: engineering extensive genetic diversity int he Vibrionaceae. Genome Res. 13:428-442.

Roszak, D. B., and Colwell, R. R. 1987. Survival strategies of bacteria in the natural environment. Microbiol. Rev. 51:365-379.

Ruiz-Echevarria, M. J., Giménez-Gallego, G., Sabariegos-Jareño, R., and Diaz-Orejaz, R. 1995. Kid, a small protein of the parD stability system of plasmid R1, is an inhibitor of DNA replication acting at the initiation of DNA synthesis. J. Mol. Biol. 247:568-577.

Sat, B., Hazan, R., Fisher, T., Khaner, H., Glaser, G., and Engelberg-Kulka, H. 2001. Programmed cell death in *Escherichia coli*: some antibiotics can trigger *mazEF* lethality. J. Bacteriol. 183:2041-2045.

Sat, B., Reches, M., and Engelberg-Kulka, H. 2003. The *Escherichia coli mazEF* suicide module mediates thymineless death. J. Bacteriol. 185:1803-1807.

Sayeed, S., Reaves, L., Radnedge, L. and Austin, S. 2000. The stability region of the large virulence plasmid of *Shigella flexneri* encodes an efficient postsegregational killing system. J. Bacteriol. 182: 2416-2421.

Shpanchenko, O.V., Zvereva, M.I., Ivanov, P.V., Bugaeva, E.Y. Rozov, A.S., Bogdanov, A.A., Kalkum, M., Isaksson, L.A., Nierhaus,K.H., and Dontsova, O.A. 2005. Stepping Transfer Messenger RNA through the Ribosome. J. Biol. Chem. 280, 18:18368-18374.

Sia, E.A., Roberts, R.C., Easter, C., Helsinki, D.R. and Figurski, D.H. 1995. Different relative importances of the par operons and the effect on conjugal transfer on the maintenance of intact promiscuous plasmid RK2. J. Bacteriol. 177:2789-2797.

Smith, A.S. and Rawlings, D.E. 1997. The poison-antidote stability system of the broad-host range *Thiobacillus ferrooxidans* plasmid pTF-FC2. Mol. Microbiol. 26:961-970.

Strauch, M. A., Spiegelman, B. G., Perego, M., Johnson, W. C., Burbulys, D., and Hoch, J. A. 1989. The transition state transcription regulator *abrB* of *Bacillus subtilis* is a DNA binding protein. EMBO J. 8:1615-1621.

Tachdjian, S. and Kelly, R.M. 2006. Dynamic metabolic adjustments and genome plasticity are implicated in the heat shock response of the extremely thermoacidophilic archaeon *Sulfolobus solfataricus.* J. Bact. 188:4553-4559.

Tamas, I., Klasson, L., Canback, B., Naslund, A.K., Eriksson, A.S., Wernegreen, J.J., Sandstrom, J.P., Moran, N.A. and Andersson, S.G. 2002. 50 million years of genomic stasis in endosymbiotic bacteria. Science 296:2376-2379.

Tsuchimoto, S., and Ohtsubo, E. 1993. Autoregulation by cooperative binding of the *pemI* and *pemK* proteins to the promoter region of the *pem* operon. Mol. Gen. Genet. 237:81-88.

van Melderen, L. et al. 1996. ATP-dependent degradation of CcdA by Lon protease. Effects of secondary structure and heterologous subunit interactions. J. Biol. Chem. 271:27730-27738.

van Melderen, L. 2002. Molecular interactions of the CcdB poison with its bacterial target, the DNA gyrase. Int. J. Med. Microbiol. 291:537-544.

Vaughn, J. L., Feher, V., Naylor, S., Strauch, M. A., and Cavanagh, J. 2002. Novel DNA binding domain and genetic regulation model of *Bacillus subtilis* transition state regulator *abrB*. Nat. Struct. Biol. 12:98-106.

Viprey, V., Perret, X. and Broughton, W.J. 2000. Host-plant invasion by rhizobia. Subcell. Biochem. 33:437-456.

Wall, D., and Kaiser, D. 1999. Type IV pili and cell motility. Mol. Microbiol 32:1-10.

Wells, D. H., and Long, S. R. 2002. The *Sinorhizobium meliloti* stringent response affects multiple aspects of symbiosis. Mol. Microbiol. 43:1115-1127.

Wolfgang, M., van Putten J. P. M., Hayes, S. F., Dorward, D., and Koomey, M. 2000. Components and dynamics of fiber formation define a ubiquitous biogenesis pathway for bacterial pili. EMBO J. 19:6408-6418.

Yarmolinsky, M. B. 1995. Programmed cell death in bacterial populations. Science 267:836-837.

Zavialov, A.V., Mora, L., Buckingham, R.H. and Ehrenberg, M. 2002. Release of peptide promoted by the GGQ motif of class 1 release factors regulates the GTPase activity of RF3. Mol. Cell. 10:789-798.

Zhang, J., Zhang, Y., and Inouye, M. 2003b. Characterization of the interactions within the *mazEF* addiction module of *Escherichia coli*. J. Biol. Chem. 278:32300-32306.

Zhang, J., Zhang, Y., Zhu, L., Suzuki, M and Inouye, M. 2004. Interference of mRNA function by sequence-specific endoribonuclease PemK. J. Biol. Chem. 279:20678-20684.

Zhang, Y., Zhang, J., Hara, H., Kato, I. and Inouye, M., 2005: Insights into the mRNA cleavage mechanism by MazF, an mRNA interferase, J. Biol. Chem., 280:3143-3150

Zhang, Y., Zhang, J., Hoeflich, K. P., Ikura, M., Qing, G., and Inouye, M. 2003a. MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. Mol. Cell, 12:913-923.

Zhang, Y.X., Guo, X.K., Wu, C., Bi, B., Ren, S.X., Wu, C.F. and Zhao, G.P. 2004. Characterization of a novel toxin-antitoxin module, VapBC, encoded by the *Leptospira interrogans* chromosome. Cell Res. 14:208-216.

List of publications

Bodogai M, Ferenczi S, Bashtovyy D, Miclea P, Papp P, Dusha I., The ntrPR operon of Sinorhizobium meliloti is organized and functions as a toxin-antitoxin module. Mol Plant Microbe Interact. 2006 Jul;19(7):811-22.

Puskas LG, Nagy ZB, Kelemen JZ, Ruberg S, **Bodogai M**, Becker A, Dusha I. Wide-range transcriptional modulating effect of ntrR under microaerobiosis in Sinorhizobium meliloti. Mol Genet Genomics. 2004 Oct;272(3):275-89

Abstract

The recent expansion of the microbial DNA and protein databases that followed the sequencing of a high number of prokaryotic genomes, promoted the identification of a large number of toxin-antitoxin (TA) modules present in bacterial plasmids and chromosomes.

The first TA modules were identified on plasmids acting as post-segregational killing systems. Their function was to prevent the proliferation of plasmid-free progeny. Subsequently, TA loci were also found on chromosomes and were considered to be associated with the modulation of the global level of translation under various stress conditions. A typical TA module consists of two small genes that form an operon, in which the first gene determines an unstable antitoxin and the second gene, a stable toxin protein. The two proteins form a complex, thus the antitoxin prevents the lethal effect of the toxin. Under stress conditions, the antitoxin is degraded by proteases, and the activity of the free toxin results in the inhibition of translation.

Seven TA gene families were described; one of them is the most abundant *vapBC* family, present in Gram-positive and Gram-negative bacteria as well as in Archaea. The antitoxin protein of this family is an AbrB/MazE homolog and the toxin partner belongs to the PIN domain family.

Previously we have identified the *ntrPR* operon in *Sinorhizobium meliloti* and showed that the organization and domain structure of this operon resembled those of the TA modules belonging to the *vapBC* family. As was demonstrated under symbiotic conditions, a Tn5 insertion in the *ntrR* gene resulted in increased transcription of nodulation and nitrogen fixation genes as compared to that of the wild type strain, and this effect was more pronounced in the presence of an external ammonium source. We supposed that the NtrR protein has a role in the nitrogen regulation of the *nod* and *nif* genes. However, when the gene expression pattern of the entire genomes of the wild type and *ntrR* mutant strains were compared, an unexpectedly large number of genes exhibited altered expression in the mutant strain, suggesting a more general function for NtrR.

Objectives

Our aim was to examine whether or not, the NtrP and NtrR proteins represent an active TA system, based on their autoregulatory properties, complex formation and function; and to determine the possible physiological role of this operon in *Sinorhizobium meliloti*.

Results

The autoregulatory functions of NtrPR have been investigated by using different methods: measurements of the *ntrPR* promoter activity, electrophoretic mobility shift assays and DNase I footprinting.

We have shown that the the *ntrPR* operon exhibits the characteristic regulatory circuit of bacterial toxin-antitoxin modules: the antitoxin NtrP is able to recognize a DNA segment in the promoter region of the *ntrPR* operon, but its binding is weak. The toxin component alone is not able to bind to the same DNA region, but the complex of NtrP and NtrR strongly binds to the promoter region resulting in the negative autoregulation. The N-terminal part of NtrP is responsible for the interaction with the promoter DNA, whereas the C-terminal part is required for protein-protein interactions.

NtrR toxicity assays revealed that the expression of this protein results in the inhibition of cell growth and colony formation.

Experiments were performed to determine the induction of the *ntrPR* module in *Sinorhizobium meliloti*. ppGpp, an inducer of the *mazEF* TA family had no effect on the *ntrPR* operon, whereas antibiotics that inhibit transcription or translation were shown to have a weak effect on this module.

If the toxin-antitoxin modules are considered to function as metabolic stress managers, *Sinorhizobium meliloti* can be a valuable test organism for studying such systems. The metabolism of this symbiotic nitrogen fixing bacterium is influenced by a wide-range of conditions under which it is capable to survive: it can be found in the soil as a free-living bacterium, or it occupies the nodules developed on the roots of leguminous plant, it is able to live under aerobic conditions or function under microoxic conditions as a bacteroid, the differentiated form of the bacterium. Another distinct feature of *Sinorhizobium meliloti* is the special metabolism in the symbiotic state: the carbon source utilized by the bacteroids in the form of dicarboxylic acids is supplied by the plant metabolism to fuel nitrogen fixation, while in exchange, fixed nitrogen is transferred from the bacteroid to the plant cell and the ammonium assimilation is inactive in the bacteroids.

Toxin-antitoxin systems present on the chromosome of the bacterium may have an important role during the transition from one way of life to the other, and may determine the adaptation to the varying environmental conditions.

Previous experiments that demonstrated the higher expression of *nod, nif* and *fix* genes in the *ntrPR* mutant compared with the *Sinorhizobium meliloti* 1021 wild type under symbiotic conditions and the wide range of transcriptional changes in the genome of this bacterium, suggest that the *ntrPR* operon may help this organism to cope with the metabolic stress induced by the transition from free-living to symbiotic state, since this transition involves drastic metabolic changes.

Based on protein homologies, domain architectures and gene neighborhood analysis we could identify 17 TA modules belonging to different gene families in the genome of *Sinorhizobium meliloti* strain 1021.

Ten of the 17 modules are members of the *vabBC* gene family and all of them are located on the bacterial chromosome. What is the reason for encoding such a high number of *vapBC* modules and why only the chromosome carries complete modules, whereas the plasmids encode only solitary and probably inactive toxins?

Taking into account the possible physiological role of these modules in stress management, their presence in high number in the genome of *Sinorhizobium meliloti* may not be surprising. The TA systems may be involved in helping these bacteria to cope with metabolic transitions.

Summarizing our results, in these experiments we characterized the first TA module in a symbiotic bacterium. In addition, we presented new information on the *vapBC* gene family, by determining the DNA binding ability of proteins containing an AbrB/SpoVTdomain. Data on the involvement of the *ntrPR* module in the adaptation to symbiotic state opens new perspectives on how Rhizobia manage to adjust their metabolic processes in accordance with the variable environmental requirements.

Összefoglalás

A közelmúltban nagyszámú prokarióta genom teljes DNS szekvenciája vált ismertté, és az eredmények feldolgozása a mikrobiális DNS és fehérje adatbázisok gyors növekedését eredményezte. Az így felhalmozódott adattömeg elősegítette a baktériumok kromoszómáján és plazmidjain jelenlévő nagyszámú toxin-antitoxin (TA) modul azonosítását is.

Az első TA modulokat, mint plazmidokon működő poszt-szegregációs elimináló rendszereket írták le. Ezek funkciója a plazmid nélküli utódsejtek elszaporodásának megakadályozása. Ezt követően kromoszómális TA lókuszokat is azonosítottak, melyek feladata különböző stresszhatásokat követően a transzláció általános szintjének modulálása. A tipikus TA modul két kisméretű génből álló operon, melyben az első gén egy instabil antitoxint, a második gén pedig egy stabil toxin fehérjét határoz meg. A két fehérje komplexet képez, így akadályozza meg az antitoxin a toxin letális hatását. Stressz körülmények között az antitoxint különböző proteázok elbontják, ekképpen a szabad toxin aktivitásának hatására a transzláció gátlása következik be.

Hét TA géncsaládot írtak le, melyek közül a legnépesebb a Gram-pozitív és Gram-negatív baktériumokban, valamint az Archaeakban is jelenlévő *vapBC* család. Ebben a családban az antitoxin fehérje az AbrB/MazE fehérjékkel, a toxin partner pedig a PIN doménnel mutat homológiát.

A *Sinorhizobium melilotiban* korábban azonosított *ntrPR* operonról kimutattuk, hogy szerveződése és az általa kódolt fehérjék domén struktúrája a *vapBC* családhoz tartozó TA modulokéval megegyezik. Az *ntrR* génben lévő Tn5 inszerció a gümőképzési és nitrogénkötési gének megnövekedett transzkripcióját okozta a szimbiózis körülményei között, és ez a hatás külső ammóniumforrás jelenlétében még fokozottabb volt. Ezért azt feltételeztük, hogy az NtrR fehérjének a *nod* és *nif* gének nitrogénszabályozásában van szerepe. Azonban a vad típusú és az *ntrR* mutáns törzs teljes genomjának génexpressziós mintázatát összehasonlítva azt az eredményt kaptuk, hogy a mutáns törzsben meglepően nagy számú gén mutatott megváltozott expressziós szintet, mely arra utalt, hogy az NtrR fehérjének általánosabb funkciója lehet.

Célkitűzés

Célunk az volt, hogy megvizsgáljuk, vajon az NtrP és NtrR fehérje autoregulációs tulajdonsága, komplex képzése és funkciója alapján aktív TA rendszert képvisel-e, továbbá, hogy meghatározzuk ennek az operonnak a lehetséges fiziológiai szerepét *Sinorhizobium melilotiban.*

Eredmények

Az NtrPR fehérjék autoregulációban játszott szerepét különböző módszerekkel vizsgáltuk: mértük az *ntrPR* promóter aktivitását, valamint a fehérje-DNS kölcsönhatás meghatározására elektroforetikus mobilitás változást detektáló és DNázI „footprint" méréseket végeztünk.

Kimutattuk, hogy az *ntrPR* operon a bakteriális toxin-antitoxin modulok jellegzetes regulációs tulajdonságaival rendelkezik, az NtrP antitoxin felismer egy DNS szakaszt az *ntrPR* operon promóter régiójában, de a kötődése gyenge. A toxin komponens önmagában nem képes kötődni ehhez a DNS régióhoz, de az NtrP és NtrR komplexe erősen kötődik a promóterhez, így a komplex negatív autoregulációt idéz elő. Az NtrP N-terminális fele felelős a promóter DNS-sel való kölcsönhatásért, míg a C-terminális rész a fehérje-fehérje kölcsönhatáshoz szükséges.

Az NtrR toxicitásának vizsgálata bizonyította, hogy ennek a fehérjének az expressziója a sejtnövekedés és a telepképzés gátlását okozza.

Megvizsgáltuk az *ntrPR* modul indukcióját *Sinorhizobium melilotiban*. A *mazEF* TA család inducerének, a ppGpp-nek nem volt hatása az *ntrPR* operonra, míg a transzkripciót vagy a transzlációt gátló antibiotikumok gyengén befolyásolták a modul expresszióját.

Ha a toxin-antitoxin modulok stressz hatására működésbe lépő modulátorként játszanak szerepet, akkor a *Sinorhizobium meliloti* igen hasznos teszt organizmus lehet a hasonló rendszerek megismerésében. Ennek a szimbiótikus nitrogénkötő baktériumnak a metabolizmusát ugyanis igen sokféle külső tényező befolyásolhatja, melyeket képes tolerálni és túlélni: megtalálható szabadon élő állapotban a talajban, vagy a pillangós növények gyökerén fejlődő gümőkben, életképes aerob körülmények között, de a baktérium differenciálódott formájában bakteroidként funkcióképes alacsony oxigén koncentrációk mellett is. Egy másik jellegzetes tulajdonsága a *Sinorhizobium melilotinak*, hogy szimbiótikus állapotban speciális metabolizmussal rendelkezik: a nitrogénkötéshez energiát biztosító dikarbonsavakat, melyek a bakteroidok által felhasznált szénforrások, a növény

metabolizmusa biztosítja, míg cserébe, a kötött nitrogén a bakteroidokból a növényi sejtekbe szállítódik, és a bakteroidok ammónium asszimilációs rendszere inaktív.

A baktérium kromoszómális toxin-antitoxin rendszereinek fontos szerepe lehet az egyik életformából a másikba való átmenet során, és meghatározhatják a változó környezeti feltételekhez való alkalmazkodást.

A korábbi kísérletek, melyekben kimutattuk a *nod*, *nif* és *fix* gének megnövekedett expresszióját az *ntrR* mutánsban szimbiótikus körülmények között, valamint a nagyarányú transzkripciós változások a baktérium genomjában mind arra utalnak, hogy az *ntrPR* operon segíthet a baktérium sejteknek azokban a metabolizmust érintő változásokban, melyeket a szabadon élő állapotból a szimbiótikus állapotba való átmenet indukál, hiszen ez az átmenet jelentős metabolikus átalakulásokkal jár.

Fehérje homológia, domén felépítés és génszomszédság analízis alapján a *Sinorhizobium meliloti* 1021 törzs genomjában 17 TA modult azonosítottunk, melyek különböző géncsaládokhoz tartoznak. A 17 modulból 10 a *vapBC* géncsalád tagja, és ezek mindegyike a baktérium kromoszómáján helyezkedik el. Vajon mi lehet az oka, hogy ilyen nagy számú *vapBC* modul van a genomban, és miért csak a kromoszóma hordoz teljes modulokat, míg a plazmidokon csak magányos és valószínűleg inaktív toxin gének találhatók? Figyelembe véve ezeknek a moduloknak a stresszkezelésben betöltött lehetséges fiziológiai szerepét, nagy számban való előfordulásuk a *Sinorhizobium meliloti* genomjában talán nem meglepő. A TA rendszerek segíthetnek a baktériumnak a különböző metabolikus átmenetek kezelésében.

Eredményeinket összefoglalva, kísérleteinkben jellemeztük az első, szimbiótikus baktériumban jelenlévő TA modult. Továbbá, az AbrB/SpoVT domént tartalmazó fehérjék DNS kötő képességét meghatározva új információt nyertünk a *vapBC* géncsaládról. Adataink, melyek az *ntrPR* modul szerepét mutatják a szimbiótikus állapothoz való alkalmazkodásban, új perspektívát tárnak fel arról, hogy a *Rhizobiumok* miként alkalmazkodnak metabolikus folyamataikkal a változó környezeti követelményekhez.