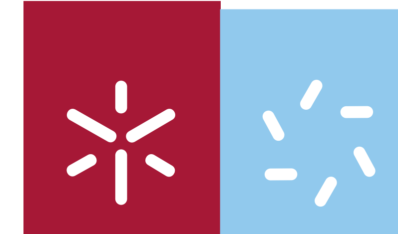
Ana Raquel Maceiras de Oliveira **On the role of a 5'-leader region in controlling the levels of the aromatic-responsive transcriptional activator DmpR**

UMinho | 2011

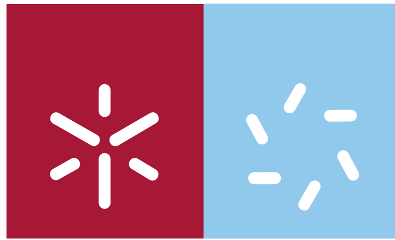


Universidade do Minho
Escola de Ciências

Ana Raquel Maceiras de Oliveira

On the role of a 5'-leader region in controlling the levels of the aromatic-responsive transcriptional activator DmpR

Outubro de 2011



Universidade do Minho
Escola de Ciências

Ana Raquel Maceiras de Oliveira

On the role of a 5'-leader region in controlling the levels of the aromatic-responsive transcriptional activator DmpR

Dissertação de Mestrado
Mestrado em Genética Molecular

Trabalho realizado sob a orientação da
Professora Doutora Victoria Shingler
Professor at the Department of Molecular Biology
Umeå University, Umeå, Sweden

Outubro de 2011

Declaração

NOME: Ana Raquel Maceiras de Oliveira

ENDEREÇO ELECTRÓNICO: raquel.maceiras@gmail.com

BILHETE DE IDENTIDADE: 13364306

TÍTULO DA TESE DE MESTRADO:
On the role of a 5'-leader region in controlling the levels of the aromatic-responsive transcriptional activator DmpR
O papel de uma região 5' líder no controlo dos níveis do activador transcripcional DmpR

ORIENTADOR:
Professora Doutora Victoria Shingler

ANO DE CONCLUSAO: 2011

MESTRADO EM:
Genética Molecular

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA TESE APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE;

Universidade do Minho, ____/____/_______

Assinatura:

____

Acknowledgements

First of all I want to thank my supervisor Victoria Shingler for accepting me in her group and to give me the opportunity to work and learn so much. I'm also grateful for helping me with the stay arrangements and all the paper work. For everything, thank you so much.

A special thanks to all the group members: to Eleonore for teaching me how to work in the lab and for the patience with all my questions; to Sofia for being available to help me and for always asking about my work on your visits; to Anjana for always smiling and for keeping me company during lunch time; last but not least, to Teresa for teaching me, helping me and for being a friend.

I also want to thank my family and friends for the support during this year even 4000 kilometres apart. A special thank to Lídia and Cristina for having the courage to face a 24h trip to visit me.

Carlos, thank you for being there for me and for being who you are.

To my mother, I want to thank for made this possible and for giving me all the support and courage I needed.

On the role of a 5'-leader region in controlling the levels of the aromatic-responsive transcriptional activator DmpR

Metabolically versatile bacteria play an important role in recycling carbon in the environment. For certain bacteria this metabolic versatility extends to seeming obnoxious toxic carbon sources such as aromatic compounds that can cause environmental pollution. One such example is *Pseudomonas putida* CF600 that carries the *dmp*-system for catabolism of dimethylphenols, mono-methylated phenols, and phenol on a catabolic plasmid. The *dmp*-system consists of the *dmp*-operon encoding the specialized catabolic enzymes divergently transcribed from the *dmpR* gene. The *dmpR* gene encodes the aromatic-responsive transcriptional activator DmpR whose activity is strictly required for transcription of the *dmp*-operon. Because DmpR is a sensor-regulator that is activated upon binding substrates of the *dmp*-pathway enzymes, the *dmp*-system is always silent unless substrates are available. However, like other auxiliary catabolic pathways, regulation of expression of the Dmp-enzymes is also highly integrated within the host global regulatory network such that the system is also silent if more energetically favourable carbon sources are present. Failure to engineer such integration within host physiology has lead to unpredictable performance of artificial constructed catabolic pathways under field conditions. This provides a practical impetus to gain a greater understanding of the mechanisms involved. Much previous work had focused on the multiple roles of a bacterial alarmone that converge to stimulate activity of the promoter that drives transcription of *dmpR* to maximize performance of the *dmp*-system under low-energy / stress conditions. However, the 5'-leader region of the *dmpR* mRNA has also been implicated in playing a regulatory role. In this work, it is presented evidence, from *in vivo* and *in vitro* assays, that the DNA encoding the 5'-leader region and the cognate region of the resulting mRNA exert control of the levels of DmpR by at least three different mechanisms: I) at the level of transcription through a ATAAATA motif within the 5'-leader region DNA, II) at the level of translation by binding of Crc to the 5'-leader region RNA, and III) by a less well defined, Crc-independent mechanism, that likely involved coupling of translation between a small open-reading frame with the 5'-leader region and that of the downstream *dmpR* gene. The results of these analysis and their physiological and mechanistic implications are discussed.

O papel de uma região 5' líder no controlo dos níveis do activador transcripcional DmpR

Bactérias metabolicamente versáteis são importantes na reciclagem de carbono no ambiente. Em algumas delas, a sua versatilidade abrange fontes de carbono aparentemente tóxicas, como compostos aromáticos, e causadoras de poluição ambiental. Um exemplo é a espécie *Pseudomonas putida* CF600 que possui o sistema *dmp* que permite o catabolismo de dimetil-fenois, metil-fenois e fenol. O sistema *dmp* consiste no operão *dmp*, que codifica enzimas catabolicas especializadas, e o gene *dmpR* divergentemente transcrito. Este último codifica o ativador transcriptional DmpR cuja atividade é estritamente necessária para ocorrer transcrição do operão. Sendo o DmpR um sensor / regulador apenas ativo após a ligação a substractos da via metabólica, o sistema *dmp* encontra-se sempre silenciado, exceto, quando substratos estão presentes. No entanto, como qualquer outra via metabólica auxiliar, a regulação da expressão das enzimas Dmp está também integrada nas vias regulatórias globais da célula; desta forma, o sistema é silenciado quando fontes de carbono mais favoráveis estão presentes. A falha em construir esta integração com a fisiologia do hospedeiro tem levado a resultados imprevistos por parte de vias catabólicas artificialmente construídas quando submetidas a condições de campo. Este facto impulsiona a obtenção de um melhor entendimento dos mecanismos envolvidos. Uma grande parte do trabalho previamente efetuado focou-se nos múltiplos papéis de uma alarmona bacterial, os quais convergem para estimular a atividade do promotor do gene dmpR, de modo a maximizar a performance do sistema em condições de baixa energia / stress. No entanto, a região 5' líder do mRNA do gene *dmpR* parece também estar implicada na regulação dos níveis da proteína. Neste trabalho, são apresentadas evidências, de ensaios realizados *in vivo* e *in vitro*, em como o DNA codificante desta região e a correspondente região do mRNA controlam os níveis de DmpR através de pelo menos 3 mecanismos: ao nível da transcrição através do motivo ATAAATA presente no DNA; ao nível da tradução através da ligação da proteína Crc ao mRNA e através de um mecanismo pouco definido mas que parece envolver a tradução acoplada entre uma pequena ORF (dentro da região 5' líder) e o gene dmpR. A discussão dos resultados desta análise, as implicações fisiológicas e os mecanismos associados são apresentados.

Contents

Acknowledgements .... **III**

Summary .... **V**

Resumo .... **VII**

Contents .... **IX**

Abbreviations .... **XI**

1. Introduction **15**

1.1. Regulation of gene expression 18

- Sigma factors 19
- (p)ppGpp and DskA 21
- 5' -Leading Regions (5'-LRs) 23
- Small RNAs (sRNAs) 23

1.2. *Pseudomonas putida* 24

1.3. Carbon Catabolite Repression 25

- The Pseudomonas Crc – Catabolite repression control protein 25

1.4. Introduction to the *dmp*-experimental system 26

1.5 Negative regulation by the extensive and conserved 5'-LR 28

2. Aims **31**

3. Methods **35**

3.1. Bacterial strains and culture conditions 37

3.2. Plasmid constructions 37

3.3. Crc protein for IVTT reaction 40

3.4. Electro-mobility shift assays (EMSA) 40

3.5. *In vitro* transcription-translation (IVTT) assays 40

3.6. *In vivo* Luciferase assays 41

3.7. Western analysis 42

4. Results 43

4.1. DNA encoding the 5’-LR mediates inhibition at the level of transcription 45

4.2. The 5’-LR controls expression of DmpR at the level of translation 46

4.3. Crc controls translation of DmpR through binding to the 5’-LR 47

4.4. A highly conserved promoter proximal region of 5’-LR is also involved in controlling DmpR levels 50

4.5. The 5’-LR encompasses an upstream ORF (uORF) and a potential target site for a small regulatory RNA. 54

5. Discussion 57

How does the initially transcribed (5’-LR) DNA control output from the Pr promoter? 60

How does Crc-binding to the mRNA result in repression of translation of DmpR? 61

How might the secondary structure of the 5’-LR potentially influence binding of Crc? 62

How might translational coupling through an upstream open-reading frame (uORF) enhance translation of DmpR? 62

6. Conclusions 65

References 69

Abbreviations

bEBPs	bacterial enhancer binding proteins
bp	Base pairs
Cb	Carbenicillin
CCR	Carbon Catabolite Repression
CRC	Catabolite repression control
DNA	Deoxyribonucleic acid
Gm	Gentamycin
h	Hours
IHF	Integration host factor
IVTT	*in vitro* transcription and translation
Km	Kanamycin
LB	Luria-Bertani/Lennox broth
LR	Leading region
mins	Minutes
nt	Nucleotides
ORF	Open-reading frame
ppGpp	Guanosine tetraphosphate
pppGpp	Guanosine pentaphosphate
(p)ppGpp	ppGpp and pppGpp
RBS	Ribosome binding site
RNA	Ribonucleic acid
sRNA	Small RNA
Tp	Trimethoprim
UAS	Upstream activating sites

"The important thing is not to stop questioning.

Curiosity has its own reason for existing."

Albert Einstein

1. Introduction

Bacteria are able to occupy a wide range of niches, including hostile environments. Physico-chemical parameters such as temperature, pH, osmotic pressure, and nutrient (e.g. carbon and nitrogen) sources are some of the major factors that influence bacterial survival and growth. To accommodate fluctuating and/or unfavourable conditions, bacteria have evolved a broad metabolic capacity and highly sophisticated regulatory networks that combine plasticity with tight regulation. The genome size and content (and consequently the metabolic profile) of a particular bacterial species mainly depends on the selective pressure exerted by the environment and lifestyle of the organism (1-3). Therefore, bacteria such as *Pseudomonas putida,* which naturally inhabit continually changing and highly competitive soil and aquatic environments, have larger genomes and more regulators per gene than intracellular pathogens and endosymbionts that inhabit the comparably stable environments of host cells (4,5).

Continual competition for limiting carbon sources has selected diverse biochemical pathways for the conversion of a wide range of organic compounds to intermediates of central metabolism. Due to promiscuity of some of the enzymes and regulatory circuits involved, several of these pathways can also degrade synthetic (man-made) compounds, albeit often inefficiently. These catabolic pathways serve as an enormous "library" of enzymes and regulatory components for biotechnological applications and for bacteria to evolve new pathways as novel chemicals become available as nutrient sources (6,7).

The ability of certain bacteria to degrade different aromatic compounds that are considered pollutants (e.g. benzene, toluene, xylene [BTX], phenols, naphthalenes, atrazine, nitroaromatics, biphenyls, polychlorinated biphenyls [PCBs] and chlorobenzoates) has been exploited for bioremediation – the process of degradation or bioconversion of hazardous components of wastes or *in situ* pollutants in the environment using microorganisms. The specificity of pathway enzymes as well as specific and global regulatory circuits that control their expression can limit both the efficiency and range of compounds that can be degraded by microorganisms (6). The prospective of exploiting bacteria capable of catabolism of aromatic compounds in bioremediation protocols has led to an extensive study of their ecology, biochemistry, gene regulation and physiological adaptation processes in order to optimise their potential usefulness (8). This thesis is centred on the regulatory circuit that controls degradation of aromatic methylphenols pollutants by a soil/water microorganism –

Pseudomonas putida CF600. Therefore, in the following sections I briefly overview regulation of gene expression in bacteria, focussing in particular on factors that are important for controlling methylphenol catabolism by *P. putida* CF600.

1.1. Regulation of gene expression

To be able to accommodate environmental changes, bacteria have to rapidly respond by regulating their gene expression profile. In bacteria, this regulation can occur at the transcriptional and translational levels (Figure 1) and/or the post-translational level. Since 1957, when a regulatory model for the *lac* operon was first described, analysis of regulation of

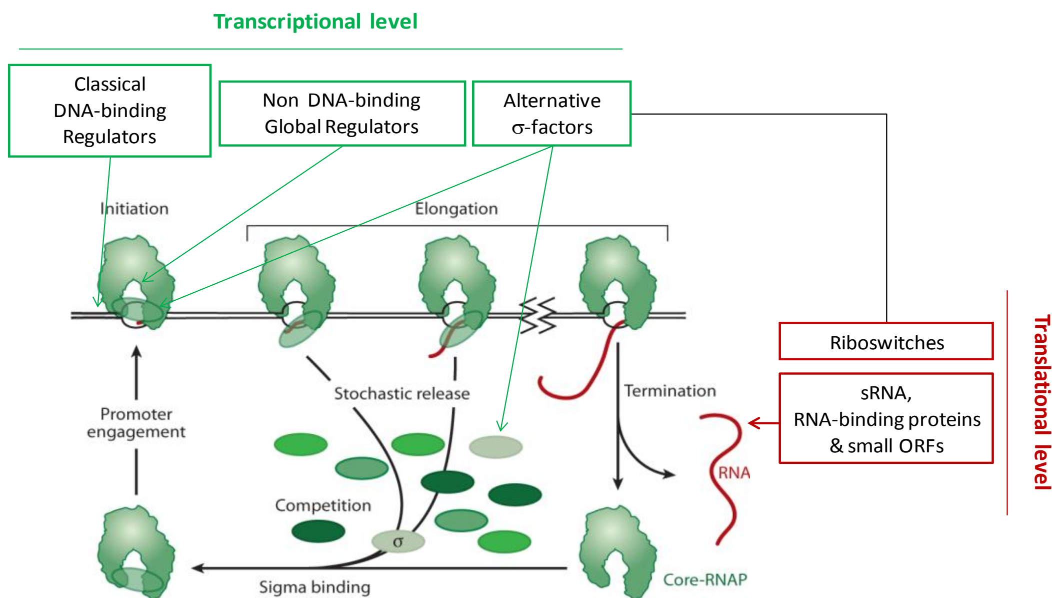


Figure 1 –The σ-cycle and regulation of gene expression in bacteria. Transcription is dived into the discreet steps of initiation, elongation, and termination. During initiation the holoenzyme RNA polymerase first forms a closed DNA complex with the promoter DNA, which is subsequently melted to form the transcription-competent open-promoter complex. Note that complete detachment of the σ -factor is not a *priori* for transition in to the elongation phase and a partially attached σ can cause elongation stalling by binding promoter-element mimics within the DNA. Nevertheless, the σ-factor is stochastically released during the first ~200 nt of elongation of the RNA. The released σ then joins the pool of free σ -factors that compete for association to core RNA polymerase. Thus, active transcription and the σ-cycle is the key process for re-orchestrating the composition of the holoenzyme pool within the cell. Pertinent examples of regulatory processes employed by bacterial cells to regulate transcription and subsequent translation of the RNA product are superimposed on the transcription cycle. Figure is adapted from (10). See text for details.

gene expression has remained an active field of research that continually reveals novel strategies to achieve control. Recently, new high-throughput technologies have demonstrated that many regulatory processes, traditionally thought as exceptions, seem to have more general applicability and may actually now be considered as common regulatory mechanisms (9). As depicted in Figure 1, on the transcriptional level these include proteins and other regulatory molecules that directly target RNAP to alter its function and internal promoters within operons that affect transcript levels of internal genes. In addition, translational coupling through small upstream open-reading frames (ORFs), RNA-binding proteins, riboswitches and small RNAs (sRNAs) can affect transcription and/or translation.

- **Sigma factors**

Bacteria use alternative sigma (σ) factors to redirect the transcriptional machinery to different classes of promoters in the genome (reviewed in (10)). The catalytic core RNA polymerase (core RNAP, subunit composition $\alpha_2\beta\beta'\omega$) is able to synthesize RNA from DNA, but cannot specifically recognise or initiate transcription from promoters. For the latter, core RNAP needs to be programmed by a σ-factor to form the holoenzyme (σ-RNAP) in which the σ-factor builds in the distinct promoter DNA-binding specificities.

All bacteria have a general (or "housekeeping") σ-factor (designated σ^{70} in *Escherichia coli* and *Pseudomonas putida*) that is responsible for transcription of the majority of the genes under rapid growth conditions. Many bacteria also encode additional "alternative" σ-factors. However, the number of alternative σs that each organism possesses differs, with the multiplicity generally reflecting the diversity of lifestyle and developmental characteristics of the organism. For example, enteric *E. coli* has just seven different σ-factors (11), while the soil and root coloniser *P. putida* has twenty-four, which probably contributes to the exquisite environmental adaptability of this organism (6).

To be able to form a holoenzyme, a σ-factor must be available so as to compete with other σ-factors for core RNAP (see Figure 1). The levels and availability of the different σ-factors in the cell are continually adjusted in response to the conditions the organism finds itself in. This control is achieved by multi-facetted regulation at the transcription, translation and protein stability levels, by signal-responsive release from "anti-σ-factors" that otherwise sequester specific σs from interaction with core RNAP, and by other processes that alter the

competition parameters between different σ-factors for core RNAP (reviewed in (10)). Hence, these changes in the composition of the holoenzyme pool determine the extent to which different promoter classes can be occupied, and it is against these dynamic changes in the holoenzyme RNAP pool that classical DNA-binding transcriptional activators and repressors must act (10).

The σ^{70}-family: Most alternative σ-factors are related in sequence and structure to σ^{70} and together constitute the so-called σ^{70}-family of proteins. All σ-factors of this family form holoenzymes that recognize distinct promoter signature sequences located approximately -35 and -10 positions from the transcription initiation site. The σ^{70}-family has been divided into four groups on the basis of phylogenetic relatedness (gene structure and function) (5,12). Members of the σ^{70}-family have up to four conserved sub-regions. Regions 2 and 4 are the most conserved regions and are present in all σs. These two regions are involved in recognition of the -10 and -35 promoter motifs, and are the only regions present in group 4 σs. Region 3 is involved in recognition of extended -10 promoter motifs, while region 1 is only found in group 1 household σs. This latter region is responsible for auto-inhibition of DNA binding of free σ^{70} and, within the context of the holoenzyme, can make DNA contacts just downstream of -10 motif at some promoters (reviewed in (13)). A non-conserved region of highly variable length intersperses region 2 of some housekeeping Group 1 σ-factors. For *E. coli* σ^{70} this region is thought to aid dissociation of the σ-factor and thus alleviate pausing caused by binding to promoter-mimic DNA during the early stages of transcription (14).

σ^{54} – in a class of its own: In addition to alternative σs of the σ^{70}-family, many bacteria also encode a σ^{54}-factor. The σ^{54}-factor, also called σ^{N}, is widely distributed among bacteria and differs in amino acid sequence and transcription mechanism from those of the σ^{70}-family. The unusual σ^{54}-factor also recognizes substantially different promoter motifs located -24 and -12 relative to the transcriptional start [consensus TTGGCACG-N4-TTGC]. Although bacteria frequently have several alternative factors of the σ^{70} family, two forms of σ^{54} rarely coexist in the same organism and orthologues of *E. coli* σ^{54} are the only members of the σ^{54} family. The σ^{54}-factor controls transcription of genes whose products have a wide range of different

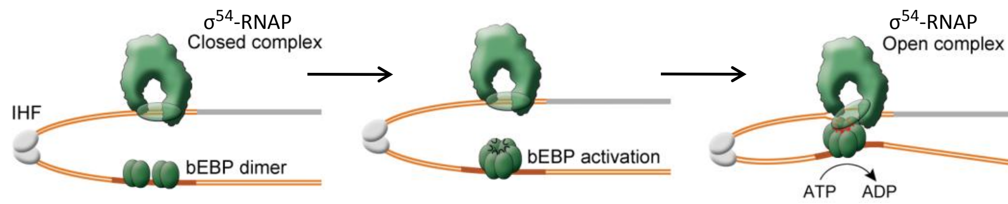


Figure 2 - Transcriptional initiation by σ^{54}-RNAP. Schematic illustration of the sequential steps of activation of σ^{54}-RNAP by bEBPs. The bEBP activators act on pre-bound σ^{54}-holoenzyme that is locked in a closed-complex. Multimerization of the bEBP to the transcriptional-promoting active form requires a bEBP-specific signal (e.g. phosphorylation, ligand-binding, or relief of repressive protein-protein interaction) and binding of ATP. Figure adapted from (17).

functions (15,16). In *P. putida*, this factor controls physiological processes such as the assembling of motility organs, expression of chemotaxis transducers, nitrogen assimilation, and the utilization of unusual carbon sources.

A key feature of σ^{54}-RNAP, which contrast σ^{70}-holoenzymes, is its inability to spontaneously melt (isomerise) double stranded DNA to form an open DNA promoter complex required for transcriptional initiation. This step strictly requires assistance from mechano-transcriptional activators that utilize ATP hydrolysis to drive conformational changes essential for this transition (17). Because these activators bind and activate transcription from sites located unusually far (80 to 200 bp) from the promoter (see Figure 2), they are usually referred to as bacterial enhancer binding proteins (bEBPs). Physical interaction between the bEBP bound to its enhancer site (also known as upstream activating sites, UASs) and the promoter-bound σ^{54}-RNAP requires looping out of the intervening DNA. At some σ^{54}-promoters, this process is facilitated by integration host factor (IHF), which induces DNA bending up to 160° at specific sites (18). Allied to the isomerization and transcriptional initiation by σ^{54}-RNAP is the activation of the hidden ATPase of the bEBP. These activators are usually present in an inactive dimeric form, and only take up their ATPase active and transcriptional promoting multimeric form when they receive an appropriate signal (reviewed in (17)).

- **(p)ppGpp and DskA**

In contrast to most other alternative σ-factors, the levels of σ^{54} in the cell are constant in *E. coli* and *P. putida* in different growth phases and under different growth conditions, and there is no known anti-σ-factor for σ^{54} (19,20). Control of σ^{54}-RNAP holoenzyme formation

appears primarily regulated by the alarmones guanosine tetraphosphate (ppGpp) and guanosine pentaphosphate (pppGpp) – collectively known as (p)ppGpp. These unusual nucleotides are probably the most global and far reaching bacterial signalling molecules known (21).

The levels of (p)ppGpp in the cell vary from very low, when cells are rapidly growing in rich medium (i.e. high-energy conditions), to very high when cells are cultivated in a medium with low quantities of nutrients or when they enter into the exponential-to-stationary transition phase (i.e. low energy conditions). The (p)ppGpp nucleotides are synthesized from GTP and GDP by RelA (ppGpp synthetase I) and the bi-functional SpoT protein (ppGpp synthetase II). RelA is associated with ribosomes. Under amino acid starvation conditions, where uncharged tRNAs bind to the ribosomal 'A' site and stalls protein synthesis, the synthetase activity of RelA becomes activated. The synthetase activity of the SpoT protein, which is able to both synthesizes and hydrolyzes (p)ppGpp, is induced by other stresses including deprivation of phosphate, iron, carbon source or fatty acids (21).

(p)ppGpp influences transcription by binding to RNAP and affecting the transcriptional initiation and elongation properties of RNAP (18). However, its most potent effects lie in altering transcriptional initiation at kinetically sensitive promoters – depending on the characteristics of the promoters the effect may result in an increase or decrease in productive initiation events (13). Because (p)ppGpp affects transcriptional initiation, it also influences the global transcription profile indirectly through altering the activity of promoters that control other global regulators and thus consequently causing regulatory cascade effects. One important indirect effect of ppGpp is its apparent capability to direct preferential use of alternative σ-factors, including σ^{54} (reviewed in (10)). Substantial evidence exists that this is bought about through (p)ppGpp-directed modulation of competition between σ-factors for core RNA polymerase to enhance the levels of alternative holoenzymes ((19,20) and references therein).

The RNAP-binding protein DksA frequently assists (p)ppGpp in both its direct and indirect effects on transcription. DksA is a member of a growing family of structurally, but not sequence, related proteins that directly access the active site cleft of RNAP through the secondary channel. In so doing, DksA mediates long-range structural changes within RNAP that alter interaction with the -6 to +6 region at σ^{70}-promoters (13,22). The mechanistic details of

how DksA binding to RNAP co-modulates stimulatory and inhibitory effects of (p)ppGpp on transcriptional initiation have not yet been determined. However, DksA clearly sensitizes RNAP to the cellular levels of (p)ppGpp to account for their synergistic action (23).

- **5'-Leading Regions (5'-LRs)**

5'-leading regions (5'-LRs) of mRNA are used to control the expression of some genes. In a number of cases the 5'-LR are long – up to hundreds of base pairs – and are the object of one or multiple regulatory mechanisms (24-26). 5'-LRs can form secondary structures that can, for example, act to block recognition sites for ribosome binding (RBS) or for mRNA-binding proteins in double stranded loops. Such occlusion mechanisms are often associated with other regulation process that result in signal-responsive alterations in the secondary structures that free the binding site for occupancy by the cognate molecule. Some mRNA-binding proteins are translational repressors that bind to specific sequences within the target mRNA. These sequences usually flank or overlap the RBS, thus preventing recognition and binding of the ribosome or, when located downstream, cause a road-block in translation (27,28).

Translation of small peptides encoded within the 5'-LR can also exert control – at the transcriptional levels (by regulating premature termination) or – at the translational level (through translational coupling and/or by disrupting secondary structures). Classical examples are transcriptional attenuation mechanisms of amino acid biosynthetic operons (24-26) that allow the cell to sense amino acid levels and regulate expression accordingly. A conceptually analogous and probably even more frequent mode of regulation is through riboswitches located within 5'-LRs. Riboswitches are mRNA regions that recognise and bind small molecules (cellular metabolites), and in so doing induce changes in the structure which in turn affects transcription, translation or mRNA stability (reviewed in (29)).

- **Small RNAs (sRNAs)**

As alluded to above, occlusion mechanisms in which target sites for binding are masked within secondary structures of the RNA, are frequently coupled to signal-responsive control of the secondary structure of the mRNA. In many cases this is mediated by small RNAs, usually non-coding RNAs of 50-250 nt. Binding of the sRNAs alters the secondary structure to modulate transcription, translation, and/or change mRNAs half-life (25). Although

a small minority, some sRNAs encode small peptides; however, only few have been described as functional and their function is far from completely understood ((26) and references there in).

sRNAs can target proteins or act by antisense mechanisms. In the latter case, they can be further divided in *cis*-acting, if the target gene is within the same locus, or *trans*-acting if it targets one or multiple genes widely distributed in the genome. Unlike *cis*-encoded sRNAs, the regulatory sequences of *trans*-encoded sRNAs are usually short and the complementarily with the target(s) mRNA is not necessarily perfect. Typically, these molecules act as negative regulators of target genes by binding to mRNAs and forestalling recognition of the RBS by the ribosome, or by preventing translation to proceed. However, they can also have a positive effect by binding to mRNA leading regions and inducing a shift from a secondary structure where the RBS is sequestered within a double-stranded portion of the mRNA to a structure where the RBS is freely accessible for ribosome recognition (reviewed in (30)).

sRNA that target proteins appear to be comparatively rare; nevertheless, the ones described to date target mRNA-binding proteins and thereby regulate their action. In these cases, the sRNAs usually possess multiple binding sites for the protein, thus sequestering the protein from binding to the mRNAs they regulate.

1.2. *Pseudomonas putida*

The *Pseudomonas* genus is characterized by the capacity of these organisms to live in a wide range of environmental niches such as soil and water ecosystems and in association with plants and animals, including humans. This ability is primarily due to their versatile metabolism and regulatory processes that finely control preferential utilization of different carbon sources through carbon catabolite repression (31). *P. putida* strains are Gram-negative rod-shaped soil bacteria which frequently inhabit the rhizosphere of plants. As with other members of the Pseudomonad group, *P. putida* strains are able to grow on the expense of a variety of carbon sources. In some strains their catabolic capacity is expanded through possession of plasmids that encode the ability to degrade toxic pollutants such as naphthalene, phenol, methylphenols, and toluene (32-34). This capacity lends these species as useful for bioremediation of these compounds from soil (35,36).

1.3. Carbon Catabolite Repression

To face changes in carbon-source availability bacteria have global regulation systems that allow them to coordinate the expression of different metabolic pathways in order to optimize efficiency and ecological fitness. One of these regulatory systems is carbon catabolite repression (CCR) or catabolite repression control (CRC). These regulatory processes allow bacteria to preferentially assimilate the compound that provides the most efficient growth when several carbon sources are available. Strategically, these systems operate by repressing the genes that encode proteins needed for the catabolism of non-preferred compounds until conditions cue the need for their activity.

Although CCR has been observed in most free-living bacteria with versatile metabolisms, the molecular mechanism involved can vary greatly from one species to another (37,38). For example, the major CCR mechanisms in *Pseudomonas* are quite different from the one described in *E. coli*. For *E. coli*, as for the majority of enteric bacteria, the preferred carbon source is glucose. Therefore, in this organism the main CCR mechanism involves the $EIIA^{Glc}$ protein, which is a specific glucose kinase that also mediates its passage through the cytoplasmic membrane. $EIIA^{Glc}$ also indirectly regulates expression of genes responsible for the metabolism of other carbon sources by inducer exclusion and by controlling cAMP levels, which in turn controls the DNA-binding ability of the global transcriptional regulatory protein CRP (37,38). For *Pseudomonas,* on the other hand, glucose and other hydrocarbons are not the preferred carbon source; instead some organic acids and amino acids are. The CCR mechanisms known in *Pseudomonas* are, therefore, also different and involve the Cyo terminal oxidase system, the PTS^{Ntr} system and the Crc protein (reviewed in (37,38)).

- **The *Pseudomonas* Crc – Catabolite repression control protein**

As carbon sources, Pseudomonads prefer amino acids over hydrocarbon but different amino acids are not all equally preferred. Therefore, the cell needs to repress the assimilation of non-preferred ones as well as favouring the ordered assimilation of amino acid or organic acid if several are available. Crc is the regulatory factor responsible for this organised assimilation to result in efficient use of the available nutrients by optimization of metabolic-flow, which consequently leads to a maximum growth rates (39).

Crc also controls other important aspects of the cell biology in these organisms. In *P. aeruginosa*, an opportunistic human pathogen, Crc is necessary for the formation of fully functional type IV pili that are mandatory for twitching motility and biofilm formation that is associated with chronic colonization of human tissues (40).

Crc is a RNA-binding protein that regulates gene expression at the post-transcription level. Crc binds to short unpaired A-rich sequences (consensus AAnAAnAA) in the 5'-end of mRNAs; this prevents translation without destroying the mRNA by cleavage or degradation (41,42). The precise mechanism that results in inhibition of translation is still uncertain; it is likely that the formation of the ribosomal-mRNA complex is inhibited by competition for closely located binding site, but it is also possible that Crc does not prevent binding of the 30S ribosomal subunit *per se,* but rather can trap it in an inactive form that is not able to continue the translation process (41).

Crc levels and availability vary according to the growth status of the cell. For example, when *P. putida* cells are cultivated in rich media, Crc levels levels are four- to five-fold higher at mid-exponential phase than in stationary phase (43). Although the ultimate signal(s) that control Crc levels in the cell are not yet known, levels are adjusted in response to the carbon source present in the medium (43-45). Moreover, even when Crc is present, its availability can be controlled through sequestering by sRNAs. In *P. aeruginosa*, it has been found that a sRNA of 407 nt (named CrcZ) contains five exposed Crc motifs, and that CrcZ sequesters Crc. As a result, in this organism, levels of free Crc are modulated in order to control the strength of CCR effect that is dependent on Crc (42). Two CrcZ-like RNAs are encoded in the KT2440 genome, suggesting the presence of a similar sequestering mechanism for Crc in this organism.

1.4. Introduction to the *dmp*-experimental system

The extensively studied *dmp*-system of the pVI150 plasmid of *P. putida* CF600 provides the capability to grow on the expense of phenol and methylphenols (6,46). The *dmp*-system is controlled by two promoters dependent on different σ-factors, see Figure 3. The non-overlapping σ^{70}-Pr and σ^{54}-Po promoters are located within a 406 bp intergenic region. The DmpR-regulated σ^{54}-Po promoter drives transcription of the *dmp*-operon that contains fifteen *dmp* structural genes which encode a multicomponent phenol hydroxylase and a subsequent

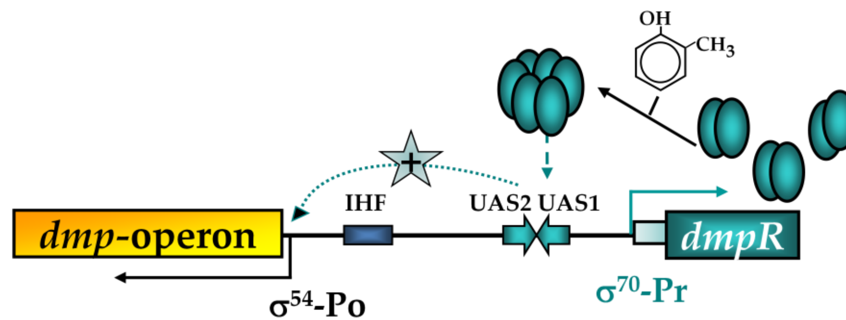


Figure 3 - The *dmp*-regulatory circuit. Schematic illustration (not to scale) of the locations of the Po and Pr promoters as well as DNA binding sites for DmpR (UASs) and IHF.

meta cleavage pathway that are responsible for complete metabolize phenolic compounds (reviewed in (47)). The Dmp-enzymes are only required when the pathway substrates are present. Specific regulation in response to the presence of phenolics is achieved through the sensor-regulator DmpR, a 63.3 kDa protein encoded by *dmpR* gene. The *dmpR* gene is transcribed from a σ^{70}-Pr promoter and encodes a bEBP that binds *dmp*-pathway substrates (or structural analogues) to take up its active multimeric form (48-52). Thus, it is only when pathway substrates are present that the σ^{54}-Po promoter is active and the *dmp*-operon encoded enzymes are expressed.

The σ^{54}-Po promoter belongs to the class of promoters recognized by the σ^{54}-RNAP. As mentioned previously, this class of promoters is sensitive to the indirect influence of (p)ppGpp and DksA through their effects on σ^{54}-RNAP holoenzyme levels. (p)ppGpp and DksA also affect the output of σ^{54}-Po by stimulating transcription of IHF genes (53), which in turn enhances productive interaction of DmpR (bound to UAS1 and UAS2) and Po-bound σ^{54}-RNAP (6,19). A final direct input of (p)ppGpp and DksA in this regulatory circuit is their stimulatory effect on σ^{70}-RNAP activity at the Pr-promoter (53). The direct stimulatory effect of (p)ppGpp/DksA at the Pr promoter has been traced to its extremely suboptimal -10 element, more specifically the lack of A residue at the -11 position. Pr requires (p)ppGpp and DksA to both stimulate binding of σ^{70}-RNAP and to accelerate the rate of open-complex formation (54).

It has previously been shown that divergent but non-overlapping transcription from σ^{54}-Po promoter stimulates transcription from the σ^{70}-Pr promoter of *dmpR* (53). Because σ^{54}-Po activity is dependent on DmpR, this interplay between σ^{54}-dependent and σ^{70}-dependent transcription generates a feed-forward loop in which DmpR stimulates its own synthesis as

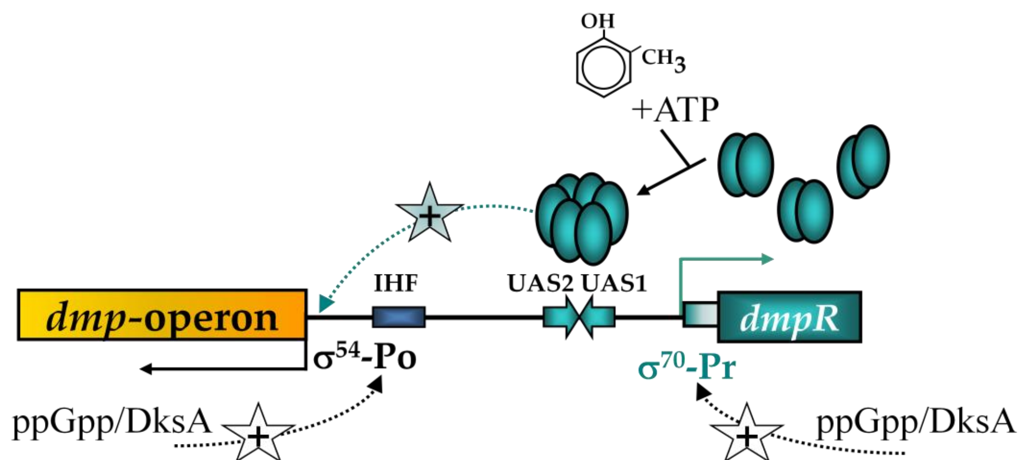


Figure 4 – Model of aromatic-effector activation of DmpR and a consequence feed-forward loop (blue stars) mediated through interplay of the Po and Pr promoters and the effects of (p)ppGpp/DksA; adapted from (53), see text for details.

illustrated in Figure 4. This feed-forward loop, which is integrated with (p)ppGpp production upon stress and/or nutrient limitation as described above, both reinforces the silencing or the transcription of the σ^{54}-Po promoter under high energy conditions (low ppGpp levels) when this auxiliary pathway is not needed, and promotes high level productions of the enzymes under low energy conditions when an appropriate substrate is present.

1.5 Negative regulation by the extensive and conserved 5'-LR

In addition to the positive stimulatory action of (p)ppGpp/DksA via the feed-forward loop, the levels of DmpR are also modulated through an initially transcribed region of the *dmpR* gene that encodes a 123 nt long 5'-LR. This effect was firstly detected during monitoring of transcription from the Pr promoter using luciferase (Lux) transcriptional reporters that either had the 5'-LR or lacked this DNA (53). The presence of the 5'-LR has a 5- to 7-fold inhibitory effect on the number of transcripts from the Pr promoter as assessed by transcriptional reporter gene assays and quantitative RT-PCR (see Figure 5). Moreover, the mRNA stability is not affected by the absence or presence of the 5'-LR region (V. Shingler, personal communication). As outline in preceding sections, 5'-LRs of mRNA are known to affect gene expression by a wide variety of mechanisms. How the 5'-LR of *dmpR* is involved in controlling the levels of DmpR is the subject of this thesis.

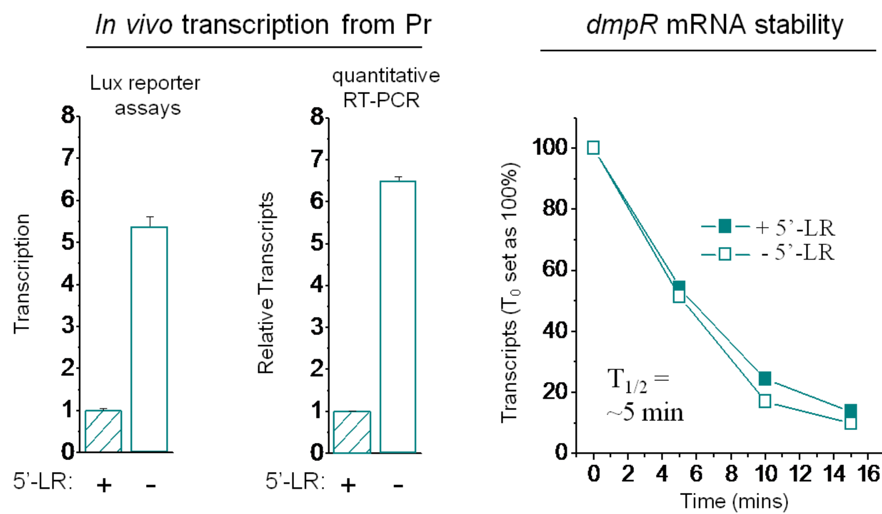


Figure 5 – Effects of the 5' -LR of *dmpR* on the level of transcripts from the Pr promoter. The two bar charts on the left compare the *in vivo* transcription values from the Pr promoter in the presence (+) and absence (-) of DNA encoding the 5'-LR of *dmpR*, obtained by quantitative transcriptional reporter assays and quantitative RT-PCR. Data are from early stationary phase cultures, the same fold-inhibition is observed throughout the growth curve. The graph on the right shows data from a comparison between the half-lifes of the transcripts from these reporters with and without the 5'-LR. Data were obtained after adding rifampicin to early stationary phase cultures to prevent new RNA synthesis (0 time point), RNA extraction at the indicated time points, followed by quantification of mRNA levels by qRT-PCR. Unpublished data courtesy of V. Shingler.

2. Aims

To avoid disadvantageous energy-fluxes, regulation of natural pathways for toxic aromatic compounds is attuned to host physiology and the energy status of the cell. Because such compounds are toxic even to bacteria that can degrade them, the regulatory circuits that control the expression of the specialized catabolic enzymes interface both metabolic- and stress-responsive regulatory networks. Failure to engineer such integration into specifically designed synthetic pathways destined for use in environmental clean-up is one of the major reasons for the partial or complete failure of some systems to perform their task under field conditions (55).

Available data suggest that population dynamics and physiological control of catabolic gene expression prevail over any artificial attempt to engineer optimal performance of the desired catalytic activities in determining a successful outcome (55). Therefore, to be able to rationally design bacteria for bioremediation applications, a full understanding of the multiple mechanisms which couple the expression of these pathways to host physiology is required. In the case of *Pseudomonas putida* CF600, expression of the master regulator DmpR is the major check point. As overviewed in the introduction, multiple mechanisms involving the bacterial alarmone (p)ppGpp converge to stimulate transcription from the intrinsically weak Pr promoter under stress conditions. On the other hand, a 5'-LR seems to also have an important role in the regulation of this system.

Therefore, the overall objective of this thesis was to unveiling the regulatory mechanism(s) associated with the 5'-LR. As the research progressed, the specific aims became:

I) To define the region of the 5'-LR that mediated transcriptional repression through the activity of the Pr promoter.

II) To elucidate if the Crc mRNA-binding protein mediates control of DmpR levels through modulation of translation.

III) To determine if an open-reading frame and/or potential target site for a small regulatory RNA (PhrS) – which both lie upstream of the *dmpR* coding region in the mRNA – have roles in directing the levels of DmpR produced.

3. Methods

3.1. Bacterial strains and culture conditions

P. putida and *E. coli* strains (Table 1) were cultured at 30°C and 37°C respectively, in Luria-Bertani / Lennox medium (LB, AppliChem GmbH) or M9-minimal salts medium (56) supplemented with appropriate antibiotics (Table 2) for the strain and/or plasmid selection.

Table 1 – Bacterial strains used in this study

Strain	Relevant Properties	Source [A] or Reference
Escherichia coli		
EC51	DH5; prototrophic	(57)
EC1655	S17λ*pir* Tp^R; strain that provides mobilization functions from the chromosome	(58)
Pseudomonas putida		
PP2	KT2440 mt-2, prototrohic parent strain	(59)
PP980	KT2440Km^R, Po-*luxAB* transcriptional reporter cassette on the chromosome	(60)
PP3044	KT2440Km^R *crc*::Gm Po-*luxAB*; Crc null derivative of PP980	lab collection
Pseudomonas aeruginosa		
PAO1	Prototrophic parent strain	(61)
PAO6671	PAO1 Δ*PhrS;* PhsR null strain	(26)

[A] Shingler "lab collection" strains are as yet unpublished and were constructed by co-workers.

Table 2 – Antibiotics used to supplement the media

Antibiotic	Concentration for *E. coli* (μg ml^{-1})	Concentration for *P. putida* and *P. aeruginosa* (μg ml^{-1})
Carbenicillin (Cb)	100	1000
Trimethoprim (Tp)	100	-

3.2. Plasmid constructions

Plasmids (Table 3) were introduced into *E. coli* strains by transformation, into *P. putida* strains by electroporation, and introduced into *P. aeruginosa* by conjugation from *E. coli* S17λ*pir*. Standard DNA cloning techniques were used to generate the plasmids constructed in this work. Fidelity of DNA regions generated by PCR or by the insertion of synthetic double stranded linker DNA was confirmed using a Big Dye terminator sequencing kit (Applied Biosystems) and primer 135 listed in Table 4.

Table 3 – Key plasmids used in this study

Plasmid	Relevant properties [A]	Source[B] or Reference
Cb^R narrow-host-range, high copy number plasmids		
pBlueScriptSK+	Multipurpose cloning vector	Stratagene
pVI399	-555 to +1842 Pr 5'-LR WT *dmpR* (WT)	(48,60)
pCon2117	-555 to +1842 Pr 5'-LRΔ9-101 *dmpR* (Δ8)	lab collection
pCon2147	-555 to +1842 Pr 5'-LR– *dmpR* (mutD)	lab collection
pCon2202	-555 Pr 5'-LR39TAA41 *dmpR* (stop)	This study
Cb^R broad-host-range vectors and luciferase (*luxAB*) promoter probe vectors		
pMMB66HE	RSF1010-based *$lacI^q$*-P_{tac} expression vector; polycloning site: HindIII/PstI/SalI/BamHI/SmaI/EcoRI	(62)
pMMB66EH	RSF1010-based *$lacI^q$*-P_{tac} expression vector; polycloning site: EcoRI/SmaI/BamHI/SalI/PstI/HindIII	(62)
pVI398	pMMB66HE derivative deleted of *$lacI^q$*-P_{tac} ; polycloning site: PstI/NotI/BamHI/SmaI/EcoRI	(60)
pCon1640	pMMB66 derivative deleted of *$lacI^q$*-P_{tac} and carrying the promoterless *luxAB* genes downstream of a polycloning site: EcoRI/SmaI/XhoI/BglII/NotI/SpeI/StuI/SacI used for -38 to +127 and -266 to +127 Pr-*luxAB* derivatives below	lab collection
pCon2126	As above but with polycloning site: SmaI/SaI/BglII/NotI/KpnI/NdeI/SpeI/StuI/SacI used for -555 Pr-*luxAB* derivatives below	lab collection
Cb^R broad-host-range *luxAB* transcriptional reporter plasmids		
pVI466	Pr-*dmpR* Po-*luxAB* reporter (native configuration) on pVI398	(63)
pCon1565	-38 to +1 Pr 5'-LR (WT) *luxAB*	lab collection
pCon 1850	-38 to +127 Pr 5'-LR (WT) *luxAB*	lab collection
pCon 1631	-38 to +127 Pr *luxAB* with a BglII site downstream from +1	lab collection
pCon 1586	-38 to +127 Pr (+1 to +127 random DNA) *luxAB*	lab collection
pCon 1633	-38 to +127 Pr (+1 to +127 in opposite orientation) *luxAB*	lab collection
pCon 1598	-38 to +127 Pr 5'-LR Δ9-78 *luxAB* (Δ3)	lab collection
pCon 1599	-38 to +127 Pr 5'-LR Δ9-93 *luxAB* (Δ4)	lab collection
pCon 2100	-38 to +127 Pr 5'-LR Δ9-108 *luxAB* (Δ5)	lab collection
pCon 2101	-38 to +127 Pr 5'-LR Δ9-93, Δ109-123 *luxAB* (Δ6)	lab collection
pCon 2114	-38 to +127 Pr 5'-LR Δ9-105 *luxAB* (Δ7)	lab collection
pCon 2115	-38 to +127 Pr 5'-LR Δ9-101 *luxAB* (Δ8)	lab collection
pCon 1817	-266 to +127 Pr 5'-LR WT *luxAB* (WT)	lab collection
pCon 1821	-266 to +127 Pr 5'-LR 74AAAAAA79 *luxAB* (mutA)	lab collection
pCon 1822	-266 to +127 Pr 5'-LR 91AAAA94 *luxAB* (mutB)	lab collection
pCon 1833	-266 to +127 Pr 5'-LR 77AAAAAA82 *luxAB* (mutC)	lab collection
pCon 1841	-266 to +127 Pr 5'-LR 83AAGTAA89 *luxAB* (mutD)	lab collection
pCon 1843	-266 to +127 Pr 5'-LR 77AAAAAAAAGTAA89 *luxAB* (mutE)	lab collection
pCon 1485	-555 to +127 Pr 5'-LR WT *luxAB*	lab collection
pCon 2204	-555 to +127 Pr 5'-LR 39TAA41 *luxAB* (NotI to NdeI fragment from pCon2202 into pCon2126)	This study

Table 3 continued on next page

Table 3 – continued

Plasmid	Relevant properties [A]	Source [B] or Reference
Cb^R broad-host-range DmpR expression plasmids		
pVI401	-555 to +1842 Pr 5'-LR WT *dmpR* (WT-NdeI)	(60)
pCon2119	-555 to +1842 Pr 5'-LR Δ9-101 *dmpR* (Δ8)	lab collection
pCon 1483	-555 to +1842 Pr 5'-LR Δ9-78 *dmpR* (Δ3)	lab collection
pCon 1482	-555 to +1842 Pr 5'-LR Δ9-93 *dmpR* (Δ4)	lab collection
pCon 2203	-555 to +1842 Pr 5'-LR39TAA41 *dmpR* (NotI fragment from pCon2202 into pVI398)	This study
pCon 2150	-555 to +1842 Pr 5'-LR 83AAGTAA89 *dmpR* (mutD) (NotI fragment from pCon2147 into pVI398)	This study
pCon 2127	-266 to +1842 Pr 5'-LR WT *dmpR*	This study
pCon 2128	-266 to +1842 Pr 5'-LR 74AAAAAA79 *dmpR* (mutA)	This study
pCon 2129	-266 to +1842 Pr 5'-LR 91AAAA94 *dmpR* (mutB)	This study
pCon 2132	-266 to +1842 Pr 5'-LR 77AAAAAA82 *dmpR* (mutC)	This study
pCon 2133	-266 to +1842 Pr 5'-LR 83AAGTAA89 *dmpR* (mutD)	This study
pCon 2134	-266 to +1842 Pr 5'-LR 77AAAAAAAAGTAA89 *dmpR* (mutE)	This study

[A] Co-ordinates are relative to the +1 transcriptional start from Pr. Deletions (Δ) or substitutions of the 5'-LR (1 to 123) are indicated; WT denotes wild-type. The abbreviated designations in brackets are those used in the Figures.
[B] Shingler "lab collection" plasmids are as yet unpublished and were constructed by co-workers.

Table 4 – Oligonucleotides used during this study

Number	Sequence [A]	Purpose
2646f	5'-CGGATCCGCAGGACATCAAGCAACGGC	Amplification of BamHI to NdeI fragments - constructs pCon2127-2134
2647r	5'-CGCGCATATGAGCGAGGCCCCTATTTATTT	
2646f	As above	Overlapping PCR mutagenesis: internal HindII to BglII fragment used to replace the wild-type region pVI399 to generate pCon2202
2846int	5'-CCCC**ATGTAA**CCATCTGGAATCGCCGCCTGCCT	
2845int	5'-CGATTCCAGATGG**TTACAT**GGGGAAAATCGGCAGT	
354r	5'-CAGATTTCCACCTCGAAGGAGTC	
135f	5'-CTTTTTAAGCATTTGATCAATTGCC	Sequencing verification for faithful PCR amplification and linker insertions

[A] Restriction sites artificially introduced via oligonucleotides are underlined, while initiation and termination codons are shown in bold.

3.3. Crc protein for IVTT reaction

P. putida Crc with an N-terminal His tag was a generous gift from Fernando Rojo. Crc was provided in high salt / 30% glycerol buffer and stored at -80°C until use. Before added to the buffer-sensitive IVTT reactions, aliquots of the proteins were first equilibrated into 10 mM Tris-HCl pH 7.6 using BioRad P30 spin columns. Concentration of stock and buffer-exchanged proteins was determined using the Pierce BCA protein assay kit (based on the biuret reaction) following the microtiter well plate protocol. Standard curves were obtained used BSA-acetylated stock (initial concentration of 20 mg/ml; Ambion), diluted with the corresponding buffer of the protein preparation.

3.4. Electro-mobility shift assays (EMSA)

EMSA analysis of Crc binding to RNA was performed by Fernando Rojo and colleagues essentially as described in (41), using increasing concentrations of *P. putida* Crc-His and an RNA probe 5'-UUCCCCAUCU**AAAAAUAA**AUAGGGGC-3' (encompassing the suspected Crc-binding region of the 5'-LR of the *dmpR* mRNA, which is highlighted in bold). 20 μl reactions (in 10 mM Hepes-KOH pH 7.9, 35 mM KCl, 2 mM $MgCl_2$) and containing 1 μg yeast tRNA, 0.1 nM [^{32}P]-labelled RNA probe were supplemented with the indicated concentration of Crc-His (0, 53, 106, 212, 425, and 850 nM). These levels of Crc-His do not alter the migration of an unrelated RNA probe lacking a Crc site (Fernando Rojo, personal communications). Reactions were incubated for 30 mins at 20°C prior to addition of 4 μl of loading buffer (60% glycerol, 0.025% xylene cyanol) and analysis on a non-denaturing 4% polyacrylamide gel containing TMB buffer (45 mM Tris-HCl, pH 8.3, 43 mM boric acid, 2 mM $MgCl_2$, 5% glycerol). Electrophoresis was performed in TBM buffer at 4°C, and the results documented by exposure of the dried gel to X-ray film.

3.5. *In vitro* transcription-translation (IVTT) assays

IVTT assays were performed using the Promega *E. coli* S30 extract kit for circular DNA. Reactions (total volume 25 μl) contained S30 extract (7.5 μl), reaction pre-mix (10 μl) and all

amino acids except methonine (2.5 µl) as recommended by the manufacturer. Reactions were additionally supplemented with a) 1 µg of plasmid DNA (in 2 µl) , b) L-[^{35}S]-Methionine (5 µCi at >1000 Ci/mmol, Perkin Elmer; 0.75 µl), c) RNAse inhibitor (Ambion; 0.25 µl) and d) 2 µL of 10 mM Tris-HCl pH 7.6 containing or not Crc (to the final concentrations indicated). Reactions were incubated at 37°C for 60 mins at which point 5 µl samples were taken and proteins precipitated by the addition of 20 µl of ice cold acetone. After 15 mins on ice, tubes were centrifuged at 12,000 rpm for 5 mins and the recovered air-dried protein pellet re-suspended and heated to 100°C for 5 mins in final sample buffer (125 mM Tris pH 7.5, 2.5% SDS, 8% glycerol, 10% β-mercaptoethanol and 0,1% bromophenol blue) prior to separation on an 11% SDS-polyacrylamide gel. After electrophoresis, the gel was stained with coomassie blue, de-stained, dried and newly synthesised proteins detected using AGFA medical X-ray film.

3.6. *In vivo* Luciferase assays

To perform luciferase plate tests, single colonies of the strains were streaked into LB plates containing Cb and 2-methylphenol and allowed to grow over night. The next day, approximately 100 µl of 1:100 diluted decanal solution was dispensed in the plate lid and the inverted plates were incubated for some minutes; light emission was then acquired by exposing AGFA medical X-ray films.

Quantitative luciferase *luxAB* reporter assays were performed on cultures grown and assayed at 30°C as described in (64). Overnight cultures were diluted 1:50 and grown into exponential phase followed by a second dilution to a final OD_{600} between 0.05-0.08 to ensure a balanced growth of all strains prior to initiation of the experiment. After the second dilution, when required for DmpR activity, 2 mM (in case of *P. putida*) or 0.5 mM (in case of *P. aeruginosa*) of 2-methyphenol was added to the cultures. A 100 µl samples were collected of at the indicated time points. Optical density (OD at 600 nm) and light emission of samples were measured using an Infinite M200 luminometer (Tecan); for light emission, a 100 µl of a 1:2000 diluted decanal solution was dispended as substrate. Samples were taken every 45 minutes for approximately 9 hours. Data points are the average of duplicate determinations from each of two or more independent cultures ± standard errors. Data treatment was performed using Origin software.

3.7. Western analysis

Samples were collect from cultures of luciferase assays at time points 3.5 h and 7.5 h (that correspond to exponential and stationary phase, respectively). The procedures for preparation of crude extract, SDS-PAGE, and electrotransfer to Amersham Biosciences PVDF membrane (0.45 mm, Amersham Biosciences) were as previously described (49). Anti-DmpR antibodies were affinity purified polyclonal rabbit antibodies raised against the N-terminal 232 residues of DmpR (19). Antibody-decorated bands were revealed using ECL-Plus reagents (GE Healthcare) following the instructions of the manufacture, and documented using AGFA medical X-ray films.

4. Results

4.1. DNA encoding the 5’-LR mediates inhibition at the level of transcription

The presence of DNA encoding the 123 nt long 5’-LR of *dmpR* results in a decreased number of transcripts originating from the Pr promoter, as assessed using transcriptional reporter gene assays and qPCR, through a mechanism that is independent of mRNA stability (Figure 5). To further investigate the mechanism underlying the repressive effect of the 5’-LR on Pr output, a series of luciferase transcriptional reporter plasmids were constructed in which the promoterless *luxAB* genes were placed under the control of the Pr promoter and downstream of variants of the 5’-LR ((65), see Figure 6).

To enable manipulation of DNA encoding the 5’-LR, a unique BglII site was introduced immediately downstream of the +1 transcriptional start from Pr. This allowed construction of derivatives with i) a reconstituted 5’-LR in its native orientation, ii) with the 5’-LR DNA inverted, or iii) with the 5’-LR DNA substituted by random DNA. As summarised in Figure 6, luciferase plate tested showed that the 5’-LR DNA exerts a repressive effect in either orientation, but that repression is sequence specific because random DNA cannot mimic the effect.

The next step was to map the region responsible for the repressive effect. Therefore, a

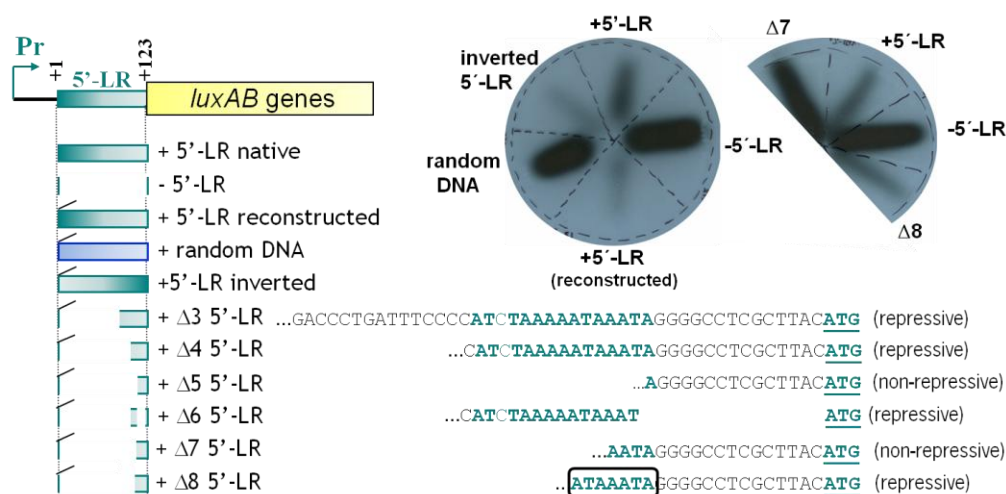


Figure 6 – Mapping the repressive effect of the 5’-LR on Pr activity. Left shows a schematic of a series of constructions used to analyse and trace the sub-region of the 5’-LR responsible for the repressive effect on output from Pr. The upper autoradigrams show representative results of luciferase plate test as examples of repressive and non-repressive derivatives. The sequences of the 5’-LR present in six key deletions derivatives are given with the bases common between Δ3, Δ4, Δ5 and Δ8 that maintain the repressive effect highlighted in cyan. The minimal sequence deduced as mediating the repressive sequence is outlined by a black box superimposed on the sequence of Δ8.

series of deletions were constructed (see Figure 6). The Δ3 and Δ4 constructions, in which promoter proximal parts of the *dmpR* 5'-LR were deleted, maintained the repressive effect, thus tracing the responsible region to the promoter distal portion of the 5'-LR. This region was further reduced to 15 bp after observing that Δ5, in which the promoter proximal region was further reduced, no longer maintained the repressive effect, while Δ6, which has the same deletion of the promoter proximal part as Δ4 combined with a 15 bp deletion just upstream the ATG start, still maintained it. Although Δ7 has a smaller deletion than Δ5, it did not manifest the repressive effect. However, with Δ8, which possesses just 3 additional bases more than Δ7, the repressive effect was restored. Taken together, the data trace the region responsible for repressing Pr output to a short AT-rich DNA region – ATAAATA – present in all the derivatives that exhibit the repressive phenotype but absent in those that do not.

4.2. The 5'-LR controls expression of DmpR at the level of translation

Because the *luxAB* reporter genes possess their own ribosome binding site (RBS), the transcriptional analysis described above only documents effects at the level of transcription from the Pr promoter. To determine if the 5'-LR also plays a role at the level of translation, the Δ8 deletion – the largest deletion that still maintained the wild-type repressive effect on transcription from Pr – was engineered into a Pr-*dmpR* expression plasmid (pCon2119). The levels of DmpR produced from this construct were compared to those produced by an otherwise identical Pr-*dmpR* expression plasmid containing the native 5'-LR (pVI401). These two plasmids were independently introduced into PP980, a *P. putida* KT2440 strain that carries a Po-*luxAB* transcriptional cassette on the chromosome. As schematically illustrated in Figure 7A, because activity of the σ^{54}-dependent Po promoter is strictly dependent on DmpR, luciferase activity from the Po-*luxAB* cassette indirectly reports on the *in vivo* levels of DmpR.

Quantitative luciferase assays of both strains, cultivated in rich liquid media in the presence of 2-methylphenol (a stronger inducer of DmpR activity), was followed for nine hours. The result (Figure 7B) showed that the activity of Po varied greatly during the exponential phase, suggesting that DmpR levels from the Δ8 were likewise elevated during this phase of growth. That this was indeed the case was confirmed by quantitative Western analysis, which showed that exponential phase levels of DmpR produced by the Δ8 derivative were 2- to 4-fold

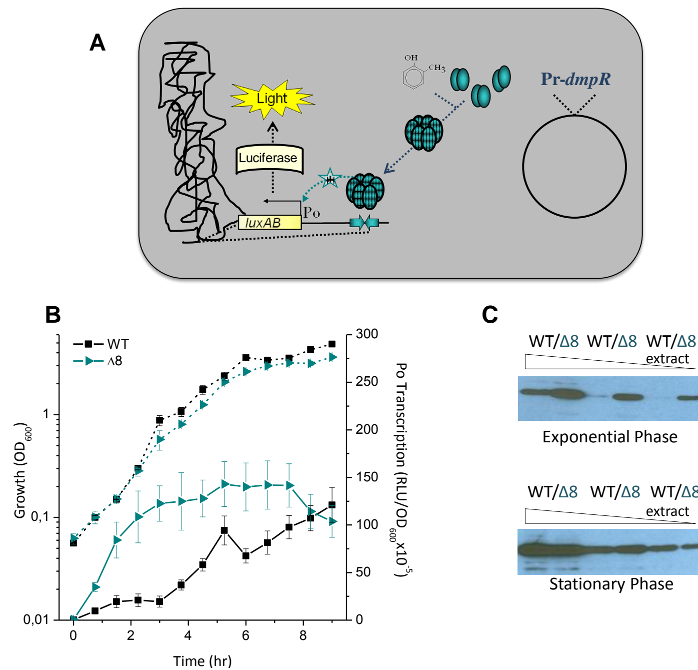


Figure 7 – *In vivo* effect of the 5'-LR at the translational level. (A) Schematic of the experimental system used to monitor DmpR levels through output from the DmpR-dependent Po promoter. **(B)** Graph showing the luciferase activity profiles (continuous lines) and corresponding growth curves (dashed lines) for the strain Po-*luxAB* reporter strain PP980 carrying plasmids encoding the wild-type Pr-*dmpR* fusion (pVI401; black squares) and the Pr-Δ8-*dmpR* construction (pCon2119; dark cyan triangles). **(C)** Western blot analysis of DmpR levels; cells were harvested at time points 3.5 h and 7.5 h (exponential and stationary phase respectively) and 20 μg, 10 μg and 5 μg of soluble protein extract were separated on 12% SDS-PAGE gels.

higher than those produced by the wild-type counterpart (Figure 7C). This contrasts the results from stationary phase cultures in which DmpR levels are very similar. This data strongly indicates the existence of a second regulatory role of the 5'-LR, namely control at the translation level during the exponential phase of growth.

4.3. Crc controls translation of DmpR through binding to the 5'-LR

Many different mechanisms could account for control of DmpR translation through the 5'-LR of its mRNA. However, three observations suggested Crc as a likely candidate for mediating the observed effect. Firstly, the mRNA-binding Crc protein is responsible for hierarchal control of carbon source catabolism. Secondly, Crc levels (and availability) vary during different phases of growth, being high during the exponential phase (when the inhibitory effect through the wild-type 5'-LR was observed) but only at basal levels during stationary

phase (when the effect is no longer observed). Thirdly, the 5'-LR sequence posses of a potential Crc RNA-binding motif (Figure 8A) that i) overlaps the RBS region and ii) would be destroyed in Δ8. Thus, if Crc does influence this system, then the Δ8 derivative would represent a Crc target-site deficient mutant.

To test for the potential involvement of Crc, a Po-*luxAB* reporter plasmid that also carried the Pr promoter, the 5'-LR, and *dmpR* in its native configuration was used (Figure 8B). This plasmid was introduced into wild-type *P. putida* KT2440 and an otherwise identical Crc null counterpart. The result (Figure 8C) showed that DmpR-dependent Po output is greatly affected by the lack of Crc, particularly during the exponential phase where levels are up to 22-fold higher than in the wild-type strain. It is notable however, that the Crc null strain has a marked slower growth rate. This would be predicted to increase (p)ppGpp levels within the cell. Thus, part of the observed effect in the Crc null strain may be due to abnormal levels of (p)ppGpp because elevated (p)ppGpp would stimulate transcription from the Pr promoter and hence have a positive stimulatory effect on the system. This may at least in part account for much greater effect observed in the Crc null strain as compared to the effect of the Δ8 Crc target-site deficient mutant. Taken together, the data in Figures 7 and 8 provide strong evidence that Crc controls translation (but not transcription) of DmpR to result in elevated levels of DmpR *in vivo.*

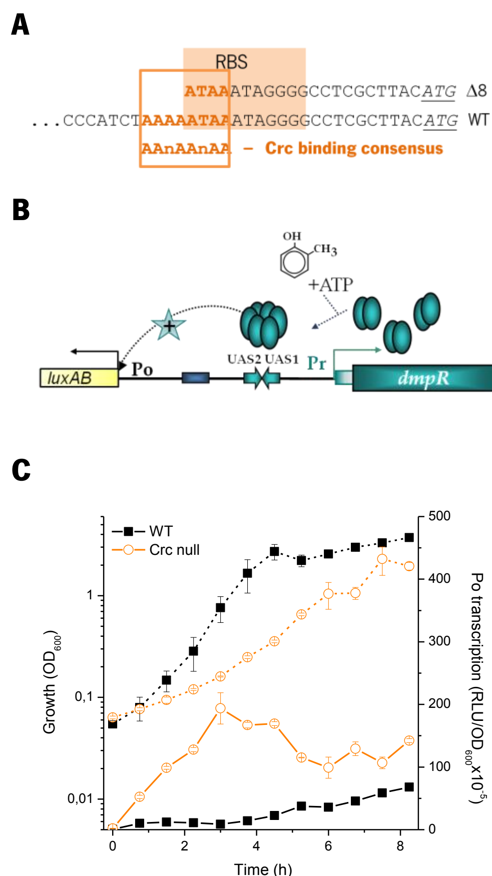


Figure 8 – Effect of Crc on DmpR-dependent transcription *in vivo.* **(A)** Promoter distal sequences of WT and Δ8 Pr-*dmpR* 5'-LRs indicating the location of a potential Crc binding motif and RBS (consensus sequence from (41). **(B)** Schematic of the reporter plasmid (pVI466) used to test the impact of Crc in WT versus Crc null strains. **(C)** The graph shows the growth (dashed lines) and luciferase activity (lines) profiles of the two strains carrying pVI466 over 8.25 h of growth; WT (PP2; black squares) versus Crc null (orange circles).

To further analyse the role of Crc in inhibiting translation of DmpR, two *in vitro* approaches were used. In the first approach, DmpR production was monitored using a plasmid encoding the native Pr-*dmpR* system in a coupled *in vitro* transcription-translation reaction (IVTT) in the presence of different amounts of Crc. The results showed that the levels of DmpR decrease as the levels of Crc increase (Figure 9A upper panel). Notably, the inhibitory action of Crc was abolished when a plasmid carrying the Crc target-site mutation (Δ8) was used (Figure 9A lower panel). Recall that the Δ8 deletion does not affect transcript levels, hence these results strongly suggest that Crc binds to the 5'-LR of the *dmpR* mRNA and thereby inhibits translation.

The ability of Crc to potentially bind the *dmpR* mRNA was directly assessed using an RNA electro-mobility shift assays (EMSA) (performed by Fernando Rojo (Spain)), using a RNA probe that encompassed the potential Crc binding site within the 5'-LR (Figure 9B). The presence of different concentration of Crc resulted in a dose-dependent increase in the amount of RNA-Crc complex formed (see bands labelled C in Figure 9B). No RNA-Crc complex was observed in similar assays using RNA probes that carried C-substitutions within the potential Crc motif (orange in Figure 9B; data not shown). The free RNA probe ran as three different

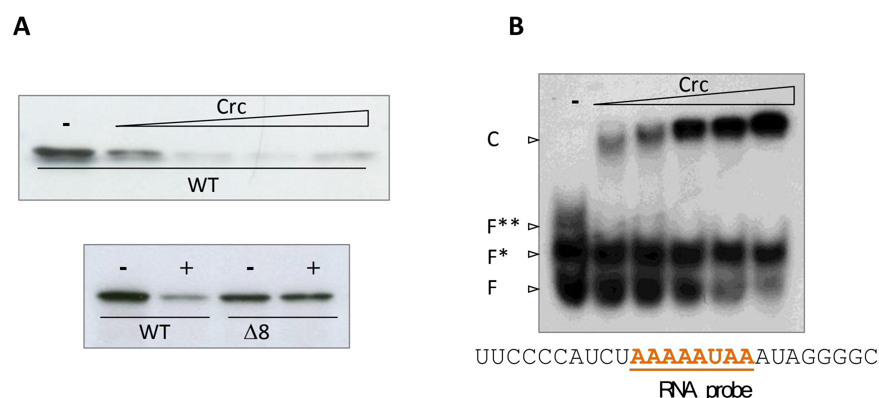
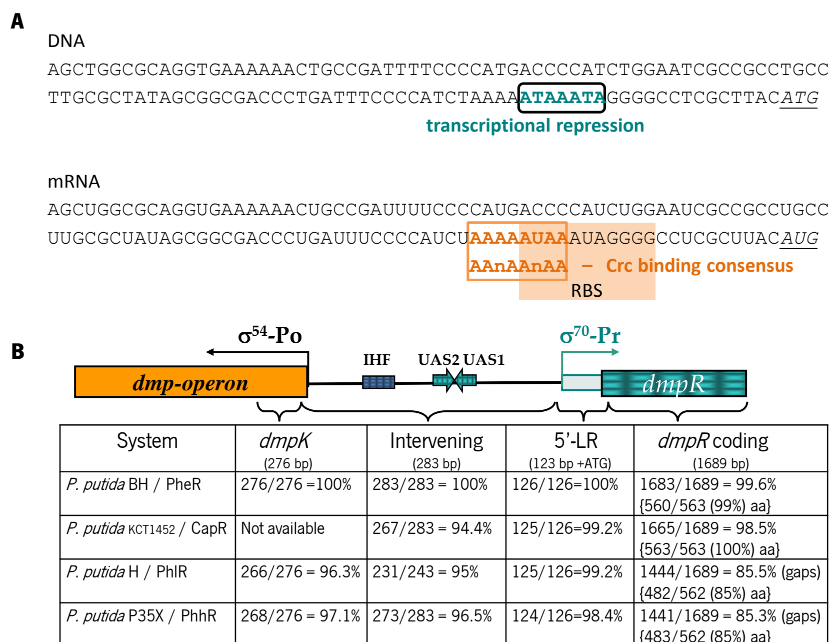


Figure 9 – Effect of Crc on DmpR levels and binding of Crc *in vitro*. (A) Coupled *in vitro* transcription-translation assays. The upper panel shows a representative assay using a Pr-*dmpR* plasmid template with a native 5'-LR (WT, pVI399) in the absence (-) or in the presence of increasing amounts of the Crc protein (0.5, 1.0, 2.0 and 4.0 μM). The lower panel shows assays with the same plasmid (WT, pVI399) and a Crc target site deficient Δ8 derivative (pCon2117) in the absence (-) or presence (+) of 0.8 μM of Crc. **(B)** EMSA assay of a 5'-LR RNA probe (+89 to +115) containing the potential Crc binding site (orange) in the presence of different amounts of Crc (0, 53, 106, 212, 425 and 850 nM); C indicates RNA-Crc complexes, while F, F* and F** indicate the different forms/structures observed for the free RNA probe (Courtesy of Fernando Rojo, Spain).

forms on these non-denaturing gels (denoted F, F* and F** in Figure 9B), although only one form is observed when assessed on denaturing gels (data not shown). It is notable that these three forms are differentially bound by Crc: F and F** can both be bound by Crc, while F* cannot. Because Crc can only bind to sequences that are unpaired (41), it seems likely that F* is an alternative structure where the binding site is masked in a double strand form, while F and F** represent alternative forms where the Crc-site is freely accessible.

4.4. A highly conserved promoter proximal region of 5'-LR is also involved in controlling DmpR levels

The preceding sections document two regulatory processes associated to the 5'-LR of DmpR – one at the levels of transcription and one at the levels of translation. However, the regions involved only cover 11 (~9%) of the 123 bp/nt that comprise this region (Figure 10A). When the sequence of the 5'-LR and DNA in its surroundings – *dmpR*, the intergenic Po-Pr region, and the first gene of the *dmp*-operon (*dmpK*) – are compared to those present in other phenol-degradative systems, the 5'-LR is the most highly conserved region of all (Figure 10B).



System	*dmpK* (276 bp)	Intervening (283 bp)	5'-LR (123 bp +ATG)	*dmpR* coding (1689 bp)
P. putida BH / PheR	276/276 =100%	283/283 = 100%	126/126=100%	1683/1689 = 99.6% {560/563 (99%) aa}
P. putida KCT1452 / CapR	Not available	267/283 = 94.4%	125/126=99.2%	1665/1689 = 98.5% {563/563 (100%) aa}
P. putida H / PhlR	266/276 = 96.3%	231/243 = 95%	125/126=99.2%	1444/1689 = 85.5% (gaps) {482/562 (85%) aa}
P. putida P35X / PhhR	268/276 = 97.1%	273/283 = 96.5%	124/126=98.4%	1441/1689 = 85.3% (gaps) {483/562 (85%) aa}

Figure 10 - The 5'-LR is more highly conserved than coding regions in related systems. (A) Sequence of the 5'-LR with the region involved in transcription regulation (DNA) and translational regulation (RNA) indicated**. (B)** Table of relative conservation in related systems with a schematic of the sub-regions used in Blast searches superimposed above.

This raised the question of whether the other 112 nt of the 5'-LR are involved in additional roles that impact DmpR levels.

Previously, the potential secondary structure of the 5'-LR had been analysed using the M-fold and Vienna RNA folding programmes (66). The result predicted two possible structures: structure 1 (left-hand side, Figure 11A) has a large central stem-loop and two alternative forms of a short stem-loop, while the second structure (right-hand side, Figure 11A) exhibits two long stem-loops. Five different mutations, denoted mutA to mutE, designed to disrupt either one or the other of these potential structures had previously been introduced into Pr-*luxAB* reporter plasmids and found to have little, if any, effect on transcription from the Pr promoter in either the exponential or stationary phase of growth (Figure 11B; (66)). To be able to assess whether these mutations could mediate any effect on translation of DmpR, they were reconstituted in the context of Pr-*dmpR* expression plasmids and introduced into the *P. putida* PP980 Po-*luxAB* reporter strain described under Figure 7. These *in vivo* experiments showed that all five of the mutations resulted in a down-regulation of DmpR in exponential phase as assessed through output from the DmpR-dependent Po-promoter (Figure 11C). The differences were 2- to 5-fold for all but mutB, which represents the mutation with the least effect (Figure 11C, bar chart). Because the major effect of these mutations appeared limited to the exponential phase of growth, the activities of these Pr-*dmpR* plasmids were also monitored in a Crc null background (Figure 11D). Note that promoter output is generally lower in the Crc-null strain for all derivatives – a phenomenon that is also observed for transcription from other promoters dependent on a variety of different σ-factors (V. Shingler, personal communication). Most importantly, however, the profiles differ between mutations as compared to those seen in the *P. putida* wild-type (compare Figure 11C and D); mutA, mutD and mutE presented similar profiles as the wild-type (WT) Pr-*dmpR* construct in the Crc null strain, indicating that the reduced output observed in the wild-type strain is associated with the action of Crc i.e. that these three mutations in some way affect binding of Crc to the mRNA and thereby translation of DmpR. Therefore, these mutations lend support to the idea that the secondary structure of the 5'-LR may be important for correct presentation of the Crc-binding site. For mutB and mutC this was not the case. Surprisingly, mutC exhibited ~2.5-fold higher output as compared to the WT Pr-*luxAB* derivative in the Crc null strain, while mutB exhibited ~3-fold lower output. Although this data does not give any insight into which of the potential secondary structure is

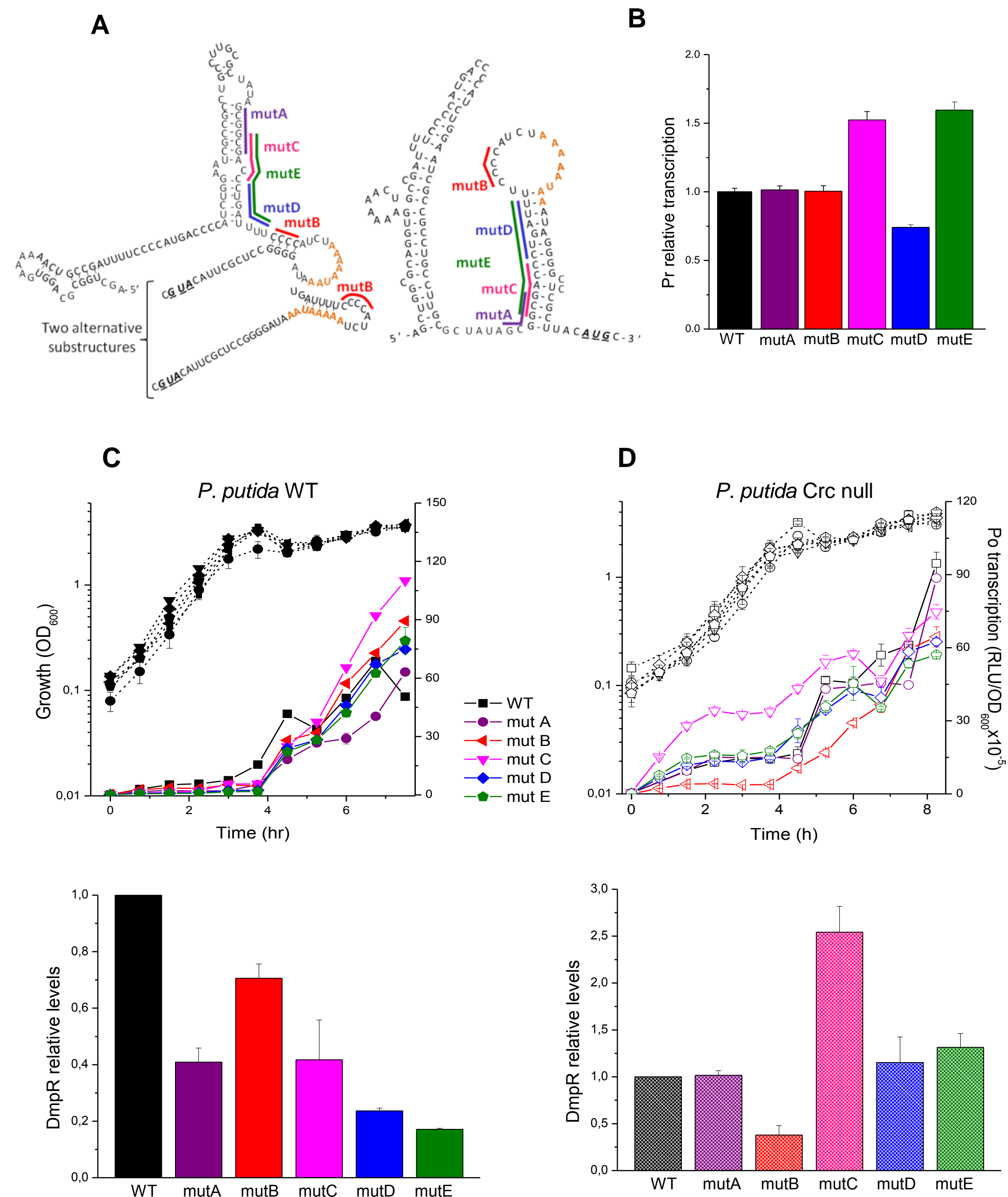


Figure 11 – Effect of 5'-LR structural mutations on transcription and translation. (A) Two potential secondary structures of the 5'-LR of *dmpR* are depicted with the locations of the substitution mutations mutA to mutE indicated in relation to the Crc binding site (orange). **(B)** Relative transcription values of PP2 harbouring Pr-*luxAB* WT (pCon1817) or the otherwise identical derivatives carrying the five substitution mutations mutA to mutE (pCon1821 [mutA], pCon1822 [mutB], pCon1833 [mutC], pCon1841[mutD], or pCon1843 [mutE]: WT pCon1817 values were set as 1. The differences were constant over the entire growth curve (as followed for nine hours). **(C)** DmpR-dependent Po-output of PP980 (Po-*luxAB* reporter strain) harbouring Pr-*dmpR* expression plasmids (pCon2127 [WT], pCon2128 [mutA], pCon2129 [mutB], pCon2132 [mutC], pCon2133 [mutD], or pCon2134 [mutE]) cultured on rich medium contain 2-methylphenol as the inducer of DmpR activity as under Figure 7. Bar graphs are the average differences during the exponential phase (time points 0.45 to 3.5 h) with that of WT set as 1. **(D)** As for (C) except assays were performed with the Crc null derivative of PP980 (PP3044). Exponential phase time points used in the bar diagram were 1.5 to 4.5 h for this slower growing strain.

correct, the results with the mutB and mutC derivatives do provide the first evidence that the structure of the 5'-LR mRNA may be important for other regulatory processes that impact translation of DmpR by a mechanism(s) that is independent of Crc-binding.

To start to address this possibility, two deletions (Δ3 and Δ4) of the 5'-LR that had previously been analysed in the context of the 5'-LR in Pr-*luxAB* transcriptional reporter plasmids (Figure 6) were incorporated in the context of Pr-*dmpR* expression plasmids. Both these deletions would remove most of the potential secondary structure of the 5'-LR, but would leave the Crc binding site and RBS intact (see Figure 12A). These plasmid derivatives carry a longer portion of the Pr upstream region than those used in the analysis shown in Figure 11. Because this has repercussions for the absolute values obtained (53), an equivalent derivative

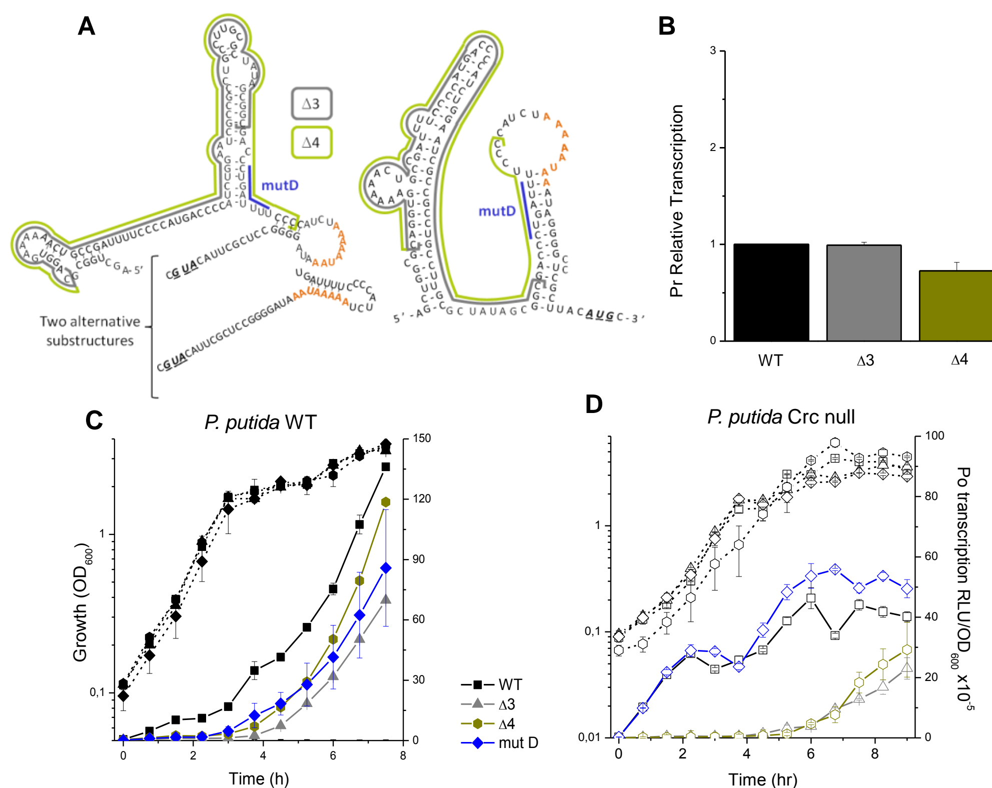


Figure 12 – Effect of Δ3 and Δ4 on transcription and translation. (A) Schematic of the two potential secondary structures showing the extent of the regions deleted in Δ3 (gray) and Δ4 (yellow) as well as the mutated region in mutD (blue) and the location of the Crc binding site (orange). **(B)** Relative transcription values of PP2 harboring Pr-*luxAB* reporter plasmids: WT (pCon1580), Δ3 (pCon1598), Δ4 (pCon1599). Differences were constant over the 9 h growth curve; data are the average of all time point comparisons with WT set as 1. **(C)** Growth curves (dotted lines) and luciferase values (continuous lines) obtain from Po-*luxAB* reporter strain PP908 harboring Pr-*dmpR* plasmids: WT (pVI401), Δ3 (pCon1483), Δ4 (pCon1482) or mutD (pCon2150). **(D)** As (C) but with the same plasmids in PP3044 – the Crc null counterpart of PP980.

harbouring the mutD substitution was also constructed for comparison purposes. These plasmids were subjected to the same series of assays as described for mutA to mutE above. Consistent with the finding in Figures 6, neither Δ3 nor Δ4 had any detectable effect on output from the Pr promoter i.e. they do not have any effect on the level of transcription (Figure 12B). Importantly, however, both deletions caused a decrease in output from the DmpR-dependent Po-*luxAB* fusion in wild-type *P. putida* and markedly decrease output in the isogenic Crc null derivative (Figure 12C and D). Thus, these data again support the suggestion that the secondary structure of the 5'-LR is important for efficient translation of DmpR (and thus corresponding output from DmpR-dependent Po promoter).

4.5. The 5'-LR encompasses an upstream ORF (uORF) and a potential target site for a small regulatory RNA.

As detailed in the introduction, 5'-LRs can control translation by different mechanisms including modulation of translation efficiency through coupling of translation with that of a short upstream open-reading frame (uORF). Control of translation of PsqR of *P. aeruginosa* is a prime example of such regulation (26). In that system, the 5'-LR of the PqsR mRNA contains an uORF, and active translation through the uORF is required for efficient translation of the downstream *pqsR* gene. Moreover, this translational coupling was found to be controlled through the secondary structure of the 5'-LR, which in turn is modulated by interactions with a small regulatory RNA denoted PhrS (26). The region of PhrS responsible for modulating translation has been traced to a sub-region of PhrS denoted creg. A homologue of PhrS exists in *P. putida* and exhibits a 100% identity in the creg region (26). This raised the possibility that a similar mechanism may also operate through the 5'-LR of the DmpR mRNA.

Re-examination of the 5'-LR region in the light of the findings with PhrS/PqsR in *P. aeruginosa* revealed the presence of both a potential uORF within the *dmpR* mRNA and a potential target site for interaction with the regulatory RNA PhrS (Figure 13). The potential PhrS target site within the uORF of 5'-LR of the *dmpR* mRNA has both higher continuity and complementary than that of the identified target site within *pqsR* mRNA (Figure 13, lower). The *dmpR* uORF has two potential start sites; a GTG Val start (20 codon ORF) or an ATG Met start (12 codon ORF) that are bounded by a UAG termination codon. The UAG termination codon is

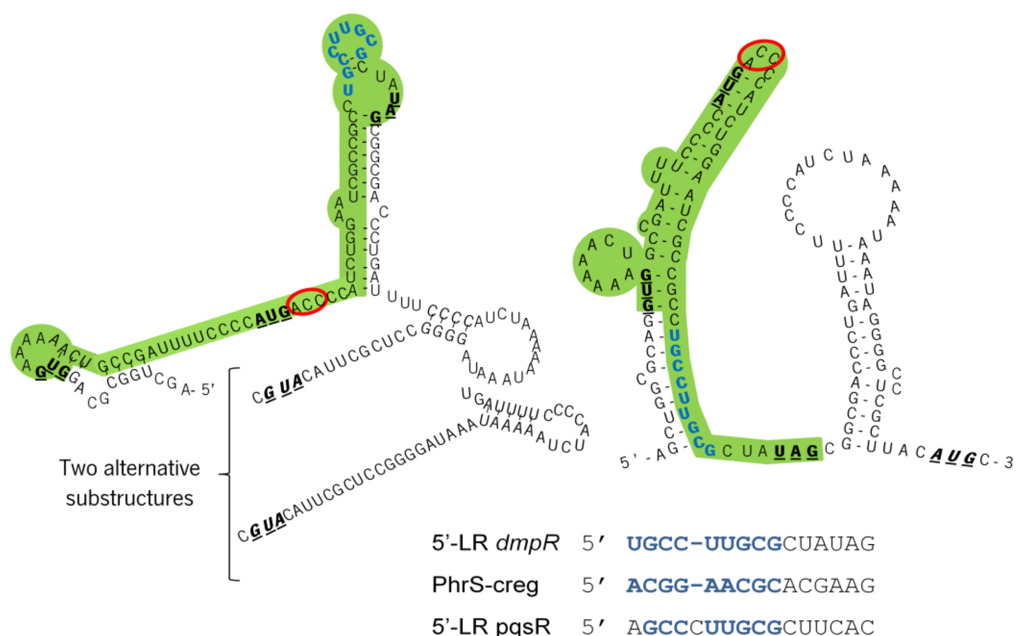


Figure 13 – The 5'-LR has a potential uORF and target site for PhrS. The sequence of the 5'-LR in alternative structures is shown with the potential uORF highlighted in green. Start and stop codons are shown in bold, while the red circle indicates the codon targeted for mutagenesis as described in the text. The region with complementary to the PhrS-creg region is shown in blue. The alignment in the lower part of the figure shows the greater continuity and higher complementarity of the PhrS-creg region with the 5'-LR of *dmpR* than its document target within the mRNA of *pqsR*.

the poorest translation stop signal, particularly when followed by a C (67). Inefficient termination at the stop codon would be consistent with coupling between translation of the uORF and translation of DmpR. These observations prompted an evaluation of these potential regulatory features in controlling DmpR levels.

To assess a potential regulatory role of the PhrS RNA, the Po-*luxAB* Pr-*dmpR* reporter plasmid used in analysis of the *in vivo* effects of Crc (Figure 8D) was employed. This reporter plasmid was introduced into wild-type *P. aeruginosa* and its PhrS null counterpart. However, luciferase activity assays across the growth curve did not show any difference between the two strains and, therefore, did not provide any support for the notion that PhrS is involved in controlling translation of DmpR (Figure 14A).

To assess if translational coupling between the uORF and the *dmpR* ORF could potentially control DmpR levels, the ACC codon immediately downstream of the Met start of the uORF was chosen for mutagenesis. This codon was targeted because substitutions to a strong TAA termination codon would not disrupt any of the bonds in the predicted secondary structures (see red circles in Figure 13). However, the ACC to TAA substitution would be

expected to terminate any potential translation of the uORF and thus disrupt potential translational coupling between the uORF and that of *dmpR*. The ACC to TAA substitution, when incorporated in the 5'-LR in the context of a Pr-*luxAB* transcriptional reporter plasmid, had little or no effect on the transcriptional output from Pr as assessed by luciferase plate assays (Figure 14B, upper). However, plate test of the Pr-*dmpR* expression plasmid, showed that the ACC to TAA substitution resulted in reproducible decrease in output from the DmpR-dependent Po promoter in the *P. putida* Po-*luxAB* reporter strain PP980 (Figure 14B, lower). Although, the differences were not dramatic under the experimental conditions used, they do suggest that translational coupling may exist.

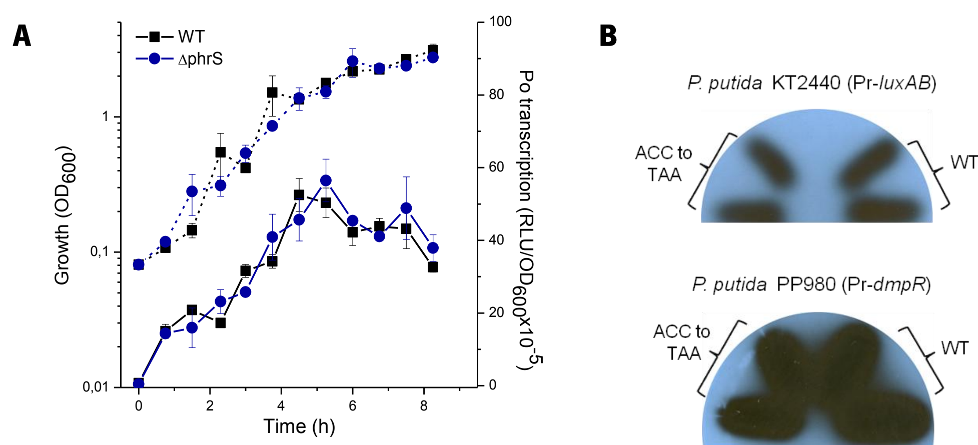


Figure 14 – Assessing potential regulation through PhrS and translational coupling. (A) The graph shows the growth (dashed lines) and luciferase (continuous lines) profiles of WT (PAO1, black squares) and its PhrS null counterpart (PA6671, blue circles) harbouring the Po-*luxAB* Pr-*dmpR* reporter plasmid pVI466. **(B)** Autoradiograms of luciferase plate test assays. The upper image shows a comparison of PP2 harbouring either a wild-type (WT) Pr-*luxAB* reporter (pCon1485) or an equivalent plasmid carrying a ACC to TAA (stop) substitution within the 5'-LR (pCon2204) cultivated on rich media. The lower image shows a comparison of PP980 Po-*luxAB* reporter strains harbouring either a wild-type (WT) Pr-*dmpR* expression plasmid (pVI346) or an equivalent plasmid carrying a ACC to TAA (stop) substitution within the 5'-LR (pCon2203). In this case, strains were cultivated on rich media containing 2-methylphenol required for DmpR activity.

5. Discussion

Expression of natural systems for the catabolism of toxic carbon sources is finely tuned to host physiology to allow appropriate expression of the specialised catabolic enzymes only when they are of benefit to the host. For the *dmp*-system for degradation of methylphenols, all such integration appears to converge on controlling expression of the aromatic-responsive transcriptional activator DmpR. In this work, we present evidence that the DNA encoding the 5'-leader region (5'-LR) of *dmpR* exerts control of transcription from the Pr promoter, while the resulting 5'-LR of the mRNA is employed in at least two independent mechanisms that regulate translation of the DmpR gene product – one involving binding of the Crc protein to repress translation, and a second mechanism that likely involves translational coupling through the 5'-LR to enhance translation of DmpR. The findings raise the following mechanistic questions:

- *How does the initially transcribed (5'-LR) DNA control output from the Pr promoter?*
- *How does Crc-binding to the mRNA result in repression of translation of DmpR?*
- *How might the secondary structure of the 5'-LR potentially influence binding of Crc?*
- *How might translational coupling through an upstream open-reading frame (uORF) enhance translation of DmpR?*

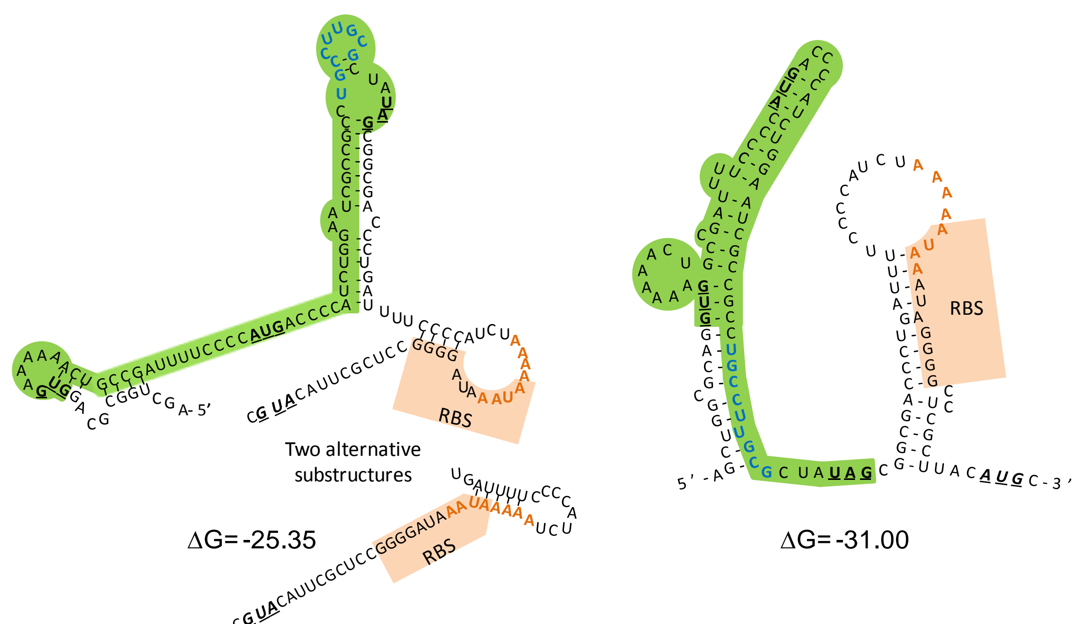


Figure 15 – Potential regulations at translational level addressed during this study. The two potential secondary structures and sub-structures predicted for the *dmpR* 5'-LR presented with the features involved in the potential regulatory mechanisms identified highlighted: Crc binding site in orange letter, RBS in pale orange, uORF in green and the PhrS-creg consensus region in blue letter.

How does the initially transcribed (5'-LR) DNA control output from the Pr promoter?

Simple possession of a symmetrical promoter-distal ATAAATA motif within the DNA encoding the 5'-LR is all that is needed to result in a reduced number of transcripts from the Pr promoter (Figures 5 and 6). How the presence of this motif causes this effect is yet unknown, but does not involve any process that reduces the mRNA half-life (Figure 5). This suggests that the mechanism involves either the number of productive initiation events at the Pr promoter or the rate with which RNAP progresses through the 5'-LR DNA. Because the ATAAATA motif lies quite distal from the Pr promoter, the latter of these two alternatives appears the most likely.

Two possible mechanisms can be envisioned to account for the repressive effect of this motif on the number of transcripts produced from the Pr promoter. The first of these involves transcriptional pausing at the ATAAATA motif due to its similarity to a promoter element (-10 TATAAT) motif and its location 102 bp downstream of the +1 transcription start. As outlined in the introduction, σ-factor can stay partially attached to the RNAP-core during the first ~200 nt of elongation of the RNA and cause pausing by binding to promoter mimics within the DNA. This would impede the progression of RNAP and thus result in a lower net number of full length transcripts.

A second and perhaps more likely explanation is that the ATAAATA motif serves a binding site for an as yet unknown repressor. Repression at the level of the transcription is also seen in the heterologous host *E. coli* DH5 containing the minimal system of the Pr promoter and 5'-LR controlling the expression of the *luxAB* reporter genes (66). Therefore, such a repressor would have to be encoded within the genomes of both *E. coli* and *P. putida* KT2440. The ATAAATA motif bears similarity to the consensus sequence for *E. coli* H-NS (TCGATAAATT). H-NS-like proteins all bind AT-rich regions and are global regulators of environmentally regulated genes in many Gram-negative bacteria (reviewed in (68)). *E. coli* has two known H-NS like proteins (H-NS and StpA), while the genome of *P. putida* KT2440 encodes multiple such proteins (PP0017, PP1366, PP2947, PP3693, and PP3765). However, systematic evaluation of the possible involvement of a H-NS-like protein by monitoring transcription from Pr with or without the 5'-LR in *P. putida* strains individually devoid of each of these H-NS like proteins refuted this idea (T. del Peso-Santos and V. Shingler, unpublished

data). Thus, the identity of the potential repressor that could bind the ATAAATA motif remains a mystery. A future approaches to pursue such a repressor is to try affinity purification of proteins that binds the 5'-LR DNA by established methods (69).

How does Crc-binding to the mRNA result in repression of translation of DmpR?

During this study, much evidence that Crc specifically binds to the 5'-LR of the *dmpR* mRNA and inhibits translation both, *in vivo* (Figures 7 and 8) and *in vitro* (Figure 9) has been acquired. The actual mechanism that brings about this effect remains to be verified. However, as highlighted in Figure 15, the Crc binding site overlaps 5 nt of the potential RBS of *dmpR*. This, therefore, suggests that the most probable mechanism involves a competition between Crc and ribosomes to recognize and bind their respective sites within this region.

This hypothesis could potentially be directly assessed by toe-printing assays in the presence or absence of Crc. The toe-printing technique was originally developed to detect formation of RNA-ribosomal complexes through blocking a primer-extension reaction mediated by a reverse transcriptase. As reverse transcriptase copies cDNA from RNA (using an oligonucleotide primer complementary to a region downstream the RBS), bound ribosomes serve as a road-block. The length of the prematurely terminated primer extension product can then be used to determine the RBS location. By analogy, the presence of Crc or other RNA-binding proteins could also potentially be detected in the same way.

During the time of this study it was possible to establish fluorescent-primer extension based assay essentially as described in (53). However, consistent with previous findings with Crc, premature termination due to binding of Crc could not be detected by toe-printing (41). This result suggests that Crc is readily displaced by the reverse transcriptase as it copies the RNA into cDNA. Although the direct binding of Crc bind could not be determined by this method, if competition with ribosomes is indeed the mechanism through which Crc acts, its inclusion in assays would likely reduce the number of premature termination events observed due to ribosome binding. However, despite extensive test I was unable to document binding of available 30S *E. coli* ribosomes to the 5'-LR mRNA. Therefore, further development and optimisation of this assay is need in order to test the hypothesis.

How might the secondary structure of the 5'-LR potentially influence binding of Crc?

Crc can only bind to single stranded RNA (41). The analysis of the effect of structural mutants (mutA-mutE) on DmpR translation did not give any insight to which of the two potential secondary structures of the 5'-LR mRNA is the likely relevant form *in vivo* (Figure 11). However, the results not only suggested that the secondary structure does influence Crc binding (as exemplified by reversal to wild-type expression profiles in a Crc null strain), but also revealed the existence of an additional Crc-independent regulatory role(s) associated with the 5'-LR. To further pursue how the secondary structure may influence DmpR levels, it would be greatly beneficial to know the actual structure of 5'-LR mRNA. One technique that may allow the determination of this is RNA enzymatic probing as used in (26). Within this kind of analysis, the restriction pattern obtain with RNAses that only recognize and cut after specific nucleotides and/or only cut single-stranded or double-stranded RNA can be used to deduce secondary structures. Such analysis would directly determine if the proposed Crc-binding site is located within a loop, the size of such a loop, and if mutants that apparently facilitate Crc binding alter the overall structure.

How might translational coupling through an upstream open-reading frame (uORF) enhance translation of DmpR?

As outlined in the preceding sections, Crc binding likely competes with binding of ribosomes to directly reduce translational efficiency, and binding of Crc is also likely to be enhanced by exposure of the Crc binding site in a single stranded loop configuration. Introduction of a TAA stop codon immediately adjacent to the ATG Met codon within a small uORF of the 5'-LR apparently reduces translation of DmpR (Figure 14C). These finding suggest that there is some translational coupling event between the uORF and that of DmpR. Two non-mutually exclusive scenarios can be envisaged.

Firstly, because transcription and translation are coupled processes in bacteria, active translation of the uORF would likely block the potential formation of the first large stem loop of either of the structure depicted in Figure 15. This, in turn, may simply allow formation of an RNA configuration that is bound less efficiently by Crc – less efficient binding of Crc would result in enhance translation of *dmpR*. A second, possible explanation is suggested by the

inefficient termination codon context of the uORF – a TAG codon followed by a C. Inefficient termination at this codon might allow read-through of terminated-but-not-released-ribosome that would scan by lateral diffusion along the mRNA to the next available ATG start (70) i.e. that of the *dmpR* ORF. In so doing, translational efficiency may be increased by increasing the frequency of productive binding of ribosomes directly and/or by "railroading" Crc off the mRNA in much the same manner as reverse transcriptase appears to do. One approach to distinguish between these two possibilities would be to introduce an alternative Met start and stop codon downstream of the uORF termination codon but upstream of the Crc site, so as to sequester any scanning ribosomes before they could reach the Crc binding site or the ATG initiation codon of *dmpR*.

Irrespective of the mechanism involved, either explanation for the apparent translational coupling demands that the uORF is actually translated. Future experiments involving translational fusions with a reporter gene, in- and out-of-frame, and downstream of the termination codon placed within the uORF, should resolve this issue. However, translation coupling to enhance translation of DmpR combined with translational repression via Crc presents a conundrum – what is the regulatory logic of having two counteractive systems operating simultaneously? Because Crc levels and availability are controlled in response to the carbon sources and growth conditions of the bacteria, repression of translation by Crc seemingly makes biologic sense – muting of the *dmp*-system when preferred carbon sources are present. But why also have a system for translational coupling that can stimulate translation of DmpR? By analogy to the PsqR/PhsR system of *P. aeruginosa* (26), it is plausible that a small regulatory RNA (or an RNA-binding protein) may influence this event to enhance DmpR production when the methylphenol catabolic enzymes are required. An initial experiment using a PhrS null mutant of *P. aeruginosa* did not provide any supportive evidence that its *P. putida* homologue may be such factor (Figure 14A). However, given the higher similarity of a potential target site for PhrS within the uORF of the 5'-LR, further analysis of this possibility using the native *P. putida* host is certainly warranted.

6. Conclusions

A complete understanding of the physiological control of catabolic gene expression is necessary to allow predictable successful design of bacteria for biotechnological and bioremediation applications. In the case of the *dmp*-system of *Pseudomonas putida* CF600, expression of DmpR is without doubt the major regulatory check point of the system and the target for integration within host physiology. Much previous work had focussed on the Pr promoter that drives transcription of *dmpR*, but little was known about the impact of the 5'-LR of the *dmpR* gene. Therefore, this study was performed in the pursuit of more information about the role of the 123 bp long *dmpR* 5'-LR.

The main conclusions that can be drawn from the results are:

1) A short promoter distal motif – ATAAATA – within the DNA encoding the 5'-LR of the *dmpR* mRNA is responsible for a 5- to 7-fold repression at the level of transcription from Pr.
2) Crc specifically binds to a near-consensus RNA site (AAAAAUAA), localized in the distal region of the 5'-LR, and represses translation of *dmpR*.
3) The secondary structure of the 5'-LR is likely involved in the regulation of DmpR levels by i) providing a configuration suitable for Crc binding, and ii) also by other mechanisms that are not completely understood as yet.
4) The 5'-LR possesses a small open-reading frame located upstream of *dmpR*, translation of which seems to have a positive effect on DmpR levels and, therefore, on the *dmp*-system.

References

1. Ochman, H. and Moran, N.A. (2001) Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. *Science*, **292**, 1096-1099.
2. Andersson, J.O. and Andersson, S.G. (2001) Pseudogenes, junk DNA, and the dynamics of Rickettsia genomes. *Mol Biol Evol*, **18**, 829-839.
3. Tamas, I., Klasson, L.M., Sandstrom, J.P. and Andersson, S.G. (2001) Mutualists and parasites: how to paint yourself into a (metabolic) corner. *FEBS Lett*, **498**, 135-139.
4. Cases, I., de Lorenzo, V. and Ouzounis, C.A. (2003) Transcription regulation and environmental adaptation in bacteria. *Trends Microbiol*, **11**, 248-253.
5. Gruber, T.M. and Gross, C.A. (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol*, **57**, 441-466.
6. Shingler, V. (2003) Integrated regulation in response to aromatic compounds: from signal sensing to attractive behaviour. *Environ Microbiol*, **5**, 1226-1241.
7. Shingler, V. (2009) In Timmis, K. N. (ed.), *Handbook of Hydrocarbon and Lipid Microbiology*. Springer-Verlag, Berlin Heidelberg, pp. 1235-1245.
8. Dua, M., Singh, A., Sethunathan, N. and Johri, A.K. (2002) Biotechnology and bioremediation: successes and limitations. *Appl Microbiol Biotechnol*, **59**, 143-152.
9. Guell, M., Yus, E., Lluch-Senar, M. and Serrano, L. (2011) Bacterial transcriptomics: what is beyond the RNA horiz-ome? *Nat Rev Microbiol*, **9**, 658-669.
10. Osterberg, S., Del Peso-Santos, T. and Shingler, V. (2011) Regulation of alternative σ-factor use. *Annu Rev Microbiol*, **65**, 37-55.
11. Gourse, R.L., Ross, W. and Rutherford, S.T. (2006) General pathway for turning on promoters transcribed by RNA polymerases containing alternative σ factors. *J Bacteriol*, **188**, 4589-4591.
12. Paget, M.S. and Helmann, J.D. (2003) The sigma70 family of sigma factors. *Genome Biol*, **4**, 203.

13. Haugen, S.P., Ross, W. and Gourse, R.L. (2008) Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat Rev Microbiol*, **6**, 507-519.

14. Leibman, M. and Hochschild, A. (2007) A σ-core interaction of the RNA polymerase holoenzyme that enhances promoter escape. *EMBO J*, **26**, 1579-1590.

15. Buck, M., Gallegos, M.T., Studholme, D.J., Guo, Y. and Gralla, J.D. (2000) The bacterial enhancer-dependent σ^{54} (σ^{N}) transcription factor. *J Bacteriol*, **182**, 4129-4136.

16. Studholme, D.J. and Buck, M. (2000) The biology of enhancer-dependent transcriptional regulation in bacteria: insights from genome sequences. *FEMS Microbiol Lett*, **186**, 1-9.

17. Shingler, V. (2011) Signal sensory systems that impact σ^{54}-dependent transcription. *FEMS Microbiol Rev*, **35**, 425-440.

18. Sze, C.C., Laurie, A.D. and Shingler, V. (2001) *In vivo* and *in vitro* effects of integration host factor at the DmpR-regulated σ^{54}-dependent Po promoter. *J Bacteriol*, **183**, 2842-2851.

19. Bernardo, L.M., Johansson, L.U., Solera, D., Skarfstad, E. and Shingler, V. (2006) The guanosine tetraphosphate (ppGpp) alarmone, DksA and promoter affinity for RNA polymerase in regulation of σ^{54}-dependent transcription. *Mol Microbiol*, **60**, 749-764.

20. Bernardo, L.M., Johansson, L.U., Skarfstad, E. and Shingler, V. (2009) σ^{54}-promoter discrimination and regulation by ppGpp and DksA. *J Biol Chem*, **284**, 828-838.

21. Srivatsan, A. and Wang, J.D. (2008) Control of bacterial transcription, translation and replication by (p)ppGpp. *Curr Opin Microbiol*, **11**, 100-105.

22. Blankschien, M.D., Lee, J.H., Grace, E.D., Lennon, C.W., Halliday, J.A., Ross, W., Gourse, R.L. and Herman, C. (2009) Super DksAs: substitutions in DksA enhancing its effects on transcription initiation. *EMBO J*, **28**, 1720-1731.

23. Paul, B.J., Barker, M.M., Ross, W., Schneider, D.A., Webb, C., Foster, J.W. and Gourse, R.L. (2004) DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell*, **118**, 311-322.

24. Yanofsky, C. (1981) Attenuation in the control of expression of bacterial operons. *Nature*, **289**, 751-758.

25. Park, S.Y., Cromie, M.J., Lee, E.J. and Groisman, E.A. (2010) A bacterial mRNA leader that employs different mechanisms to sense disparate intracellular signals. *Cell*, **142**, 737-748.

26. Sonnleitner, E., Gonzalez, N., Sorger-Domenigg, T., Heeb, S., Richter, A.S., Backofen, R., Williams, P., Huttenhofer, A., Haas, D. and Blasi, U. (2011) The small RNA PhrS stimulates synthesis of the *Pseudomonas aeruginosa* quinolone signal. *Mol Microbiol*, **80**, 868-885.

27. Liu, M.Y. and Romeo, T. (1997) The global regulator CsrA of *Escherichia coli* is a specific mRNA-binding protein. *J Bacteriol*, **179**, 4639-4642.

28. Moreno, R., Ruiz-Manzano, A., Yuste, L. and Rojo, F. (2007) The *Pseudomonas putida* Crc global regulator is an RNA binding protein that inhibits translation of the AlkS transcriptional regulator. *Mol Microbiol*, **64**, 665-675.

29. Garst, A.D., Edwards, A.L. and Batey, R.T. (2011) Riboswitches: structures and mechanisms. *Cold Spring Harb Perspect Biol*, **3**.

30. Majdalani, N., Vanderpool, C.K. and Gottesman, S. (2005) Bacterial small RNA regulators. *Crit Rev Biochem Mol Biol*, **40**, 93-113.

31. Palleroni, N.M., ERB. (2004) In Ramos JL, e. (ed.), *Pseudomonas*. Kluwer Academic/Plenum Publishers, New York, Vol. Vol. 1, pp. 3–44.

32. Marques, S. and Ramos, J.L. (1993) Transcriptional control of the *Pseudomonas putida* TOL plasmid catabolic pathways. *Mol Microbiol*, **9**, 923-929.

33. Bartilson, M., Nordlund, I. and Shingler, V. (1990) Location and organization of the dimethylphenol catabolic genes of *Pseudomonas* CF600. *Mol Gen Genet*, **220**, 294-300.

34. Jeffrey, A.M., Yeh, H.J., Jerina, D.M., Patel, T.R., Davey, J.F. and Gibson, D.T. (1975) Initial reactions in the oxidation of naphthalene by *Pseudomonas putida*. *Biochemistry*, **14**, 575-584.

35. Gomes, N.C., Kosheleva, I.A., Abraham, W.R. and Smalla, K. (2005) Effects of the inoculant strain *Pseudomonas putida* KT2442 (pNF142) and of naphthalene contamination on the soil bacterial community. *FEMS Microbiol Ecol*, **54**, 21-33.

36. Sarand, I., Skarfstad, E., Forsman, M., Romantschuk, M. and Shingler, V. (2001) Role of the DmpR-mediated regulatory circuit in bacterial biodegradation properties in methylphenol-amended soils. *Appl Environ Microbiol*, **67**, 162-171.

37. Collier, D.N., Hager, P.W. and Phibbs, P.V., Jr. (1996) Catabolite repression control in the *Pseudomonads*. *Res Microbiol*, **147**, 551-561.

38. Rojo, F. (2010) Carbon catabolite repression in *Pseudomonas*: optimizing metabolic versatility and interactions with the environment. *FEMS Microbiol Rev*, **34**, 658-684.

39. Moreno, R., Martinez-Gomariz, M., Yuste, L., Gil, C. and Rojo, F. (2009) The *Pseudomonas putida* Crc global regulator controls the hierarchical assimilation of amino acids in a complete medium: evidence from proteomic and genomic analyses. *Proteomics*, **9**, 2910-2928.

40. O'Toole, G.A., Gibbs, K.A., Hager, P.W., Phibbs, P.V., Jr. and Kolter, R. (2000) The global carbon metabolism regulator Crc is a component of a signal transduction pathway required for biofilm development by *Pseudomonas aeruginosa*. *J Bacteriol*, **182**, 425-431.

41. Moreno, R., Marzi, S., Romby, P. and Rojo, F. (2009) The Crc global regulator binds to an unpaired A-rich motif at the *Pseudomonas putida alkS* mRNA coding sequence and inhibits translation initiation. *Nucleic Acids Res*, **37**, 7678-7690.

42. Sonnleitner, E., Abdou, L. and Haas, D. (2009) Small RNA as global regulator of carbon catabolite repression in *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A*, **106**, 21866-21871.

43. Ruiz-Manzano, A., Yuste, L. and Rojo, F. (2005) Levels and activity of the *Pseudomonas putida* global regulatory protein Crc vary according to growth conditions. *J Bacteriol*, **187**, 3678-3686.

44. Aranda-Olmedo, I., Ramos, J.L. and Marques, S. (2005) Integration of signals through Crc and PtsN in catabolite repression of *Pseudomonas putida* TOL plasmid pWW0. *Appl Environ Microbiol*, **71**, 4191-4198.

45. Yuste, L., Hervas, A.B., Canosa, I., Tobes, R., Jimenez, J.I., Nogales, J., Perez-Perez, M.M., Santero, E., Diaz, E., Ramos, J.L. *et al.* (2006) Growth phase-dependent expression of the *Pseudomonas putida* KT2440 transcriptional machinery analysed with a genome-wide DNA microarray. *Environ Microbiol*, **8**, 165-177.

46. Shingler, V., Bartilson, M. and Moore, T. (1993) Cloning and nucleotide sequence of the gene encoding the positive regulator (DmpR) of the phenol catabolic pathway encoded by pVI150 and identification of DmpR as a member of the NtrC family of transcriptional activators. *J Bacteriol*, **175**, 1596-1604.

47. Shingler, V. (2004) In Ramos, J.-L. (ed.), *Pseudomonas: Virulence and Gene Regulation*. Kluwer Academic / Plenum Publishers, New York, Vol. II, pp. 451-477.

48. Shingler, V. and Moore, T. (1994) Sensing of aromatic compounds by the DmpR transcriptional activator of phenol-catabolizing *Pseudomonas sp.* strain CF600. *J Bacteriol*, **176**, 1555-1560.

49. Shingler, V. and Pavel, H. (1995) Direct regulation of the ATPase activity of the transcriptional activator DmpR by aromatic compounds. *Mol Microbiol*, **17**, 505-513.

50. O'Neill, E., Ng, L.C., Sze, C.C. and Shingler, V. (1998) Aromatic ligand binding and intramolecular signalling of the phenol-responsive σ^{54}-dependent regulator DmpR. *Mol Microbiol*, **28**, 131-141.

51. O'Neill, E., Sze, C.C. and Shingler, V. (1999) Novel effector control through modulation of a preexisting binding site of the aromatic-responsive σ^{54}-dependent regulator DmpR. *J Biol Chem*, **274**, 32425-32432.

52. Wikstrom, P., O'Neill, E., Ng, L.C. and Shingler, V. (2001) The regulatory N-terminal region of the aromatic-responsive transcriptional activator DmpR constrains nucleotide-triggered multimerisation. *J Mol Biol*, **314**, 971-984.

53. Johansson, L.U., Solera, D., Bernardo, L.M., Moscoso, J.A. and Shingler, V. (2008) σ^{54}-RNA polymerase controls σ^{70}-dependent transcription from a non-overlapping divergent promoter. *Mol Microbiol*, **70**, 709-723.

54. Del Peso-Santos, T., Bernardo, L.M., Skarfstad, E., Holmfeldt, L., Togneri, P. and Shingler, V. (2011) A hyper-mutant of the unusual σ^{70}-Pr promoter bypasses synergistic ppGpp/DksA co-stimulation. *Nucleic Acids Res*, **39**, 5853-5865.

55. Cases, I. and de Lorenzo, V. (2005) Genetically modified organisms for the environment: stories of success and failure and what we have learned from them. *Int Microbiol*, **8**, 213-222.

56. Miller, J.H. (1972) *Experiments in molecular genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

57. Hanahan, D. (1985) *DNA cloning: A practical approach, vol. I*. IRL Press Ltd, Oxford.

58. de Lorenzo, V. and Timmis, K.N. (1994) Analysis and construction of stable phenotypes in gram-negative bacteria with Tn5- and Tn10-derived minitransposons. *Methods Enzymol*, **235**, 386-405.

59. Franklin, F.C., Bagdasarian, M., Bagdasarian, M.M. and Timmis, K.N. (1981) Molecular and functional analysis of the TOL plasmid pWWO from *Pseudomonas putida* and cloning of genes for the entire regulated aromatic ring *meta* cleavage pathway. *Proc Natl Acad Sci U S A*, **78**, 7458-7462.

60. Pavel, H., Forsman, M. and Shingler, V. (1994) An aromatic effector specificity mutant of the transcriptional regulator DmpR overcomes the growth constraints of *Pseudomonas sp.* strain CF600 on *para*-substituted methylphenols. *J Bacteriol*, **176**, 7550-7557.

61. Holloway, B.W., Krishnapillai, V. and Morgan, A.F. (1979) Chromosomal genetics of *Pseudomonas*. *Microbiol Rev*, **43**, 73-102.

62. Furste, J.P., Pansegrau, W., Frank, R., Blocker, H., Scholz, P., Bagdasarian, M. and Lanka, E. (1986) Molecular cloning of the plasmid RP4 primase region in a multi-host-range tacP expression vector. *Gene*, **48**, 119-131.

63. Sze, C.C., Moore, T. and Shingler, V. (1996) Growth phase-dependent transcription of the σ^{54}-dependent Po promoter controlling the *Pseudomonas*-derived (methyl)phenol *dmp* operon of pVI150. *J Bacteriol*, **178**, 3727-3735.

64. Sze, C.C. and Shingler, V. (1999) The alarmone (p)ppGpp mediates physiological-responsive control at the σ^{54}-dependent Po promoter. *Mol Microbiol*, **31**, 1217-1228.

65. Skalman, L.N. (2010) Thesis project, Umeå University, Umeå.

66. Moscoso, J.A. (2009) Thesis project, Umeå University, Umeå.

67. Tate, W.P., Poole, E.S., Dalphin, M.E., Major, L.L., Crawford, D.J. and Mannering, S.A. (1996) The translational stop signal: codon with a context, or extended factor recognition element? *Biochimie*, **78**, 945-952.

68. Fang, F.C. and Rimsky, S. (2008) New insights into transcriptional regulation by H-NS. *Curr Opin Microbiol*, **11**, 113-120.

69. Rescalli, E., Saini, S., Bartocci, C., Rychlewski, L., De Lorenzo, V. and Bertoni, G. (2004) Novel physiological modulation of the Pu promoter of TOL plasmid: negative regulatory role of the TurA protein of *Pseudomonas putida* in the response to suboptimal growth temperatures. *J Biol Chem*, **279**, 7777-7784.

70. Adhin, M.R. and van Duin, J. (1990) Scanning model for translational reinitiation in eubacteria. *J Mol Biol*, **213**, 811-818.