# Studies on Actinomycete Plasmid and Bacteriophage DNA

Youtaro Shibayama


A thesis submitted to the Faculty of Science,
University of the Witwatersrand, Johannesburg,
in fulfilment of the requirements for the degree of Doctor of Philosophy

May 2011

# DECLARATION

I declare that this thesis is my own unaided work. It is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other university.

May 9, 2011

Youtaro Shibayama

Date

# ABSTRACT

The actinomycetes are a diverse group of Gram-positive, high G+C content bacteria of the order *Actinomycetales*. Many species of this group are of human interest as a result of their pathogenic nature and diverse metabolic properties. Those from the genus *Nocardia* cause opportunistic infections of lung, brain and central nervous system, and cutaneous tissue. They are also producers of antibiotics and industrially important enzymes. As studies describing plasmids in this genus are limited, we have characterized a 4326 bp cryptic plasmid pYS1 from *Nocardia aobensis* IFM 10795. Three open reading frames (ORFs) were predicted. Both sequence analyses and detection of single-stranded intermediates suggested a rolling-circle mechanism as the mode of replication of pYS1. Mutageneses and deletion analyses revealed both the predicted double- and single-stranded origins to be indispensable in replication and thus lack of secondary signals for both leading and lagging strand synthesis. The replicon of pYS1 is broad-host-range and compatible to that of pAL5000 of mycobacteria, making it potentially useful in genetic manipulation of various actinomycetes. Insertion analyses showed *orf1*, despite its sequence similarity to plasmid transfer genes, is involved in plasmid stability rather than conjugation and is lethal in the absence of a functional *orf3*. This situation is somewhat analogous to the *kil*/*kor* system of pIJ101 of *Streptomyces*, except that *orf3* was unrelated to *korA* and was shown by promoter-probe assays to encode a novel transcriptional repressor negatively regulating *orf1* expression.

*Mycobacterium tuberculosis*, the causative agent of tuberculosis, is responsible for 8 to 10 million new cases of disease and 3 million deaths every year, many of which in HIV/AIDS patients. To identify novel drug targets in multi-drug resistant strains, DNA from *Rhodococcus* phages were tested and shown to be inhibitory to mycobacteria. To study the modes of action of products encodes by these genes, the lambda cI857 repressor and $P_R$ promoter system was tested for its ability to function as a tight genetic switch for toxic gene expression. Induction studies in both *E. coli* and mycobacteria indicate that strong mycobacterial promoters may be necessary to drive expression of toxic genes and/or the repressor gene.

# ACKNOWLEDGEMENTS

I would like to acknowledge Prof. Eric R. Dabbs, for his supervision on this project. A majority of the experiments were conducted in his laboratory at the School of Molecular and Cell Biology, University of the Witwatersrand, Johannesburg, South Africa.

To Prof. Yuzuru Mikami, for our collaboration on the *Nocardia* plasmid work and the phage unusual bases work. Screening for plasmids and HPLC on phage nucleic acid were conducted in his laboratory at the Center for Pathogenic Fungi and Microbial Toxicose (currently the Medical Mycology Research Center), Chiba University, Chiba, Japan.

To Dr. Katsukiyo Yazawa, for providing me with cultures of *Nocardia* strains at the Medical Mycology Research Center.

To Dr. Yoshihisa Sei at the Kagawa School of Pharmaceutical Sciences, Tokushima Bunri University, Tokushima, Japan, for conducting mass spectrometry on my phage nucleic acid samples.

To Mr. Kazunori Aoyama at the Medical Mycology Research Center, for confirming my data on phage KG3 nucleic acid by re-running its nucleosides on HPLC.

To Dr. Kurt Lightfoot at the School of Molecular and Cell Biology, for providing me with plasmid p106.

This study was funded by the Japan-South Africa Bilateral Agreement on Science and Technology, the National Research Foundation of South Africa, and the Medical Research Council of South Africa.

Data presented here have been used as the basis for manuscripts submitted for publication.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# 1. INTRODUCTION

## 1.1 The actinomycetes

The actinomycetes are a group of Gram-positive bacteria belonging to the order *Actinomycetales*. With 30 families (Fig. 1.1), this is one of the largest taxonomic groups within the eubacteria (Ventura *et al*., 2007).

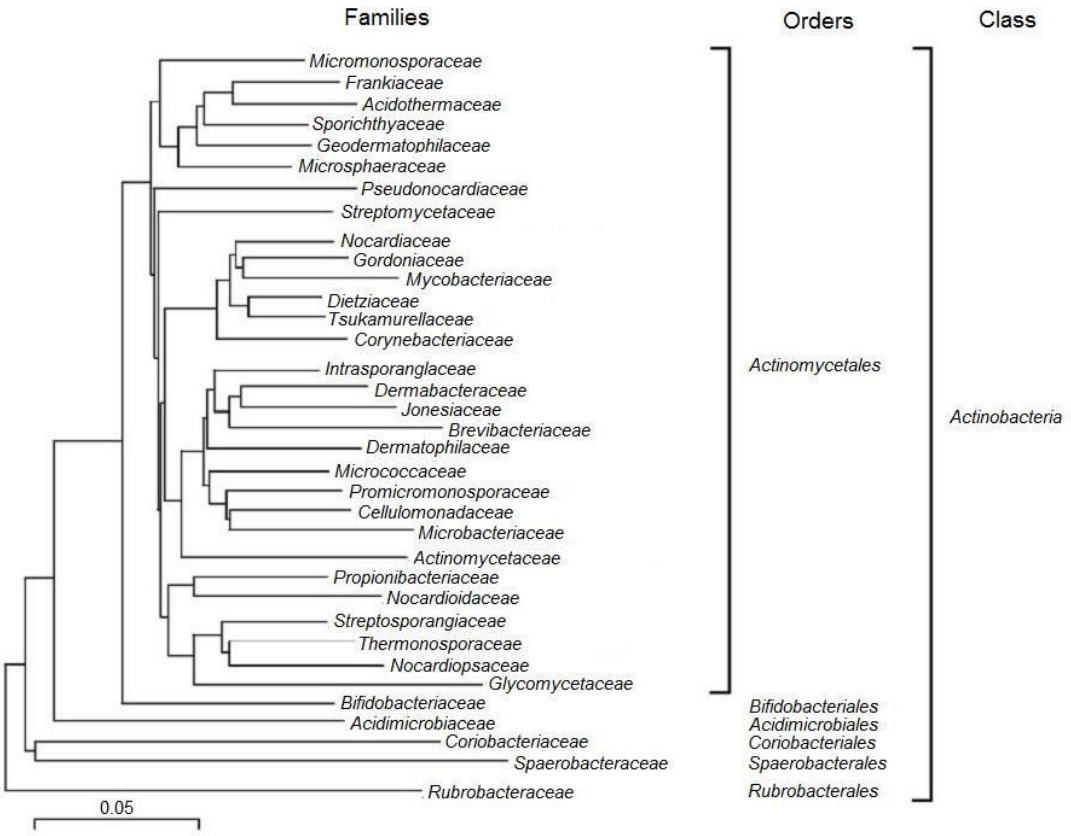



Figure 1.1. Tree showing the phylogenetics of *Actinobacteria*. Distances are based on 1500 nucleotides of 16S rRNA. Bar = 5 nucleotides. Modified from Ventura *et al*. (2007).

Actinomycetes are noted for the high G+C content of their DNA, ranging from 54% in some *Corynebacterium* (Cerdeno-Tarraga *et al*., 2003) to >70% in *Streptomyces* (Bentley *et al*., 2002; Ikeda *et al*., 2003) and *Frankia* (Normand *et al*., 2007). A variety of morphologies are displayed by actinomycete cells, including coccoid (e.g. *Micrococcus*), rod-coccoid (e.g. *Arthrobacter*), fragmenting hyphae (*Nocardia*), to extensively branched

mycelia (e.g. *Streptomyces*) (Holt *et al*., 2000). In terms of lifestyle, these bacteria range from pathogens (e.g. *Mycobacterium* spp., *Nocardia* spp., *Rhodococcus* spp., *Corynebacterium* spp., *Propionibacterium* spp.) to soil-dwellers (e.g. *Streptomyces*) to plant commensals (*Leifsonia* spp.) to nitrogen-fixing-symbionts (*Frankia*) (Holt *et al*., 2000). This diversity is reflected in the genome sequences of these bacteria (Ventura *et al*., 2007), of which 96 have so far been deposited into the public database (http://www.ncbi.nlm.nih.gov/genome).

The actinomycetes have been much studied by microbiologists and the biotechnology industry as a result of their diverse physiological and metabolic properties. They have been a rich source of secondary metabolites with pharmaceutical, chemical, agricultural, and industrial applications. For example, ~70% of antibiotics have been derived from products of actinomycetes, most notably the *Streptomyces* (Takahashi and Ōmura, 2003). The actinomycetes also secrete enzymes which degrade various xenobiotics, including polycyclic and halogenated aromatics, most of which are by-products of chemical manufacturing (Solyanikova *et al*., 2008). In this respect, the *Nocardioform* actinomycetes, including the genera *Gordonia*, *Nocardia*, *Rhodococcus*, *Tsukamurella* to mention a few, have shown promise in their biodegradative potential (Solyanikova *et al*., 2008). As there is an ongoing need to discover natural products with novel activities, significant efforts are spent into screening environmental samples for rare actinomycetes as well as those commonly known for previously overlooked compounds (Gathogo et al., 2004; Kurtböke, 2010; Takahashi and Ōmura, 2003). In support of these efforts, predictive modelling suggests that >150,000 bioactive metabolites are yet to be discovered from *Streptomyces* alone (Watve *et al*., 2001).

Contrary to the economic and medical benefits brought about by the actinomycetes, a large fraction of morbidity and mortality caused by bacterial infections are due to the same group of organisms. *Mycobacterium tuberculosis*, the causative agent of tuberculosis (TB), is responsible for 8 to 10 million new cases of disease and 3 million deaths every year, many of which in HIV/AIDS patients (World Health Organization [WHO], 2006). A major reason for its success as a pathogen is its ability to adapt to a wide range of conditions inside the human host, including the ability to stay dormant in a latent phase, the molecular mechanisms of which are not clearly understood (Ducati *et al*.,

2006). The genus *Mycobacterium* includes other pathogenic species of major significance, such as *M. leprae*, the causative agent of leprosy.

The HIV/AIDS pandemic has also caused an increase in infections by other pathogenic actinomycetes, such as species from the genera *Nocardia*, *Rhodococcus*, *Gordonia*, and *Corynebacterium*. They typically cause a granulomatous inflammatory reaction, which may progress to abscess formation (McNeil and Brown, 1994), although symptoms may vary at each of potential infection sites including lungs, brain, bones, pericardium, skin, among others (Salinas-Carmona, 2000). Combination antibiotic therapy is generally used to combat these infections, although multiple resistance to all major classes of drugs by some of these organisms has been reported (Cox *et al*., 2010; Glupczynski *et al*., 2006; Otsuka *et al*., 2006; Corti and Fioti, 2003; Hsueh *et al*., 1998; McNeil and Brown, 1994; Yazawa *et al*., 1993; Nordmann *et al*., 1992).

### 1.2 The *Nocardia*

The genus *Nocardia* currently contains 92 species characterized by molecular and phenotypic methods (Euzéby, 2010), although species have attracted widely different levels of attention. Although *Nocardia* are ubiquitous in soil, >50 species have been identified as human or animal pathogens (Ambrosioni *et al*., 2010) causing nocardiosis, which is a rare opportunistic disease mainly affecting patients with deficiencies in cell-mediated immunity such as acquired immunodeficiency syndrome (AIDS) patients or organ transplant patients. Incidence of disease in these groups of patients is 140-340 times higher than in the immunocompetent population (Filice, 2005). The disease is manifested in three forms – pulmonary, cutaneous, and disseminated nocardiosis. Pulmonary nocardiosis is the most common form among immunosuppressed patients, while the cutaneous form is more frequent among immunocompetent patients (Ambrosioni *et al*., 2010). Disseminated nocardiosis frequently involves spreading of infection from the lungs to the central nervous system, brain, skin, and soft tissues (Ambrosioni *et al*., 2010). High rate of mortality is associated with the disease, reported to be between 7 and 44% for disseminated nocardiosis (McNeil and Brown, 1994). Treatment usually involves combination antibiotic therapy with sulfonamides and other

drugs and lasts for at least six months in the case of pulmonary nocardiosis (Corti and Villafañe-Fioti, 2003).

The only genome sequence currently available for *Nocardia* is that of *N. farcinica* IFM 10152 (Ishikawa *et al*., 2004). The genome was reported to be most similar to those of *M. tuberculosis* (Cole *et al*., 1998) and *Corynebacterium glutamicum* (Kalinowski *et al*., 2003), and many genes with potential involvement in virulence were identified, although further work is needed to understand the mechanisms of virulence of this organism. The sequence also revealed at least 103 oxygenases, implying diverse metabolic capability of the organism (Ishikawa *et al*., 2004).

### 1.3 Plasmids

Plasmids are extrachromosomal elements capable of autonomous replication. Plasmids influence the biology of the host, sometimes dramatically; however, previous studies suggest no correlation between occurrence of plasmids in *Nocardia* and their virulence (Provost *et al*., 1996; Qasem *et al*., 1999). Studies on plasmids may provide understandings of basic biological processes (Thomas, 2000). Plasmids also display great diversity in size, mode of replication and transfer, host range, and the set of genes they carry (Thomas, 2000), making them interesting elements for analysis. Information of such nature is limited in *Nocardia*. Plasmids are major vehicles of horizontal gene flow, contributing to bacterial evolution. Among these, integrative and conjugative elements are increasingly recognized to allow rapid genetic adaptation to environmental alterations or occupation of entirely new niches (Burrus and Waldor, 2004). As tools, plasmids are indispensable in molecular biology, commonly used for cloning, expression, and mutagenesis.

#### 1.3.1 Occurrence of plasmids in *Nocardia*

A few previous studies have investigated the association between occurrence of plasmids in *Nocardia* and their virulence or drug resistance. Provost *et al*. (1996) found plasmids in 27 out of 87 clinical isolates and could not associate virulence or drug resistance to them. These plasmids ranged in size from <8 kb to >50 kb, and multiple plasmids could be isolated from some strains (Provost *et al*., 1996). Although a bias was found in the

incidence of plasmid-bearing strains with respect to the species, area of isolation of the strains, and region of infection of the human body (Provost *et al*., 1996), the underlying reasons are not understood. Qasem *et al*. (1999) found plasmids in 8 out of 40 *Nocardia asteroides* strains, ranging in size from <5 kb to >50 kb, and observed that their occurrence is not specific to particular species, source of origin, virulent strains, or drug resistance. This is in contrast to the closely related *Rhodococcus equi*, a pathogen of foals and immunocompromised humans, whose virulence is associated with the possession of *vapA* and *vapB* plasmids (von Bargen and Haas, 2009).

### 1.3.2 Sequenced plasmids from *Nocardia*

So far full sequences of only five plasmids from *Nocardia* have been deposited in the public database (www.ncbi.nlm.nih.gov). These include pNF1 and pNF2 (Ishikawa *et al*., 2004), pXT107 (Xia *et al*., 2006), pC1 (Shen *et al*., 2006), and pYS1 (this work). For closely related actinomycetes such as *Rhodococcus*, *Corynebacterium*, and *Streptomyces*, full sequences of 22, 34, and 28 plasmids, respectively, are available in the database (www.ncbi.nlm.nih.gov), highlighting the comparative rarity of plasmid sequences from *Nocardia*.

#### 1.3.2.1 pNF1 and pNF2

Sequenced as parts of the genome of *N. farcinica* IFM 10152 (Ishikawa *et al*., 2004), pNF1 and pNF2 are both circular plasmids of 184,027 bp and 87,093 bp, respectively. pNF1 was predicted to contain 160 protein-coding genes, 53 of which could be assigned potential functions, and was estimated to have a copy number of approximately one per chromosome. pNF2 was predicted to contain 90 protein-coding genes, 23 of which could be assigned potential functions, and was estimated to have a copy number of approximately two per chromosome. Sequences similar to *parA* and *parB* genes were identified on both plasmids, which are probably responsible for plasmid partitioning during cell division (Ishikawa *et al*., 2004).

In a different study, using codon adaptation index (Sharp and Li, 1987) as an indicator of gene expressivity and ribosomal protein genes as a reference, highly expressed genes were predicted in the *N. farcinica* genome (Wu *et al*., 2006). Genes on pNF1 and pNF2

were not as “expressive” as those on the chromosome, and only 1 out of predicted 571 highly expressed genes was on the plasmid. No function has been assigned to this gene (Wu *et al*., 2006).

#### 1.3.2.2 pXT107

Plasmid pXT107 is a circular cryptic plasmid of 4335 bp from *Nocardia* sp. 107 (Xia *et al*., 2006). Of seven predicted open reading frames (ORFs), one had resemblance to those encoding rolling-circle-replication initiation proteins. Consistent with this mode of replication, the double- and single-stranded origins of replication (DSO and SSO) were predicted. The replicon of this plasmid was not functional in *Streptomyces* (Xia *et al*., 2006).

#### 1.3.2.3 pC1

Plasmid pC1 is a circular cryptic plasmid of 5841 bp from *Nocardia* sp. C-14-1 (Shen *et al*., 2006). Of five predicted ORFs, one resembled those encoding a DNA recombinase and another to those encoding Θ-type replication proteins. The replicon of this plasmid was functional in *Nocardia* and *Rhodococcus* but not in *Streptomyces* or *Amycolatopsis*. The replication gene of pC1 was not sufficient to allow stable inheritance of a recombinant plasmid, and its copy number was estimated to be approximately ten per chromosome (Shen *et al*., 2006).

### 1.3.3 Other plasmids from *Nocardia*

There are several plasmids from *Nocardia* that have been described but not sequenced. These include pHG33 (Sensfuss *et al*., 1986), pHG201 and pHG205 (Kalkus *et al*., 1990), and pNI100 (Liu *et al*., 2000). Plasmids pHG33 is a circular plasmid of approximately 110 kb from *Nocardia opaca* MR22 and carries genetic information for thallium resistance, which could be transferred to other strains of *Nocardia* by conjugation at a frequency of 1e-1 to 1e-2 per donor (Sensfuss *et al*., 1986). Plasmid pHG201 and pHG205 are linear megaplasmids of 270 kb and 280 kb, respectively, from two different strains of *N. opaca*. Both plasmids carry genetic information for soluble hydrogenase and ribulose-biphosphate which could be transferred to other stains by conjugation (Kalkus *et*

*al*., 1990). Plasmid pNI100 is a cryptic circular plasmid of 19 kb from *Nocardia italica* CCRC 12359 which contains a gene predicted to encode a protein that initiates rolling-circle-replication. This conjugative plasmid has a replicon which is functional in *Streptomyces* (Liu *et al*., 2000).

### 1.4 Aim of part one – Plasmid from *Nocardia*

More plasmids from *Nocardia* should be described to reveal their sizes, modes of replication and transfer, host ranges, and functions of genes they carry. The aim of this work is to characterize a plasmid from *Nocardia* and highlight these features.

### 1.5 Bacterial resistance to antibiotics

Soon after any antibacterial agent is introduced for clinical use, resistance arises. There are four general resistance mechanisms: drug inactivation, target alteration, prevention of drug influx, and drug efflux (Putman *et al*., 2000). Scientists have responded to resistance by modifying already existing drugs or with new classes of drugs. For example, penicillinase-stable penicillins, and second generation aminoglycosides were produced to counteract resistant organisms. However, resistance has emerged to these newer agents in all cases, thereby repeating the cycle (Bush, 2004). Usage of a combination of existing drugs is not a complete solution, as evidenced by the emergence of multidrug-resistant pathogens (Livermore, 2004). In order to keep ahead of resistance, development of new antimicrobial agents is vital.

#### 1.5.1 Emergence of multi-drug-resistant strains of *M. tuberculosis*

A recent joint report by the US Centers for Disease Control and Prevention (CDC) and the WHO (CDC, 2006) has highlighted the emergence of extensively drug resistant (XDR) TB, caused by organisms resistant to at least isoniazid and rifampicin among the first-line drugs (i.e. multiple drug resistant [MDR] TB) and to at least three of six main classes of second-line drugs. XDR-TB has been reported in at least 17 countries in all continents (CDC, 2006). In a hospital in Kwazulu-Natal Province of South Africa, there has been a deadly outbreak of XDR-TB which killed 52 out of 53 patients, with a median survival period of only 16 days since the first sputum specimen collection (Ghandi *et al*.,

2006). It is clear that there is an urgent necessity to identify additional molecular targets and develop novel antimicrobial agents.

### 1.6 Antibacterial drug discovery

Despite the emergence of multidrug-resistant bacteria, there has been a serious fall-off in development of novel antibacterials in the past several decades. The oxazolidinone linezolid and cyclic lipopeptide daptomycin have been the only structurally novel antibiotics to be introduced in over thirty years (Hamad, 2010). Most antibiotics introduced since 1968 have been modifications of already existing ones (Powers, 2004). One reason for this decline is difficulty in target identification (Projan and Shlaes, 2004). Traditionally, antibiotics were developed by screening natural sources for molecules with antimicrobial effects. Most recently, searches for antibacterials have utilized bioinformatics to identify essential and conserved ORFs in bacterial genomes which may be potential targets (Thomson *et al*., 2004). Currently, complete genome sequences of 1312 bacterial strains are publicly available (http://www.ncbi.nlm.nih.gov/genome), including those of many clinically important pathogens. However, after more than 15 years since the beginning of the genomics era, not a single antimicrobial agent has been developed by this approach (Thomson *et al*., 2004). Other novel methods currently employed for the discovery of new targets are transcriptional profiling and proteomic analysis (Bandow *et al*., 2003). Besides searching the bacterial genomes, various potential antimicrobial agents are currently under investigation. These include antimicrobial peptides, bacteriophages, and probiotic bacteria (Gillor *et al*., 2005).

#### 1.6.1 Eukaryotic antimicrobial peptides

Gene-encoded, ribosomally synthesized antimicrobial peptides (AMPs) are widely distributed in nature and produced by both eukaryotes and prokaryotes. AMPs are produced in phagocytes and mucosal epithelial cells of mammals, haemolymph of insects, and skin of amphibians, where they provide an early defence against invading microorganisms (Boman, 1995). Some of well-characterized animal AMPs are defensins of humans, cecropins of insects, and magainins of frogs. Animal AMPs display a broad spectrum of activity against both Gram-positive and -negative bacteria. Although various

AMP structures exit, those for which the mode of action has been elucidated act by membrane permeabilization (Papagianni, 2003). Examples of eukaryotic AMPs are shown in Table 1.1.

Table 1.1. Examples of eukaryotic AMPs grouped according to structure

| **Structure** | **Peptide** | **Source** |
|---|---|---|
| High content of certain amino acid | Abaecin | Insects |
| | Bac-5 | Sheep |
| | Prophenin | Pigs |
| Intramolecular disulfide bridges | α-Defensins | Humans |
| | Insect defensins | Insects |
| | Plant defensins | Plants |
| Amphiphilic α-helical structure | Magainins | Frogs |
| | Bombinins | Frogs |
| | Cecropin | Insects |

(Modified from Papagianni, 2003)

### 1.6.2 Prokaryotic antimicrobial peptides

Prokaryotic AMPs, known as bacteriocins, are a large and diverse group of toxins found in both bacteria and archaea. They are narrow spectrum antimicrobials and use a variety of killing mechanisms. Lactic acid bacteria produce the best-known Gram-positive bacteriocins, the lantibiotics, which include nisin, a commercially used food preservative (Asaduzzaman and Sonomoto, 2009). Bacteriocins produced by Gram-positive bacteria act by membrane depolarization, membrane leakage, and cell wall synthesis inhibition (Pag and Sahl, 2002). Well known bacteriocins from Gram-negative bacteria include colicins, microcins, both of which are produced by *E. coli*, and pyocins produced by *Pseudomonas* spp. Modes of action by colicins include depolarization of cell membrane, degradation of DNA and RNA, and inhibition of cell wall and protein synthesis (Braun *et al*., 1994). Microcins target cell membranes and biosynthesis of DNA, RNA, and proteins (Destoumieux-Garzón *et al*., 2002; Pons *et al*., 2002). Pyocins act by forming pores on cell membranes and degrading DNA and RNA (Duport *et al*., 1995; Gillor *et al*., 2005; Parret and De Mot, 2000). Different groups of bacteriocins and their targets are summarized in Table 1.2.

Table 1.2. Bacteriocin types and their targets

| **Bacteriocin** | **Class/type** | **Target** |
|---|---|---|
| Lactic acid bacteria bacteriocins | IA, II | cell membrane |
| | IB | cell wall synthesis: lipid II |
| | III | Unknown |
| Colicins | A, B, E1, Ia, Ib, K, N | cell membrane |
| | E2, E7, E8, E9 | DNA |
| | E3, E4, E6 | protein synthesis: 16S rRNA |
| | E5, D | protein synthesis: tRNA |
| | M | cell wall synthesis: lipid carrier |
| Microcins | Modified | DNA replication: β subunit, DNA gyrase (*gyrB*) |
| | | Transcription: β' subunit, RNA polymerase |
| | | Protein synthesis |
| | Non-modified | cell membrane |
| Pyocins | R, F | cell membrane |
| | S | DNA |
| | | cell membrane |
| | | protein synthesis: tRNA |

### 1.6.3 Bacteriophage-mediated antimicrobials

Bacteriophages hold potential for the development of novel antimicrobials. This may be possibly done in three different ways: 1) phage therapy, 2) phage lytic enzymes, and 3) target identification.

### 1.6.3.1 Phage therapy

The idea of using bacteriophages, the most abundant organisms on the planet, as antimicrobials is not new. After d'Hérelle coined the term "bacteriophage" in 1916, he successfully treated a patient with severe dysentery using a phage preparation taken orally and set up phage therapy trials across the globe (Stone, 2002). Phage therapy was extensively tested, succeeding against diseases such as dysentery, typhoid, paratyphoid, cholera, pyogenic and urinary-tract infections (Thacker, 2003). However, overall results were variable; due to ineffective trials and the advent of antibiotics, the use of bacteriophages as antimicrobials was abandoned. In hindsight, the ineffective trials were caused by prescription of wrong phages or those which were no longer viable (Pirisi, 2000).

Although abandoned in the west, phages continued to be used as antibacterial agents in the former Soviet Union and Eastern Europe. For example, during a civil war in the early 1990s, Georgian soldiers carried spray cans with a phage suspension effective against *Staphylococcus aureus*, *Escherichia coli*, *Pseudomonas aeruginosa*, *Streptococcus pyogenes*, and *Proteus vulgaris* (Stone, 2002). In some Russian villages, civilians relied exclusively on phages as the cheaper option compared to antibiotics (Pirisi, 2000). Although some data have been published, most do not exist in English.

Due to the emergence of multidrug-resistant pathogenic bacteria, there has been a recent revival in interest in phage therapy. One advance in phage therapy is against vancomycin-resistant enterococci, for which clinical trials are currently underway (Bradbury, 2004).

### 1.6.3.2 Phage lytic enzymes

Instead of phage therapy, there has been a recent interest in purification and characterization of phage lytic enzymes, which can be used to combat target bacteria. Phage enzymes directed against Gram-positive pathogens, including *Streptococcus pneumoniae* and *Bacillus anthracis*, have been isolated and shown to successfully destroy these pathogens both *in vitro* and *in vivo* (Loeffler *et al*., 2001; Schuch *et al*.,

2002). For example, lytic enzyme Pal from *S. pneumoniae* phage Dp-1 has been shown to kill fifteen common serotypes of the bacterium, including highly penicillin-resistant strains (Loeffler *et al*., 2001). Lysin PlyG from *B. anthracis* phage γ is host specific, binding to cell wall antigens not recognized by current antibiotics (Schuch *et al*., 2002).

#### 1.6.3.3 Phage-mediated target identification

Another approach to phage-mediated antimicrobial development is target identification. This has been exemplified by a study by Liu *et al*. (2004), who screened the genomes of *S. aureus* phages for ORFs which were inhibitory to the host bacterium. Targets of these antibacterial ORFs were then biochemically identified, followed by a high-throughput screen to discover small molecules which could inhibit the host cell by the same mechanism as the phage proteins. Using this approach, 31 novel polypeptide families were identified from 26 *S. aureus* phages, several of which were inhibitory to DNA replication and transcription (Liu *et al*., 2004). This approach may be a powerful method to identify effective targets in the host, as phages are believed to have evolved multiple strategies to disable host cells.

Although there is potential in the traditional phage therapy or in the use of their lytic enzymes, the development of novel antibiotics by recruiting phages in the identification of new targets may be the best phage-mediated antimicrobial. Small antimicrobial molecules that are commonly used have been selected to have a sufficient degree of distribution in human tissues and appropriate levels of other pharmacokinetic properties. Therefore, novel antibiotics are expected to better meet these criteria of an effective medicine compared to a phage particle or even a lytic protein, whose complexity is much higher and size much larger (Projan, 2004).

## 1.7 Phage diversity

A major factor contributing to the potential of phage-mediated target identification is the abundance of these viruses in nature. For instance, >50 different phages capable of infecting *E. coli* have been included in the Universal Virus Database

(http://ncbi.nlm.nih.gov/ICTVdb/Ictv/index.htm). In addition, it has been calculated that there are ~10 million bacterial species (Curtis *et al*., 2002). If each of these is a host for ≥ 10 different phages, then the number of phage species may be at least 100 million.

Despite the predicted abundance of phages, the use of phage ORFs in identifying novel targets will be effective only if phages employ diverse mechanisms in inhibiting their hosts. If all phages possess similar genes and use identical mechanisms to act against their hosts, there are limited prospects of discovering new bactericidal mechanisms.

Recent comparative genomic analyses of phages infecting *Mycobacterium smegmatis* (Hatfull *et al*., 2010; Hatfull *et al*., 2006; Pedulla *et al*., 2003), *P. aeruginosa* (Kwan *et al*., 2006), and *S. aureus* (Kwan *et al*., 2005) imply enormous diversity in their genetic makeup. Less than 10% of predicted protein families from 60 sequenced mycobacteriophage genomes could be assigned functions based on a database search (Hatfull *et al*., 2010). From 18 *P. aeruginosa* phage genomes and 27 *S. aureus* phage genomes, 35% and 45%, respectively, of predicted ORFs had sequence similarity to those previously reported (Kwan *et al*., 2006; Kwan *et al*., 2005). Moreover, many of these phage genomes display little relationship to each other (Hatfull *et al*., 2010; Kwan *et al*., 2005; Kwan *et al*., 2006), suggesting that each individual phage possesses unique genes. Further support for these findings comes from a shotgun sequencing of uncultured viral communities in sea water, which have shown that ~75% of phage sequences are novel (Breitbart *et al*., 2002).

Average number of ORFs from the sequenced *M. smegmatis* phages*, P. aeruginosa* phages, and *S. aureus* phages, are 114, 105, and 80, respectively (Hatfull *et al*., 2010; Kwan *et al*., 2006; Kwan *et al*., 2005). If a conservative assumption is made that all 100 million phage species in the world possess 50 ORFs, of which 50% are unknown, then there are 2.5 billion phage ORFs to be discovered (Rohwer, 2003). It is likely that only a small fraction of these ORFs participate in killing the host; however, phages present a potentially enormous pool of genes from which a large number of novel bactericidal mechanisms may be identified.

### 1.8 Phage lethal genes

Expression of phage genes may be early, middle, or late, depending on the time at which their transcription begins after initiation of infection. Many lethal genes are expressed early, as they take part in shutoff of host macromolecular biosynthesis (Miller *et al*., 2003). Knowledge of mechanism by which a phage functions in a host is limited. Similarly, an overwhelming majority of genes believed to be lethal are uncharacterized. Studies (Miller *et al*., 2003) also suggest that a single phage may possess several such genes. A small number of them have been studied; when tested individually, a majority of these bactericidal genes have been proven to be dispensable during infection (Abedon *et al*., 2001; Kutter *et al*., 1994a; Kutter *et al*., 1994b), indicating that a single phage has multiple strategies for killing a host. Therefore, with respect to novel antimicrobial target identification, not only does the potential of phages attribute to the number of species and their genetic heterogeneity, but also to the variety of targets attacked by individual phage species. It should also be noted that killing of hosts may not be the only function of many lethal genes; they may have other benefits for phage replication, while being deleterious to the host. Some of the studied lethal phage genes are summarized below.

#### 1.8.1 Phage T4

Studies on *E. coli* phage T4 began in the 1940s, but of its ~300 ORFs, 127 still have no assigned function (Miller *et al*., 2003). Many of these unknown genes are unclonable, reinforcing the notion that they are lethal to the *E. coli* host, and a high level of their expression occurs early in the infection cycle (Miller *et al*., 2003).

One example of a T4 lethal gene is *alc*, which uniquely recognizes the rapidly elongating form of RNA polymerase complex, leading to termination of transcription (Kutter *et al*., 1994). Another example is the Alt protein, which is packaged in the capsid and transported with the phage DNA into the cell upon infection. Alt functions by ADP-ribosylating the α subunit of the host RNA polymerase (Igarashi *et al*., 1991).

T4 possesses at least two other ADP-ribosyltransferase genes, *modA* and *modB*, which are known to be lethal to the host. ModA ADP-ribosylates both α subunits of the host RNA polymerase (Tiemann *et al*., 1999), whereas ModB acts on the translation machinery of the ribosome by ADP-ribosylating the S1 protein, elongation factor EF-Tu, and the chaperone "trigger factor" (Tiemann *et al*., 1999).

Other T4 gene products inhibiting translation are Lit and Gol proteins. Lit cleaves the host EF-Tu at a region which is central to the Mg-GTP-binding domain (Yu and Snyder, 1994). This process is aided by Gol, which stabilizes the EF-Tu-GDP open complex, making it accessible for cleavage (Bingham *et al*., 2000). Gol is also a part of a T4 head capsid protein, suggesting its dual function (Bingham *et al*., 2000).

Intron- or intein-encoded DNases with highly specific recognition sequences are called homing endonucleases, which help to disseminate the intron/intein into specific regions of the genome. T4 possesses >15 such genes (Chevalier and Stoddard, 2001). Although purpose served by the introns/inteins is not clear, these endonucleases have been shown to be bactericidal (Sharma *et al*., 1992).

### 1.8.2 Phage T7

T7 is an *E. coli* phage which has a number of genes implicated in host transcription shutoff. Gp2 binds to promoters of host DNA, thereby abolishing promoter recognition by the host RNA polymerase (Nechaev and Severinov, 1999). Gp0.7 is a serine/threonine protein kinase, which phosphorylates various biosynthetic components: among the transcription machinery includes the β' subunit of RNA polymerase (Zilling *et al*., 1975); RNase III, which is responsible for mRNA processing (Mayer and Schweiger, 1983); RNase E, which is involved in mRNA decay (Marchand *et al*., 2001); among the translation machinery includes IF1, IF2, IF3, elongation factor G, and ribosomal proteins S1 and S6 (Robertson *et al*., 1994). Phosphorylation of the β' subunit of RNA polymerase is lethal (Severinova and Severinov, 2006). Whether host viability is affected by the phosphorylation of other targets is yet to be discovered.

### 1.8.3 Phage lambda (λ)

Lambda is an extensively studied *E. coli* phage, although most works have focused on non-lethal genes. One of few identified bactericidal gene products is peptidyl-tRNA hydrolase, which cleaves peptidyl-tRNA and aminoacyl-tRNA (Garcia-Villegas *et al*., 1991). In addition, lambda *p* gene causes host death by inhibiting the binding of template DNA and ATP to the DnaA protein (initiator protein), thereby preventing DNA replication to begin (Datta *et al*., 2005). CII gene product is another inhibitor of host DNA replication. Although the exact mechanism is unknown, it seems to interfere with the association of DNA helicase with origin of replication (Kędzierska *et al*., 2003). Another lethal gene product, Kil protein, inhibits cell division by interacting with components of the cell envelope (Sergueev *et al*., 2001).

### 1.8.4 Phage SPO1

Phage SPO1 of *Bacillus subtilis* possesses a cluster of early genes in a region named "host-takeover module", including ~24 of mostly unknown function but believed to be involved in inhibition of host biosynthesis (Stewart *et al*., 1998). An example of such gene is *e3*, whose expression leads to inhibition of DNA, RNA, and protein synthesis in both *B. subtilis* and *E. coli* (Wei and Stewart, 1993). Products of genes *44*, *50*, and *51* cause inhibition of host transcription, possibly by binding to RNA polymerase (Sampath and Stewart, 2004).

### 1.8.5 Phages with ssDNA or ssRNA genomes

Where phages with dsDNA genomes encode two to five proteins to carry out host lysis, phages with small, single-stranded nucleic acids usually have only a single gene required for rupturing host cells (Bernhardt *et al*., 2001). Instead of attacking pre-existing peptidoglycan, these phages interfere with cell wall synthesis, thereby causing weakness in the wall, which collapses from osmotic pressure from within. For example, ssDNA phage ϕX174 encodes a single lytic enzyme, E, which blocks MraY, a bacterial membrane protein responsible for transfer of murein precursors to lipid carriers that transport it through the cell membrane (Bernhardt *et al*., 2000). In

the ssRNA genome of phage Qβ is the $a_2$ gene, encoding a protein inhibiting MurA, another essential enzyme for cell wall synthesis (Bernhardt *et al*., 2001). Protein $A_2$ has multiple functions including adsorption to the host sex pilus and protection against ribonucleases (Bernhardt *et al*., 2001).

### 1.8.6 Inhibitory genes from *Rhodococcus* phages

In a previous study (Shibayama, 2006), four novel *Rhodococcus* phages (FND1, WTS1, KZA1, and PGS1) were isolated from soil and characterized. Libraries of their nucleic acid were constructed and screened for clones inhibitory to *Rhodococcus erythropolis*. Nine clones were characterized, and minimum necessary DNA for inhibitory activity sequenced. Of 18 ORFs predicted on these DNA, 13 could not be assigned a function. Genes similar to ones in databases apparently interfered with DNA metabolism, protein synthesis, or integrity of plasma membrane. Sequence analyses of inhibitory clones are summarized in Table 1.3.

Table 1.3. Summary of sequence analyses of phage DNA inhibitory to *Rhodococcus erythropolis*

| Phage | Library | Clone # | BLASTx alignment | E value | GenBank Accession# | # ORFs | # ORFs with no match |
|---|---|---|---|---|---|---|---|
| Fairland1 | *Bgl*II | 7 | none | | DQ981382<br>DQ981383 | 2 | 2 |
| | | 8 | thymidylate synthase complementing protein, *Streptomyces coelicolor* | 2e-59 | DQ981384 | 2 | 1 |
| | | 10 | none | | DQ981385 | 2 | 2 |
| | | 19 | none | | DQ981386<br>DQ981387 | 3 | 3 |
| | *Pst*I | 8 front | phage head maturation peptidase, *Mycobacterium* sp. MCS and KMS | 5e-04 | DQ981388 | 1 | 0 |
| | | 8 back | phage capsid protein, *Streptococcus agalactiae* | 4e-04 | DQ981389 | 1 | 1 |
| | | | dihydropteroate synthase, *Nocardioides* sp. JS614 | 0.003 | | | |
| | | 14 | HNH endonuclease, *Lactobacillus plantarum* phage LP65 | 0.006 | DQ981390 | 1 | 0 |
| | | 16 | none | | DQ981391<br>DQ981392 | 3 | 3 |
| Perouges1 | *Hind*III | 1 | head decoration protein, prophage MuMc02, *Roseobacter* sp. MED193 | 8e-07 | DQ981393 | 2 | 1 |
| | | | phage-related tail protein, *Xanthomonas axonopodis* | 4e-05 | | | |
| | | | cell wall surface anchor family protein, *Streptococcus agalactiae* | 5e-04 | | | |
| | | 7 | α subunit, DNA polymerase III, *Aquifex aeolicus* | 2e-24 | DQ981394 | 1 | 0 |
| | | | | | TOTAL | 18 | 13 |

### 1.9 Currently available inducible expression systems in *Mycobacterium*

To study the action of the above inhibitory genes, a system that allows their expression to be tightly switched off is necessary. Inducible expression systems have been described for both *Rhodococcus* and *Mycobacterium*. In *Rhodococcus* pTip vectors (Nakashima and Tamura, 2004), with which expression is induced by thiostrepton, have been described. In *Mycobacterium* pMind vector (Blokpoel *et al*., 2005), pTACT and pHLEGM vectors (Carroll *et al*., 2005), pUV and pME vectors (Ehrt *et al*., 2005), and pMHA vectors (Hernandez-Abanto *et al*., 2006), all with which expression is induced by tetracycline, have been described. Expression levels using some of these vectors can be controlled by the dosage of the inducer. However, leaky expression has been detected with each of the above vectors, which may be problematic for genes encoding potent inhibitors.

### 1.10 Lambda cI repressor

The *E. coli* phage lambda is capable of both lytic and temperate lifestyles. The latter is characterized by the integration of the phage genome into the host chromosome. Maintenance of this otherwise lethal state is achieved by the phage-encoded cI repressor, whose function is to prevent transcription of *cro* and other genes necessary to enter the lytic lifestyle. Their transcription is negatively regulated by the binding of the lambda repressor to the operators $O_R1$ and $O_R2$ which overlap the promoter $P_R$ that is responsible for driving transcription of the *cro* gene (Ptashne, 2004). As leaky expression of *cro* should be lethal to the host, the cI repressor is a good candidate for a component of a tight genetic switch. A temperature sensitive mutant of the repressor, cI857 (Sussman and Jacob, 1962), has been widely used to allow successful inducible expression in *E. coli* and other Gram-negatives, and with limited success in Gram-positives (Valdez-Cruz *et al*., 2010). This repressor is functional at ~33°C but loses its ability to bind the operators at ~42°C, thereby allowing a temperature-sensitive transcriptional control of genes downstream of $P_R$. This repressor may potentially be used to construct a tight genetic switch in *Mycobacterium*.

### 1.11 Aim of part two – Testing of phage genes for inhibition against *Mycobacterium* and construction of a tight genetic switch to control their expression

Multidrug-resistance in *M. tuberculosis* urges for identification of new susceptibility targets. As the preliminary genetic analysis of above phages has suggested the presence of novel bactericidal mechanisms, these inhibitory genes should be tested for activity in *Mycobacterium*. To understand the modes of action of the encoded proteins, an inducible expression system functional in this genus is necessary. The aim is therefore to construct a system where expression of a cloned inhibitory gene can be controlled by the cI857-$P_R$ system constituting a tight genetic switch.

### 1.12 Phage anti-restriction

Bacteria have evolved protective mechanisms against attacking phages at almost all stages of phage life cycle (Krüger and Bickle, 1983). One of these is the restriction and modification system, which attacks foreign DNA inside the cells while protecting their own. This function is important enough to have apparently evolved independently several times, as exhibited by the presence of different classes of restriction modification systems. In order to survive restriction, phages have evolved anti-restriction mechanisms including 1) blocking of restriction enzyme, 2) phage-encoded modification of DNA, 3) stimulation of host modification functions, 4) possession of unusual bases, 5) co-injection of protective proteins, 6) destruction of endonuclease co-factors, and 7) loss of restriction sites (Krüger and Bickle, 1983). Of these mechanisms, ones that are functional *in vitro* upon purification of DNA are the modification of DNA, presence of unusual bases, and absence of restriction sites. These have been mostly studied in *E. coli* phages and *B. subtilis* phages.

Modified bases refer to those whose modifications take place after the synthesis of normal DNA. *B. subtilis* phages SPβ and ϕ3T possess a methyltransferase gene that methylates the central cytosine in the sequence GGCC, which is the recognition sequence of endonuclease *Bsu*RI of *B. subtilis* (Cregg *et al*., 1980). *E. coli* phage Mu produces the Mom protein, which acetimidates ~15% of its adenine residues, thereby producing $N^6$-(1-acetamido)adenine, making its DNA resistant to type I and III enzymes and partially resistant to type II enzymes (Hattman, 1980). *E. coli* phages T2 and T4 encode a methyltransferase that methylates a fraction of adenine residues to produce 6-methylaminopurine (Gold *et al*., 1966). *Shigella sonnei* phage DDVI

methylates the 7 position of guanine in about a quarter of its DNA, producing 7-methylguanine (Nikolskaya *et al*., 1979).

Unusual bases are those which are not adenine, cytosine, guanine, or thymine, and are produced at the level of nucleotide metabolism. Various unusual bases in phages have been discovered, summarized in Table 1.4. Their structures are illustrated in Figure 1.2.

Table 1.4. Unusual bases in phage DNA

| **Phage** | **Host** | **Base change** | **% of normal base replaced** | **Reference** |
|---|---|---|---|---|
| T4 | *E. coli* | 5-hydroxymethylcytosine (glucosylated) replacing cytosine | 100% | Lehman and Pratt, 1960; Wyatt and Cohen, 1953 |
| SPO1 | *B. subtilis* | 5-hydroxymethyluracil replacing thymine | 100% | Kallen *et al*., 1962 |
| PBS1 | *B. subtilis* | Uracil replacing thymine | 100% | Takahashi and Marmur, 1963 |
| SP15 | *B. subtilis* | 5-dihydroxypentyluracil replacing thymine | 41% | Brandon *et al*., 1972 |
| SP10 | *B. subtilis* | α-glutamylthymine replacing thymine | 15 – 20% | Krajewska and Shugar, 1971 |
| χP12 | *Xanthomonas oryzae* | 5-methylcytosine replacing cytosine | 100% | Ehrlich *et al*., 1975 |
| S2L | *Synechococcus elongates* | 2-aminoadenine replacing adenine | 100% | Kirnos *et al*., 1977 |
| ϕW14 | *Pseudomonas acidovorans* | α-putrescinylthymine replacing thymine | 50% | Bott and Strauss, 1965 |

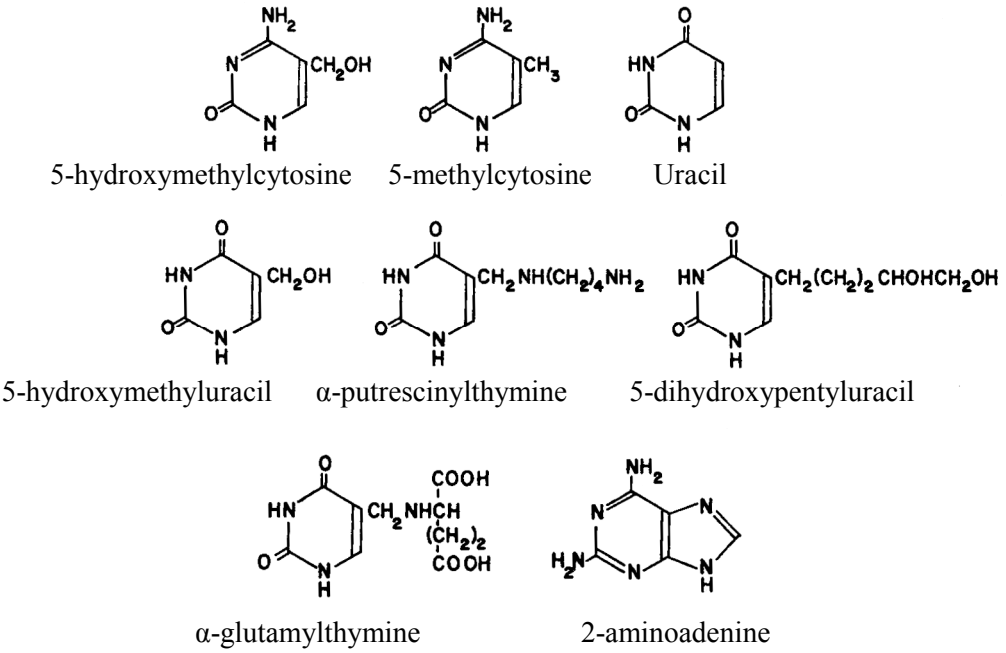



Figure 1.2. Structures of unusual bases in phage DNA (Modified from Warren, 1980).

Loss of endonuclease restriction sites has been shown in many phages, by comparing the actual number of recognition sites to their statistically expected number. For example, the recognition sequence of *B. subtilis* endonuclease *Bsu*RI, GGCC, does not occur once in *B. subtilis* phage ϕ1, although 400 such sites can be predicted statistically in its genome (Kawamura *et al*., 1981). Phage SPO1, which possesses the unusual base hydroxymethyluracil, also lacks the same recognition sequence (Reeve *et al*., 1980), showing multiple strategies used to counter restriction.

In a previous study (Shibayama, 2006) anti-restriction mechanisms were detected in two *Rhodococcus* phages. Results from DNA digestion and cloning have suggested the presence of modified or unusual bases in phages WTS1 and KZA1.

### 1.13 Aim of part three – Unusual or modified bases in *Rhodococcus* phage nucleic acids

The aim of this study is to gain further evidence supporting the presence of unusual or modified bases in phage WTS1 and KZA1 nucleic acids. Where evidence affirms their presence, the aim is to elucidate their structures.

### 1.14 Aims of this study

To recapitulate, the aim of each part of this study is:

**Part one – Plasmid from *Nocardia***

To characterize a plasmid from *Nocardia* to highlight features such as mode of replication and transfer, host range, and functions of genes it carries.

**Part two – Testing of phage genes for inhibition against *Mycobacterium* and construction of a tight genetic switch to control their expression**

To test the previously isolated *Rhodococcus* phage genes for inhibitory activity towards *Mycobacterium* and to construct an inducible expression system using the lambda cI857-$P_R$ that will allow the control of expression of these inhibitory genes in this organism.

**Part three – Unusual or modified bases in *Rhodococcus* phage nucleic acids**

To gain further evidence supporting the presence of unusual or modified bases in the nucleic acids of phages WTS1 and KZA1 and to elucidate their structures.

# 2. MATERIALS AND METHODS

For all solutions and reagents see the Appendix.

## 2.1 Bacteria and bacteriophages

Bacteria, excluding those for plasmid screening and plasmid host range study, are listed in Table 2.1. Strains for plasmid screening are listed in Table 2.2. Those for plasmid host range study are listed in Table 2.3. Phages are listed in Table 2.4.

Table 2.1. Bacterial strains used in this study

| Species and strain | Relevant characteristics | Source/reference [a] |
|---|---|---|
| *Escherichia coli* | | |
| MM294-4 | *hsdR17 endA1 gyrA* | Quan *et al*., 1997 |
| MM294-4 [λ] | MM294-4 lysogenized with λ (cI857 S7) | Quan *et al*., 1997 |
| GM2929 | *dcm-6 dam13::*Tn*9 recF143, hsdR2* | CGSC |
| *Nocardia aobensis* | | |
| IFM 10795 | Source of pYS1 | IFM |
| *Nocardia nova* | | |
| IFM 10797 | Host for plasmid construction in *Nocardia* | IFM |
| *Nocardia mexicana* | | |
| IFO 3927 | Control strain from which total DNA was used for TLC | IFO |
| *Rhodococcus erythropolis* | | |
| SQ1 | Highly transformable derivative of ATCC 4277 | Quan and Dabbs, 1993 |
| *Mycobacterium smegmatis* | | |
| $mc^2155$ | Highly transformable derivative of $mc^26$ | Snapper *et al*., 1990 |
| $mc^2155r$ | Spontaneous rifampicin-resistant mutant of $mc^2155$ | This work |
| *Bacillus subtilis* | | |
| 1A3 | Host for phages PBS1 and SPO1 | BGSC |
| *Staphylococcus aureus* | | |
| unknown strain | Host for phage KG3 | Green, 2007 |

[a] CGSC, *Coli* Genetic Stock Center; IFM, Institute for Food Microbiology, currently Medical Mycology Research Center; IFO, Institute of Fermentation, Osaka; BGSC, *Bacillus* Genetic Stock Center.

Table 2.2. Bacterial strains screened for plasmids

| Species | IFM strain |
|---|---|
| *Nocardia farcinica* | 10757 |
| | 10759 |
| | 10764 |
| | 10773 |
| | 10779 |
| | 10792 |
| *Nocardia cyriacigeorgica* | 10752 |
| | 10762 |
| | 10778 |
| | 10793 |
| | 10796 |
| | 10802 |
| *Nocardia abscessus* | 10763 |
| | 10765 |
| | 10777 |
| | 10788 |
| | 10794 |
| *Nocardia brasiliensis* | 10745 |
| | 10748 |
| | 10781 |
| | 10798 |
| *Nocardia asteroides* | 10756 |
| | 10791 |
| *Nocardia nova* | 10768 |
| | 10797 |
| *Nocardia aobensis* | 10785 |
| | 10795 |
| *Nocardia transvalensis* | 10753 |
| | 10761 |
| *Nocardia otitidiscaviarum* | 10786 |
| *Nocardia beijingensis* | 10772 |

All strains were obtained from the Medical Mycology Research Center.

Table 2.3. Bacterial strains used for plasmid host range study

| Species and strain | Source / reference [a] |
|---|---|
| *Nocardia aobensis* IFM 10795 | IFM |
| *Nocardia aobensis* IFM 10795c | this work |
| *Nocardia nova* IFM 10797 | IFM |
| *Nocardia brasiliensis* IFM 10745 | IFM |
| *Nocardia farcinica* IFM 10757 | IFM |
| *Mycobacterium smegmatis* $mc^2155$ | Snapper *et al*., 1990 |
| *Mycobacterium parafortuitum* IFM 0490 | IFM |
| *Gordonia rubropertincta* ATCC 25593 | ATCC |
| *Rhodococcus erythropolis* ATCC 4277 | ATCC |
| *Rhodococcus rhodochrous* 01 | E. Dabbs |
| *Rhodococcus equi* ATCC 14887 | ATCC |
| *Arthrobacter oxydans* C7 | E. Dabbs |
| *Streptomyces lividans* TK23 | Hütter *et al*., 1985 |
| *Escherichia coli* MM294-4 | Quan *et al*., 1997 |

[a] ATCC, American Type Culture Collection.

Table. 2.4. Bacteriophage stains used in this study

| Phage | Host species | Source reference |
|---|---|---|
| FND1 | *R. erythropolis* | Shibayama, 2004 |
| WTS1 | *R. erythropolis* | Shibayama, 2007 |
| KZA1 | *R. erythropolis* | Shibayama, 2007 |
| PGS1 | *R. erythropolis* | Shibayama, 2007 |
| KG3 | *S. aureus* | Green, 2007 |
| λ | *E. coli* | CGSC |
| T4 | *E. coli* | CGSC |
| PBS1 | *B. subtilis* | BGSC |
| SPO1 | *B. subtilis* | BGSC |

### 2.1.1 Media and growth conditions

All cultures, except those of *Nocardia* and *Gordonia*, were grown in Luria-Bertani (LB) broth with good aeration. All cultures were grown at 37 °C except those of *E. coli* MM294-4[λ], *Rhodococcus*, *M. parafortuitum*, *A. oxydans*, and *S. lividans*, which were grown at 30 °C.

Cultures of *Nocardia* and *Gordonia* were grown in Brain Heart Infusion (BHI) broth supplemented with 1% glucose and 1% glycerol with good aeration. All *Nocardia* were grown at 37 °C and *Gordonia* at 30 °C.

For all plate cultures, growth media were solidified with 1.5% agar.

For short-term storage, all strains were kept on agar plates at 4°C. For long-term storage, cells suspended in 30% glycerol were stored at -70°C.

*Rhodococcus* phages and *Staphylococcus* phages were propagated according to previously reported multiplicities of infection (MOI) (Green, 2007; Shibayama 2004; Shibayama, 2007). Optimal MOI for the propagation of *Escherichia* phages and *Bacillus* phages were empirically determined.

### 2.2 Plasmids

All plasmids used in this study are listed in Table 2.5.

Table 2.5. Plasmids

| Plasmid | Description | Reference/source |
|---|---|---|
| pYS1 | 4326 bp cryptic plasmid from *N. aobensis* IFM 10795 | This study |
| pK18 | *E. coli* vector with kanamycin resistance marker | Pridmore, 1987 |
| pYS1R1 | Position 1904-3966 of pYS1 amplified with primers PYS1Rep48F1 and PYS1Rep9R1 and ligated into *Nhe*I site of pK18 | This study |
| pYS1R1S1 | Sequence TAGCCA of second exposed loop of predicted SSO mutagenized *in-vitro* to TGCAGA in pYS1R1 | This study |
| pYS1R1D1 | Sequence TTGGAA of predicted DSO mutagenized *in-vitro* to AGATCT in pYS1R1 | This study |
| pYS1R2 | Position 1904-3597 of pYS1 amplified with primers PYS1Rep48F1 and PYS1Rep14R1 and ligated into *Nhe*I site of pK18 | This study |
| pYS1R3 | Position 1990-3597 of pYS1 amplified with primers PYS1Rep1990F1 and PYS1Rep14R1 and ligated into *Nhe*I site of pK18 | This study |
| pYS1R4 | Position 2050-3597 of pYS1 amplified with primers PYS1Rep2050F1 and PYS1Rep14R1 and ligated into *Nhe*I site of pK18 | This study |
| pYS1R5 | Position 2100-3597 of pYS1 amplified with primers PYS1Rep2100F1 and PYS1Rep14R1 and ligated into *Nhe*I site of pK18 | This study |
| pYS1R6 | Position 2182-3966 of pYS1 amplified with primers PYS1Rep54F1 and PYS1Rep9R1 and ligated into *Nhe*I site of pK18 | This study |
| pYS1R7 | Position 2182-3597 of pYS1 amplified with primers PYS1Rep54F1 and PYS1Rep14R1 and ligated into *Nhe*I site of pK18 | This study |
| pNV18/19 | *Nocardia – E. coli* shuttle vector | Chiba *et al*., 2007 |
| pUC4K | *E. coli* plasmid containing excisable *aph* gene from Tn903 | Amersham |
| pYS1K1 | 304 bp *Pst*I fragment from pYS1 replaced with 1240 bp *Pst*I fragment of pUC4K containing *aph* gene, in *N. nova* IFM 10797 | This study |
| pYS1K2 | 1264 bp *Bam*HI fragment of pUC4K containing *aph* gene inserted into *Bgl*II site of pYS1, in *N. nova* IFM 10797 | This study |
| pYS1K3 | 1320 bp fragment of pUC4K containing *aph* gene amplified with primers PUC4K10F1 and PUC4K38R1 and inserted into the *Age*I site of pYS1, in *N. nova* IFM 10797 | This study |
| pYS1K4 | 1252 bp *Sal*I fragment of pUC4K containing *aph* gene inserted into *Xho*I site of pYS1, in *N. nova* IFM 10797 | This study |
| pYSA1 | 21 bp *Sac*I-*Xba*I fragment of pUC18 replaced with 2903 bp *Sac*I-*Xba*I fragment from pYS1, in *E. coli* GM2929 | This study |
| pYSA2 | 1320 bp fragment of pUC4K containing *aph* gene amplified with primers PUC4KFse10F1 and PUC4KFse38R1 and inserted into the *Fse*I site of pYSA1, in *E. coli GM2929* | This study |
| pYS1K5 | 4231 bp *Sac*I-*Xba*I fragment from pYSA2 containing *aph* gene ligated to 1423 bp *Sac*I-*Xba*I fragment from pYS1, in *N. nova* IFM 10797 | This study |
| pUC18 | General purpose *E. coli* vector | Fermentas |

| | | |
|---|---|---|
| p106 | *E. coli* vector containing promoter-less *cat* gene downstream of a multiple cloning site | Atchison *et al.*, 1989 |
| pNVCAT | ~1900 bp *Hpa*I-*Psc*I fragment (filled-in) containing *cat* gene from p106 ligated to 4016 bp *Hind*III-*Psc*I fragment (filled-in) of pNV18 | This study |
| pNVCATP1 | Position 4214-427 of pYS1 amplified with primers PYS1PromF1 and PYS1PromR1 and ligated into *Bam*HI site of pNVCAT in *orf1* orientation with respect to *cat* | This study |
| pNVCATP2 | Position 4214-427 of pYS1 amplified with primers PYS1PromF1 and PYS1PromR1 and ligated into *Bam*HI site of pNVCAT in *orf3* orientation with respect to *cat* | This study |
| pNVCATP3 | Position 4115-539 of pYS1 amplified with primers PYS1PromF2 and PYS1PromR2 and ligated into *Bam*HI site of pNVCAT in *orf1* orientation with respect to *cat* | This study |
| pNVCATP4 | 27 bp *Hind*III-*Bam*HI fragment of pNVCAT replaced with PCR fragment spanning positions 3605-427 of pYS1 amplified with primers PYS1ORF3F1 and PYS1PromR1, in *orf1* orientation with respect to *cat* | This study |
| pNVCATP5 | 27 bp *Hind*III-*Bam*HI fragment of pNVCAT replaced with PCR fragment spanning positions 3605-539 of pYS1 amplified with primers PYS1ORF3F1 and PYS1PromR2, in *orf1* orientation with respect to *cat* | This study |
| pYS1R1Gyr | ~200 bp segment of *gyrA* from *N. aobensis* IFM 10795 amplified with primers GyrAF1 and GyrAR1 and cloned into the *Hind*III site of pYS1R1 | This study |
| pDA71 | *Rhodococcus-E. coli* shuttle vector with an *ecoR* suicide gene | Quan and Dabbs, 1993 |
| pFB7 | ~1600 bp inhibitory *Bgl*II fragment from *Rhodococcus* phage FND1 DNA cloned into pDA71 | Shibayama, 2007 |
| pFB8 | ~2200 bp inhibitory *Bgl*II fragment containing a gene resembling those encoding a thymidilate synthase complementing protein, from phage FND1 DNA cloned into pDA71 | Shibayama, 2007 |
| pFB10 | ~4600 bp inhibitory *Bgl*II fragment from phage FND1 DNA cloned into pDA71 | Shibayama, 2007 |
| pFB19 | ~3750 bp inhibitory *Bgl*II fragment from phage FND1 DNA cloned into pDA71 | Shibayama, 2007 |
| pFP8 | ~1600 bp inhibitory *Pst*I fragment containing a gene resembling those encoding a phage capsid protein, from phage FND1 DNA cloned into pDA71 | Shibayama, 2007 |
| pFP14 | ~2700 bp inhibitory *Pst*I fragment containing a gene resembling those encoding a HNH endonuclease, from phage FND1 cloned into pDA71 | Shibayama, 2007 |
| pFP16 | ~2100 bp inhibitory *Pst*I fragment from phage FND1 DNA cloned into pDA71 | Shibayama, 2007 |
| pPH1 | ~2700 bp inhibitory *Hind*III fragment containing a gene resembling those encoding a phage head decoration protein, from *Rhodococcus* phage PGS1 cloned into pDA71 | Shibayama, 2007 |
| pPH7 | ~4600 bp inhibitory *Hind*III fragment containing a gene resembling those encoding an □-subunit of DNA polymerase III, from phage PGS1 cloned into pDA71 | Shibayama, 2007 |
| pOLYG | *Mycobacterium-E. coli* shuttle vector | O'Gaora *et al*., 1997 |
| pFBM7 | ~1600 bp *Bgl*II insert from pFB7 cloned into pOLYG | This study |
| pFBM8 | ~2200 bp *Bgl*II insert from pFB8 cloned into pOLYG | This study |
| pFBM10 | ~4600 bp *Bgl*II insert from pFB10 cloned into pOLYG | This study |
| pFBM19 | ~3750 bp *Bgl*II insert from pFB19 cloned into pOLYG | This study |
| pFPM8 | ~1600 bp *Pst*I insert from pFB8 cloned into pOLYG | This study |
| pFPM14 | ~2700 bp *Pst*I insert from pFP14 cloned into pOLYG | This study |
| pFPM16 | ~2100 bp *Pst*I insert from pFP16 cloned into pOLYG | This study |
| pPHM7 | ~4600 bp *Hind*III insert from pPH7 cloned into pOLYG | This study |
| pEcoR251 | *E. coli* vector with *ecoR* suicide gene | Zabeau and Stanley, 1982 |
| pYS13 | cI857 cloned into *Bam*HI site of pEcoR251 by shotgun cloning of *Sau*3AI partial digest of λ genome | This study |

| | | |
|---|---|---|
| pYS13* | ~900 bp *Eco*RI-*Sfu*I fragment within *ecoR* deleted from pYS13 | This study |
| pYS13-1 | *cI857* amplified from pYS13 using forward primer containing *Bam*HI site and reverse primer containing *Bgl*II site and cloned into *Bam*HI site of pEcoR251 | This study |
| pYS13-1* | ~190 bp *Bgl*II-*Pst*I fragment within *ecoR* from pYS13-1 deleted | This study |
| pYSM13-1 | ~2400 bp *Eco*R-*Psc*I fragment containing *ecoR* and *cI857* from pYS13-1 ligated to ~4200 *Eco*R-*Psc*I backbone of pNV19 | This study |
| pYSM13-1* | ~2200 bp *Eco*R-*Psc*I fragment containing deleted *ecoR* and *cI857* from pYS13-1* ligated to ~4200 *Eco*R-*Psc*I backbone of pNV19 | This study |

## 2.3 Primers

All primers, excluding those for site-directed mutagenesis, are listed in Table 2.6. Primers for site-directed mutagenesis are listed in Table 2.7.

Table 2.6. Primers

| Name | Sequence (5'-3') (introduced enzyme site underlined) | Enzyme site introduced | Region amplified |
|---|---|---|---|
| PYS1Rep48F1 | gatcgctagcacatctcggcacattgcccga | *Nhe*I | Position 1904-3966 of pYS1 spanning the SSO, DSO, and *rep* |
| PYS1Rep9R1 | gatcgctagccgatggggctgaccgttacc | *Nhe*I | |
| PYS1Rep48F1 | gatcgctagcacatctcggcacattgcccga | *Nhe*I | Position 1904-3597 of pYS1 spanning the SSO, DSO, and *rep* |
| PYS1Rep14R1 | gatcgctagctaacggtccaatctccgcta | *Nhe*I | |
| PYS1Rep1990F1 | gatctagctagccccgaaacgcactgaggg | *Nhe*I | Position 1990-3597 of pYS1 spanning a portion of SSO, DSO, and *rep* |
| PYS1Rep14R1 | gatcgctagctaacggtccaatctccgcta | *Nhe*I | |
| PYS1Rep2050F1 | gatctagctagccttgcctagccagcacgg | *Nhe*I | Position 2050-3597 of pYS1 spanning a portion of SSO, DSO, and *rep* |
| PYS1Rep14R1 | gatcgctagctaacggtccaatctccgcta | *Nhe*I | |
| PYS1Rep2100F1 | gatctagctagctcgatggtagcaactgcg | *Nhe*I | Position 2100-3597 of pYS1 spanning the DSO and *rep* |
| PYS1Rep14R1 | gatcgctagctaacggtccaatctccgcta | *Nhe*I | |
| PYS1Rep54F1 | gatcgctagcaatccggacgacctcgccggta | *Nhe*I | Position 2182-3966 of pYS1 spanning the DSO and *rep* |
| PYS1Rep9R1 | gatcgctagccgatggggctgaccgttacc | *Nhe*I | |
| PYS1Rep54F1 | gatcgctagcaatccggacgacctcgccggta | *Nhe*I | Position 2182-3597 of pYS1 spanning the DSO and *rep* |
| PYS1Rep14R1 | gatcgctagctaacggtccaatctccgcta | *Nhe*I | |
| PUC4K10F1 | gatcaccggtgctgaggtctgcctcgtgaa | *Age*I | Position 435-1754 of pUC4K spanning *aph* |
| PUC4K38R1 | gatcaccggtcggctcgtatgttgtgtgga | *Age*I | |
| PUC4KFse10F1 | attataggccggccgctgaggtctgcctcgtgaa | *Fse*I | Position 435-1754 of pUC4K spanning *aph* |
| PUC4KFse38R1 | attataggccggcccggctcgtatgttgtgtgga | *Fse*I | |
| pYS1PromF1 | tgacggatccgtcgtatctccttgcgattc | *Bam*HI | Position 4214-427 of *orf3-orf1* intergenic region of pYS1 |
| pYS1PromR1 | tgacggatcccttgccacgaatcttcgtgt | *Bam*HI | |
| PYS1PromF2 | tgacggatccaggcccttcctgtgagattc | *Bam*HI | Position 4115-539 of *orf3-orf1* intergenic region of pYS1 |
| PYS1PromR2 | tgacggatccagtgtttctcctggtcagtg | *Bam*HI | |
| PYS1ORF3F1 | tgatctaagcttgcatgaggtgtgaccctc | *Hind*III | Position 3605-427 of pYS1 spanning *orf3* and *orf3-orf1* intergenic region |
| PYS1PromR1 | tgacggatcccttgccacgaatcttcgtgt | *Bam*HI | |
| PYS1ORF3F1 | tgatctaagcttgcatgaggtgtgaccctc | *Hind*III | Position 3605-539 of pYS1 spanning *orf3* and *orf3-orf1* intergenic region |
| PYS1PromR2 | tgacggatccagtgtttctcctggtcagtg | *Bam*HI | |
| GyrAF1 | gtacaagcttgcgcatcgtggtgacggtcaa | *Hind*III | ~200 bp region of *N. aobensis gyrA*, based on *N. farcinica* genome sequence |
| GyrAR1 | gtacaagcttcctggtgcgccggatgatgac | *Hind*III | |
| RtGyrAF | gcgcatcgtggtgacggtcaa | none | ~200 bp region of *N. aobensis gyrA*, based on *N. farcinica* genome sequence, for real-time PCR |
| RtGyrAR | cctggtgcgccggatgatgac | none | |

| | | | |
|---|---|---|---|
| RtRep2100F | tcgatggtagcaactgcg | none | Position 2100-2310, within the *rep* gene of pYS1, for real-time PCR |
| RtRep2310R | ctcctgctaccgcccgct | none | |
| CI4F1 | gtacggatcctcaacacgcacggtgttaga | *Bam*HI | cI857 with $P_{RM}$, $O_R2$, and $O_R3$ from pYS13 |
| CI41R1 | gctaagatcttgccgatcagccaaacgtc | *Bgl*II | |

Some primers above appear more than once when used in combination with different primers to amplify different regions.

Table 2.7. Primers for site-directed mutagenesis

| Name | Sequence (5'-3') (mutated region capitalized; introduced enzyme site underlined) | Enzyme site introduced | Mutation introduced |
|---|---|---|---|
| SSOloop2F | ggccttcttgcctGCAGagcacgggccgg | *Pst*I | agcc of second terminal loop of pYS1 SSO mutagenized to gcag |
| SSOloop2R | ggataagccccgagggcgaaccagtgcat | none | |
| DSOF | ggcggggaggccAGATCTcaccccggcgag | *Bgl*II | ttggaa of DSO of pYS1 mutagenized to agatct |
| DSOR | tccgggcaaacccgcctgcggcgg | none | |

## 2.4 Isolation of bacteriophages

### 2.4.1 Soil assay

One gram of soil was incubated in 10 ml of LB broth supplemented with 10 mM $CaCl_2$ and 10 mM $MgCl_2$ in a 100 ml flask on a rotary shaker at ~80 rpm overnight at 30 °C. The soil suspension was centrifuged in a Beckman JA-20 rotor at 15000 rpm for 10 min, after which the supernatant was decanted. 2 ml of this supernatant was added to 50 µl of a stationary phase culture of *M. smegmatis* $mc^2155r$, vortexed for 1 sec, and left at 37 °C for 30 min. Then 2 ml of sloppy agar was added, mixed by rolling between the palms, and poured onto an LB agar plate supplemented with 10 mM $CaCl_2$, 10mM $MgCl_2$, 100 µg/ml rifampicin, and 50 µg/ml nystatin. The plate was incubated for ~2 days and checked for plaques.

### 2.4.2 Production of phage suspensions

A plaque was picked by stabbing a toothpick through the soft agar and dipping into 100 µl LB broth. The suspension was vortexed for 1 sec, left for 5 min, and microfuged for 3 min. The supernatant containing the phage was collected.

### 2.4.3 Single plaque purification

Phage suspension was serially diluted with LB broth. For the last dilution, 10 µl was added to 2 ml LB broth and 50 µl stationary phase bacterial culture, vortexed for 1 sec,

and left at 30 °C or 37 °C (whichever was the optimum growth temperature for the host bacteria) for 30 min. 2 ml sloppy agar was added and poured onto an agar plate containing 10 mM $CaCl_2$ and 10mM $MgCl_2$. The plate was incubated at 30 °C or 37 °C until plaques appeared.

### 2.5 Phage lysate production

#### 2.5.1 Plate lysate production

10 µl phage suspension was added to 2 ml LB broth and 50 µl host bacterial culture, vortexed for 1 sec, left at 30 °C or 37 °C for 30 min, and overlaid onto an LB agar plate containing 10 mM $CaCl_2$ and 10 mM $MgCl_2$. Multiplicity of infection (MOI) was adjusted by diluting the phage suspension when necessary. The plate was incubated at 30 °C or 37 °C until a significant difference was observable between this plate and the no-phage-control plate.

Onto the plate lysate, 2 ml LB broth was added. The entire top agar layer, together with all the liquid, was collected with a spatula and transferred to a centrifuge tube. The tube was vortexed for 5 sec, left for 5 min, and centrifuged in a JA-20 rotor at 15000 rpm for 10 min. The supernatant was collected. When necessary, the lysate was supplemented with antibiotics or filter sterilized.

#### 2.5.2 Small scale liquid lysate production

200 µl phage suspension was added to 250 µl bacterial culture and 10 ml LB broth supplemented with 10 mM $CaCl_2$ and 10 mM $MgCl_2$ in a 100 ml flask. MOI was adjusted by diluting the phage suspension when necessary. The flask was left at 30 °C or 37 °C for 30 min, followed by incubation on a rotary shaker at ~80 rpm until a significant difference was observable between this flask and the no-phage-control flask. The lysate was centrifuged in a JA-20 rotor at 15000 rpm for 10 min. The supernatant was collected. When necessary, the lysate was supplemented with antibiotics or filter sterilized.

### 2.5.3 Large scale liquid lysate production

The condition optimal for the production of 10 ml lysate was applied to the production of 500 ml lysate, by appropriately scaling up all components. This was done by preparing five 100 ml lysates, each in 1000 ml flasks.

## 2.6 Phage purification

To the large scale lysate NaCl was added to a final concentration of 1 M and dissolved by slow stirring on a magnetic stirrer. Then polyethylene glycol (PEG 6000) was added to a final concentration of 10% (w/v) and again dissolved in the same manner. The lysate was cooled in an ice-water slurry for at least 1 hr and centrifuged in a JA-10 rotor at 9000 rpm for 20 min at 4°C. The supernatant was discarded, and the phage pellet was purified in either a step gradient or equilibrium gradient of CsCl.

### 2.6.1 Phage purification in a step gradient

The phage pellet was dissolved in 3.5 ml of SC buffer containing 50% (w/v) CsCl and transferred to a Beckman Quick-Seal centrifuge tube using a Pasteur pipette. A step gradient was produced by layering 800 µl of CsCl solutions of increasing density (1.45, 1.50, 1.70 g/ml) at the bottom of the tube. The tube was centrifuged in a Beckman VTi 65.2 rotor at 25000 rpm for 1 hr at 4°C. The bluish band of phage was extracted using a needle and a syringe.

### 2.6.2 Phage purification in an equilibrium gradient

The phage pellet was resuspended in 5 ml of SC buffer containing 81.7% (w/v) CsCl. This was loaded into a Beckman Quick-Seal centrifuge tube and centrifuged at 38000 rpm for 24 hrs at 4°C. The phage band was extracted using a needle and a syringe.

## 2.7 DNA preparations

### 2.7.1 Phage genomic DNA preparation

#### 2.7.1.1 Dialysis of purified phage in CsCl

Dialysis tubing was prepared according to Sambrook *et al*., 1989. Tubing of ~10 cm lengths were boiled for 10 min in 800 ml of 2% $NaHCO_3$ and 1 mM EDTA (pH 8.0)

and rinsed in $dH_2O$. The tubing was then boiled for 10 min in 800 ml of 1 mM EDTA (pH 8.0) and rinsed in $dH_2O$.

150-300 µl of purified phage suspension diluted to 1 ml with $dH_2O$ was transferred into the dialysis tubing which was then dialyzed for 1 hr against 1000 volumes of a buffer containing 10 mM NaCl, 50 mM Tris-HCl (pH 8.0), and 10 mM $CaCl_2$ while gently stirring with a magnetic stirrer. The dialysis sac was then transferred to a fresh flask of buffer and dialyzed for another 1 hr.

#### 2.7.1.2 Phage genomic DNA extraction

To obtain DNA from a plate lysate or purified phage, DNaseI and RNase were added to a final concentration of 10 µg/ml and incubated at room temperature for 6 hrs before subjecting it to the following steps. To the phage suspension, EDTA (pH 8.0) was added to a final concentration of 20 mM, proteinase K to a final concentration of 50 µg/ml, and SDS to a final concentration of 0.5% (w/v). The mixture was incubated at 56°C for 1 hr, cooled to room temperature, subjected to three phenol extractions and one chloroform extraction. A salt-ethanol precipitation was performed on the phage DNA, which was then dissolved in $dH_2O$ or 4 mM Tris-HCl (pH 8.0).

### 2.7.2 Bacterial bulk total DNA extraction

A 200 ml culture of *E. coli*, *Nocardia*, *Rhodococcus*, *Bacillus*, or *Staphylococcus* was grown in appropriate media supplemented with retardatory concentration of glycine (*Nocardia* spp., 3%; *R. erythropolis*, 2%; *B. subtilis*, 1%; *S. aureus*, 7%; *E. coli*, no glycine necessary). Cells were centrifuged in a Beckman JA-10 rotor at 6000 rpm for 10 min at room temperature. The pellet was resuspended in 5 ml of TE supplemented with lysozyme at a concentration of 5 mg/ml and incubated for >1 hr at 37 °C. A minute quantity of proteinase K and a 1/10 volume of 10% TE-SDS were added and incubated for >1 hr at 37 °C. The viscous lysate was transferred to a Beckman 50Ti tube and spun at 40000 rpm for 30 min. The supernatant was transferred to a JA-20 tube, to which 4.4 g CsCl was added and dissolved. The tube was centrifuged at 15000 rpm in a JA-20 rotor for 15 min, after which the supernatant was transferred to a clean bottle. 400 µl of 1% ethidium bromide (EtBr) was added, and its refractive index was adjusted to between 1.391 and 1.392 using a refractometer. The solution was transferred into a Quick-Seal tube and centrifuged at 45000 rpm in a VTi65.2 rotor for 16 hrs. Under a UV light, the upper band representing chromosomal DNA

was extracted with a needle and syringe. Where a lower plasmid band was available and required, it was also extracted in a similar manner. EtBr from the extracted DNA was removed by butanol extraction, and CsCl was removed by ethanol precipitation.

### 2.7.3 *E. coli* bulk plasmid preparation

A 100 ml culture of *E. coli* transformant was grown to stationary phase in the presence of appropriate antibiotic, after which cells were harvested by centrifugation in a JA-10 rotor at 6000 rpm for 10 min. Cells were resuspended in 5 ml of solution I and transferred to a JA-20 tube. Then, 10 ml of solution II was added, mixed by gentle inversion, and incubated for 15 min at room temperature. Thereafter, 7.5 ml of solution III was added, shaken vigorously, and left in an ice-water slurry for 10 min. Cellular debris was removed by centrifugation in a pre-cooled (4°C) JA-20 rotor at 15000 rpm for 10 min. The supernatant was transferred to a clean JA-20 tube, warmed to room temperature, to which 12 ml of isopropanol was added and left for 10 min. The tube was centrifuged at 15000 rpm for 15 min. The supernatant was decanted, after which the pellet was washed with 2 ml of 96% ethanol and vacuum dried for 20 min. The DNA was resuspended in 4 ml TE for 2 hrs at 30 °C. 400 µl of 1% EtBr and 4.1 g of CsCl were added, inverted to mix, and the refractive index was adjusted to between 1.387 and 1.389 using a refractometer. This was loaded into a Quick-Seal tube and centrifuged in a VTi 65.2 rotor at 45000 rpm for 16 hrs. The plasmid band was extracted under a UV light using a needle and syringe. EtBr was removed from the DNA by butanol extraction, and CsCl was removed by ethanol precipitation.

### 2.7.4 *E. coli* plasmid mini preparation

1 ml cultures of *E. coli* transformants were grown to stationary phase in the presence of appropriate antibiotic. Cells were harvested by microfuging for 30 sec, the supernatant decanted, and the pellet resuspended in 80 µl of solution I. 160 µl of solution II was mixed in by gentle inversion, left for 10 min, 120 µl of solution III added, shaken vigorously, and left in an ice-water slurry for 5 min. Cellular debris was removed by microfuging for 5 min at 4°C. The supernatant was transferred to a clean microfuge tube and warmed to room temperature. 220 µl isopropanol was added and after 10 min at room temperature the tube was microfuged for 5 min, the supernatant decanted, then the pellet was washed with 150 µl of 96% ethanol and

vacuum desiccated for 20 min. The DNA was dissolved in 100 µl $dH_2O$ or 4 mM Tris-HCl (pH 8.0) containing 1 µl of 10 µg/ml RNase.

### 2.7.5 Gram-positive plasmid mini preparation

A 1 ml culture of *Nocardia* or *Mycobacterium* grown to stationary phase in the presence of retardatory concentration of glycine (*Nocardia* spp., 3%; *M. smegmatis*, 0.5%) was microfuged, and the pellet was suspended in 1 ml of resuspension solution supplemented with lysozyme at a concentration of 5 mg/ml. Following incubation for >3 hrs at 37 °C on a shaker, the cells were collected by microfuging and resuspended in 280 µl TE. 280 µl 10% TE-SDS was added and mixed by gentle inversion for 20 min. 73 µl of 4.5 M sodium acetate (pH 6.0) was added, mixed by gentle inversion, and left on ice for 1 hr. The tube was microfuged for 20 min at 4 °C, after which the supernatant was collected and purified by phenol-chloroform extraction. The DNA was salt-ethanol precipitated, vacuum dried for 20 min, and resuspended in $dH_2O$ or 4 mM Tris-HCl (pH 8.0).

### 2.7.6 Bacterial plasmid mini preparation using a kit

For purer plasmid mini preparations, GeneJET Plasmid Miniprep Kit (Fermentas) was used. For *E. coli*, the kit was used according to the manufacturer's instructions. For Gram-positive bacteria, two modifications (Parish and Stoker, 1998) were made. First, lysozyme was added to a concentration of 5 mg/ml following resuspension of the cell pellet in Resuspension Solution and incubated for >3 hrs at 37 °C on a shaker. Second, after passing the cleared lysate through the column, a 1:1 mixture of chloroform-methanol (250 µl each) was used to wash the column. This was followed by the regular washing steps using the Wash Solution provided by the manufacturer.

## 2.8 DNA manipulations

### 2.8.1 Butanol extraction

EtBr was removed from DNA by adding 1/10 volume of butanol, mixing by inversion, and microfuging for 10 sec. Upper layer containing EtBr was removed. Extraction was repeated ~5 times to remove all traces of EtBr.

### 2.8.2 Phenol-chloroform extraction

For DNA volumes of <300 µl, TE buffer was added to make it up to this volume. Then, 1/3 volume of TE-saturated phenol was added, mixed by inversion, and microfuged for 5 min at room temperature. The aqueous phase was collected, to which 1/3 volume of chloroform was added, mixed by inversion, and microfuged for 30 sec. The aqueous phase was collected, and the DNA was subjected to salt and ethanol precipitation.

### 2.8.3 Salt and ethanol precipitation

DNA was precipitated by adding 1/10 volume of 1 M NaCl and 2 volumes of 96% ethanol. This was mixed by gentle inversion and microfuged for 20 min at 4 °C. The supernatant was decanted, and the tube was blotted on paper towel, followed by vacuum drying for 20 min. The DNA was resuspended in an appropriate volume of $dH_2O$ or 4 mM Tris-HCl (pH 8.0).

### 2.8.4 Restriction digestions

Restriction endonucleases were obtained from Fermentas, Roche, or New England Biolabs and used according to their instructions. An appropriate volume of 10× digestion buffer was added to the DNA. For double digestions, a common buffer in which both enzymes showed suitable activity was chosen. Then, suitable units of enzyme were added and incubated at an appropriate temperature for >3 hrs. When necessary, the endonuclease was inactivated by either heat treatment or phenol extraction. For enzymes which could be heat-inactivated, the manufacturer's instruction was followed for the correct incubation temperature and time. Otherwise, phenol-chloroform extraction was conducted.

### 2.8.5 Removal of overhangs

DNA with overhangs was treated with Klenow fragment (Fermentas) to produce blunt ends. An appropriate volume of 10× reaction buffer and suitable units of the enzyme were added in incubated at 37 °C for 20 min. The enzyme was then heat-inactivated at 75 °C for 10 min.

### 2.8.6 Detection and digestion of single-stranded DNA

For the detection and digestion of single-stranded form of plasmids, S1 nuclease (Fermentas) was used. In a total reaction volume of 15 µl, ~300 ng of *Nocardia* total DNA, 1.5 units of enzyme, and 3 µl of 5× reaction buffer were added. The reaction was incubated at room temperature for 30 min, followed by enzyme inactivation by adding EDTA to a final concentration of 30 mM and heating at 70 °C for 10 min. The digested DNA was run on a gel and examined by Southern analysis.

### 2.8.7 Dephosphorylation of 5' ends

FastAP Thermosensitive Alkaline Phosphatase (Fermentas) was used to prevent recircularization of vector. An appropriate volume of 10× dephosphorylation buffer and suitable units of alkaline phosphatase were added and incubated at 37°C for 20 min. The enzyme was then heat-inactivated at 75°C for 5 min.

### 2.8.8 Ligation

T4 DNA ligase (Fermentas) was used for ligation reactions. For sticky-end and blunt-end ligations, 5 units and 25 units, respectively, of the enzyme were used. After mixing DNA to be ligated, the reaction volume was made up to 20 µl by adding 2 µl reaction buffer and appropriate volumes of $dH_2O$ and enzyme. For sticky-end ligations, reactions were incubated at 22 °C for ≥2 hrs. For blunt-end ligations, reactions were incubated overnight. The ligase was then inactivated at 65 °C for 10 min.

### 2.8.9 Freeze-squeeze method of DNA purification from agarose gels

Piece of agarose containing the DNA of interest was placed in a microfuged tube, crushed with a spatula, and frozen at -70°C for >2 hrs. The agarose was then completely thawed at room temperature, followed by microfuging for 15 min after which the supernatant was collected. The pellet was crushed again, and the process was repeated. The supernatants from the first and second rounds were pooled and subjected to two phenol extractions and one chloroform extraction. The DNA was salt-ethanol precipitated, vacuum dried for 20 min, and resuspended in 50 µl of $dH_2O$ or 4 mM Tris-HCl (pH 8.0).

### 2.8.10 Determination of DNA concentration

DNA concentration was measured using the NanoDrop ND-1000 Spectrophotometer at $OD_{260}$. DNA purity with respect to protein contamination was assessed by measuring $OD_{260}/OD_{280}$ ratio. Where a suitable solution to blank the reading on the spectrophotometer was not available, an alternative method was used. An aliquot of the DNA was run on a gel, which was then quantified using the UVP LabWorks Image Acquisition and Analysis Software (Ver.4.5) by comparing band intensity between the DNA and the molecular weight marker, for which the concentration had been predetermined by the manufacturer.

## 2.9 Gel electrophoresis

### 2.9.1 Agarose gel electrophoresis

Agarose stock solutions were prepared in 0.5× TBE buffer at concentrations of 0.4-1.4% depending on the size of the DNA fragment to be separated. A gel was prepared by pouring 30 ml of agarose solution containing 1 µg/ml EtBr into a gel tray inserted with a comb, which was then polymerized at 4 °C for 20 min. After removing the comb, the tray containing the gel was placed in an electrophoresis unit containing 0.5× TBE buffer with 1 µg/ml EtBr. After DNA samples were mixed with bromophenol blue tracking dye and loaded into the wells, electrophoresis was conducted at 80-100 V at room temperature. The gel was viewed and photographed using the BioRad Molecular Imager Gel Doc XR system. Molecular weight marker used was GeneRuler 1 kb Plus DNA Ladder (Fermentas).

### 2.9.2 Low gelling agarose gel electrophoresis

Low gelling agarose of 1% concentration was prepared in 0.5× TBE buffer. EtBr was added to a final concentration of 1 µg/ml, and the agarose was poured into a tray inserted with a comb and polymerized at 4°C for 20 min. The gel was placed in a pre-cooled electrophoresis unit containing 0.5× TBE with 1 µg/ml EtBr. After DNA samples were mixed with bromophenol blue tracking dye and loaded into the wells, electrophoresis was conducted at 75 V at 4°C. Using a scalpel under a UV light, a band of interest was excised from the gel, from which the DNA was extracted by either phenol-chloroform extraction or freeze-squeeze method. In case of the phenol-chloroform extraction, the piece of agarose carrying the DNA was melted at 60°C for

30 min, followed by three phenol extractions. Between each extraction, 1/15 volume of 1 M Tris-HCl (pH 8.0) was added. One chloroform extraction was then conducted, followed by precipitation of the DNA.

### 2.10 Transformation and conjugation

#### 2.10.1 $CaCl_2$-mediated transformation of *E. coli*

20 ml of pre-warmed LB containing 0.5% glucose was inoculated with 200 µl of an overnight culture of *E. coli* and incubated at 37°C with vigorous shaking for 1.75 hrs. The culture, while still being shaken, was cooled in an ice-water slurry for 5 min and centrifuged in a pre-cooled JA-20 rotor at 10000 rpm for 5 min at 4°C. The supernatant was decanted, and the pellet was resuspended in 10 ml of pre-cooled $CaCl_2$ transformation buffer and placed on ice for 15 min, followed by centrifugation at 10000 rpm for 5 min at 4°C. The supernatant was decanted, and the pellet was resuspended in 1.3 ml of pre-cooled transformation buffer. The cells were left on ice for 2-24 hrs.

100 µl of competent cells and a suitable volume of plasmid were added together, mixed by bubbling air with a pipette, and left on ice for 15 min. They were heat-shocked at 42 °C for 90 sec, after which 0.5 ml of pre-warmed LB was added and incubated at 37°C for 1 hr. The cells were spread on an LB agar plate containing an appropriate concentration of antibiotics and incubated at 37 °C overnight.

#### 2.10.2 Actinomycete electroporation

Actinomycetes were grown to late logarithmic phase in media supplemented with retardatory concentration of glycine (*Nocardia* spp., 3%; *Mycobacterium* spp., 0.5%; *Rhodococcus* spp., 2%; *G. rubropertincta*, 3%; *S. lividans*, 0.5%; *A. oxydans*, 1%). Cells were chilled in an ice-water slurry for 20 min and centrifuged in a pre-chilled JA-20 rotor at 10000 rpm for 10 min. The pellet was resuspended in electroporation buffer and washed twice in the same buffer. 200 µl of cells were mixed with 1-10 µl of DNA and transferred into a BioRad 0.2 cm electroporation cuvette. Electroporation was conducted using the BioRad Gene Pulser I with the following parameters: voltage 2.5 kV; capacitance 25 µF; resistance 1000Ω. The cells were transferred to a clean bottle to which 2 ml of appropriate growth media was added, followed by incubation

at 30 °C or 37 °C for 2 hrs. Transformants were selected by spreading on agar plates supplemented with appropriate concentration of antibiotics.

### 2.10.3 *Rhodococcus* PEG-mediated transformation

*R. erythropolis* SQ1 was grown in 5 ml of LBSG (2% glycine) at 30°C for 2 days. Cells from 1 ml of this culture were washed once in 1 ml of B buffer, resuspended in 1 ml of B buffer containing freshly added lysozyme (5 mg/ml), and incubated at 37 °C for 1.5 hrs with inversion every 10 min. During the incubation, 2 ml of P buffer was prepared. P-PEG buffer was also prepared by UV-sterilizing 0.5 g of polyethylene glycol (PEG 6000) for 10 min and dissolving it in 1 ml of P-buffer by vigorous vortexing.

The cells were microfuged for 10 sec, gently washed in 1 ml of B buffer, and resuspended in 500 µl of P buffer. 25 µl of the protoplast suspension and 5 µl of plasmid were added together, mixed by bubbling air with a pipette, and left for 10 min. An equal volume of P-PEG buffer was added, mixed by bubbling, and spotted onto a pre-cooled regeneration plate. The plate was incubated at 30°C for 12 hrs, underlaid with 250 µl of 4 mg/ml chloramphenicol, and further incubated at 30°C for 5 days.

### 2.10.4 Conjugation

*N. aobensis* IFM 10795c transformed with insertion mutants of pYS1 were used as donors, and *N. nova* IFM 10797 as recipients. Both donor and recipient cells were grown to stationary phase and washed twice in the same media without any selective agents. Donor and recipient cells were mixed at a ratio of 3:1 and spotted onto a BHI agar plate supplemented with 10 mM $MgCl_2$. After incubation at 37 °C for ~7 days, lawn of cells were scraped off the agar plate using a toothpick and suspended homogenously in media. Serial dilutions of this cell suspension were spread onto BHI agar plates supplemented with 40 µg/ml streptomycin to select for recipient cells, and undiluted suspension was spread onto plates supplemented with 40 µg/ml streptomycin and 200 µg/ml kanamycin to select for transconjugants. Conjugation frequency was calculated as the number of transconjugant/recipient cell.

## 2.11 Polymerase chain reaction (PCR)

### 2.11.1 Conventional PCR

PCR was performed with the BioRad MJ Mini Personal Thermocycler. Primers were designed using the FastPCR software (PrimerDigital) and were purchased from Inqaba Biotec. Reaction mixtures were prepared in a final volume of 50 µl or 20 µl and usually contained each dNTP at a concentration of 0.2 mM and each primer at 0.5 µM. For plasmid and chromosomal templates, ~1 ng and ~100 ng of DNA, respectively, were added to the reaction mixture. *Taq* DNA polymerase (Fermentas) was used for reactions that did not require subsequent cloning of amplified DNA. When cloning was required, *Pfu* DNA polymerase (Fermentas) was used. For *Taq* DNA polymerase, reaction buffer containing $(NH_4)_2SO_4$ was added, and $MgCl_2$ was added to a final concentration of 2 mM. For *Pfu* DNA polymerase, only buffer containing $MgSO_4$ was used. For 50 µl and 20 µl reactions, 1.25 U and 0.5 U of polymerase, respectively, was added.

Cycling parameters were as follows: initial denaturation at 95 °C for 1-3 min, followed by 25-30 cycles of denaturation at 95 °C for 0.5-2 min, annealing at 55-72 °C for 2 min, extension at 72 °C for 0.5 min/kb for *Taq* polymerase and 2 min/kb for *Pfu* polymerase, and then final extension at 72 °C for 5-15 min.

### 2.11.2 Colony PCR of *Nocardia*

*Nocardia* colony PCR was conducted at a final reaction volume of 20 µl. Instead of any template DNA, a visible amount (1-3 mm in diameter) of a *Nocardia* single colony was transferred into the reaction mixture using a clean toothpick immediately prior to starting of the reaction. Other conditions and cycling parameters were the same as conventional PCR.

### 2.11.3 Site-directed mutagenesis (inverse PCR)

Phusion Site Directed Mutagenesis Kit (Finnzymes) was used. Primers incorporating desired mutations were manually designed. PCR was carried out as instructed by the manufacturer. Following PCR, the reaction mixture was subjected to ligation to circularize the amplified plasmid, which was then used to transform *E. coli*. Plasmids from transformants were screened by restriction digestion for the desired mutation. The mutation was then verified by sequencing.

### 2.11.4 Real-time PCR

An absolute quantification method as described by Lee *et al*. (2006) was attempted using the Roche LightCycler 2.0. Reaction kit used was LightCycler FastStart DNA Master SYBR Green I (Roche). Only cartridge-purified primers were used. To determine plasmid copy number, *orf2* of pYS1 was used as the target and *gyrA* of *N. aobensis* as the reference. To generate standard curves plotting plasmid copy number against threshold cycle, serial 10-fold dilutions, ranging from $1\times10^4$ to $1\times10^8$ copies/µl, of a plasmid containing both genes (pYS1R1Gyr) was used as the template. PCR on *N. aobensis* total DNA was then conducted, after which copy number of the target was compared to that of the reference to calculate plasmid copy number.

## 2.12 Southern analysis

### 2.12.1 Southern blot

Following agarose gel electrophoresis of DNA and photographing of the gel, the DNA was partially depurinated by soaking the gel in 0.25 M HCl for 15 min with gentle shaking. The solution was removed, and the gel was rinsed briefly twice in de-ionized $H_2O$. The DNA was then denatured by soaking the gel in 0.5 N NaOH for 30 min with gentle shaking. The DNA was transferred onto a Hybond-N 0.45 micron nylon membrane (Amersham) using the BioRad Vacuum Blotter Model 785 in 10× SSC for 90 min at 5 inches Hg. The membrane was then soaked in 2× SSC for 5 min, air-dried, and baked between two sheets of filter paper at 80 °C for 30 min. The blotted membrane was now ready for hybridization.

### 2.12.2 Construction of probe

Double-stranded, biotin-labelled probe was constructed by PCR. In a reaction using plasmid pYS1 as the template and PYS1Rep54F1 and PYS1Rep14R1 as the primers, biotin was incorporated into the amplicon by adding each dNTP and biotin-11-dUTP (Fermentas) to the following final concentration: dATP, 200 µM; dCTP, 200 µM; dGTP, 200 µM; dTTP, 150 µM; biotin-11-dUTP, 50 µM. The amplicon was salt-ethanol precipitated and resuspended in TE.

Incorporation of the label was confirmed by performing a dot blot. The probe was denatured by boiling for 5 min and immediately chilled on ice for 10 min. Ten-fold

serial dilutions, ranging from 1 ng to 10 fg, was spotted onto a nylon membrane (no pre-treatment necessary). The membrane was baked at 80 °C for 30 min, and the biotin signal was detected as written below. Labelling efficiency was considered satisfactory if the 100 fg spot could be viewed with the naked eye.

### 2.12.3 Southern hybridization

The blotted membrane was submerged in 5× SSC for 2 min and slipped into a hybridization bag. Hybridization solution (20 ml per 100 $cm^2$ of membrane) was added into the bag, which was sealed with a heat sealer. The membrane was prehybridized by incubation at 68 °C for 1-2 hrs by submerging the bag into a water bath. The solution inside the bag was redistributed occasionally by gentle inversion. During this incubation, the probe was denatured by boiling for 5 min and immediately chilled in an ice-water slurry for 10 min. Following incubation, the hybridization solution was discarded completely from the bag and replaced with new hybridization solution, to which the denatured probe (0.6 µg/20 ml hybridization solution) was added. The bag was sealed and incubated at 68 °C for 6 hrs. The solution inside the bag was redistributed occasionally by gentle inversion. Following incubation, the hybridization solution was discarded (but stored at -20 °C for repeated future usage), and the membrane was washed twice for 5 min each in 2× SSC and 0.1% SDS at room temperature, and also twice for 15 min each in 0.1× SSC and 0.1% SDS at 68 °C. The hybridized membrane was air-dried and was now ready for detection.

### 2.12.4 Detection of signal

Biotin Chromogenic Detection Kit (Fermentas) was used. Using streptavidin, with strong affinity for biotin, conjugated to alkaline phosphatase which produces an insoluble blue precipitate upon cleavage of a chromogenic substrate BCIP-T, bands on the hybridized membrane could be detected with the naked eye.

## 2.13 Chromatography

### 2.13.1 Thin layer chromatography (TLC)

To prepare nucleotides for separation by TLC, 10-20 µg of phage DNA was denatured by boiling for 10 min and immediately chilling in an ice-water slurry for 10 min. The denatured DNA was treated with ~0.15 units of nuclease P1 (Sigma) at

50 °C for 2 hrs. Nucleoside triphosphates were spotted onto cellulose plates (Merck) and run using a solvent containing 3 M ammonium sulphate:1 M sodium acetate:isopropanol (40:9:1).

### 2.13.2 High performance liquid chromatography (HPLC)

To prepare nucleosides for separation by HPLC, phage DNA was denatured by boiling for 10 min and immediately chilling in an ice-water slurry for 10 min. The denatured DNA was treated with nuclease P1 (Sigma) at 50 °C for 2 hrs, followed by treatment with bacterial alkaline phosphatase (Sigma) at 37 °C for 1 hr. HPLC was conducted through a Kanto Chemical LiChroCART 150-4.6 LiChrospher 100 RP-18 (5μm) column using a Gulliver JASCO PU-987 pump with a Gulliver JASCO HG-980-31 reader. Solvent contained acetonitrile:20 mM ammonium dihydrogen phosphate (1:20). A260 readings were measured for each nucleoside as it was passed through the column.

## 2.14 DNA sequencing and analysis

All DNA sequencing was conducted by Inqaba Biotec. Sequences were analyzed using the following programs and databases.

Finch TV (www.geospiza.com/Products/finchtv.shtml)
Program for viewing chromatograms.

Genbank (www.ncbi.nlm.nih.gov)
Database of publicly available nucleic acid and protein sequences.

BLAST (www.ncbi.nlm.nih.gov/BLAST/)
Program for finding nucleic acid or protein sequence similarities to what is available in the public database.

Glimmer (www.ncbi.nlm.nih.gov/genomes/MICROBES/glimmer_3.cgi)
Program for gene prediction in bacterial DNA.

GeneMark (exon.biology.gatech.edu/)
Program for gene prediction in bacterial DNA.

FGENESB (linux1.softberry.com/berry.phtml)
Program for gene prediction in bacterial DNA.

BPROM (http://linux1.softberry.com/berry.phtml)
Program for prediction of bacterial promoters.

NEBCutter (tools.neb.com/NEBcutter2/)
Program for creating restriction maps.

Merger (emboss.bioinformatics.nl/cgi-bin/emboss/)
Program for merging two overlapping sequences.

Needle (http://emboss.bioinformatics.nl/cgi-bin/emboss/)
Program for global pairwise alignment.

Edialign (http://emboss.bioinformatics.nl/cgi-bin/emboss/)
Program for multiple alignment.

FastPCR (en.bio-soft.net/pcr/FastPCR.html)
Program for primer design and analysis.

Vector NTi (www.invitrogen.com)
Program for various general sequence analysis.

Mfold (dinamelt.bioinfo.rpi.edu/)
Program for prediction of nucleic acid secondary structure.

### 2.14.1 Accession numbers

DNA sequences of pYS1 and pYS1R1 have been submitted to Genbank. Their accession numbers are GU066867 and HQ585611, respectively.

# 3. RESULTS

## 3.1 Functional analysis of plasmid pYS1 from *Nocardia*

### 3.1.1 Detection of small plasmids in *Nocardia*

Thirty-one clinical strains (Table 2.2) of *Nocardia* were screened for small plasmids by gel electrophoresis of their total DNA. Eight plasmids in total were detected (Table 3.1), which is similar in frequency (26%) to previous reports of 31% (Provost *et al*., 1996), 20% (Qasem *et al*., 1999), and 33% (Kasweck *et al*., 1981) in other strains of *Nocardia*. The smallest plasmid was observed in *Nocardia aobensis* IFM 10795, with an estimated size of approximately 4 kb, was named pYS1 and sequenced (Genbank accession number GU066867).

Table 3.1. Detection of plasmids in *Nocardia*

| Species | Strain [a] | Estimated plasmid size (kb) [b] |
|---|---|---|
| *N. brasiliensis* | IFM 10745 | 17 |
| | | 25 |
| *N. abscessus* | IFM 10765 | 21 |
| *N. nova* | IFM 10797 | 23 |
| *N. farcinica* | IFM 10779 | 12 |
| *N. farcinica* | IFM 10757 | 15 |
| *N. aobensis* | IFM 10795 | 4 |
| *N. otitidiscaviarum* | IFM 10786 | >50 |

[a] Of the 31 screened strains, only those in which plasmids were detected are listed.
[b] Plasmid were assumed to be in their covalently-closed-circular form whose sizes were estimated from gel electrophoresis

### 3.1.2 Prediction of ORFs in pYS1

The nucleotide sequence of pYS1 showed it to be a circular molecule of 4326 bp with a G+C content of 65.1% containing three ORFs, designated *orf1*, *orf2*, and *orf3* (Fig. 3.1).

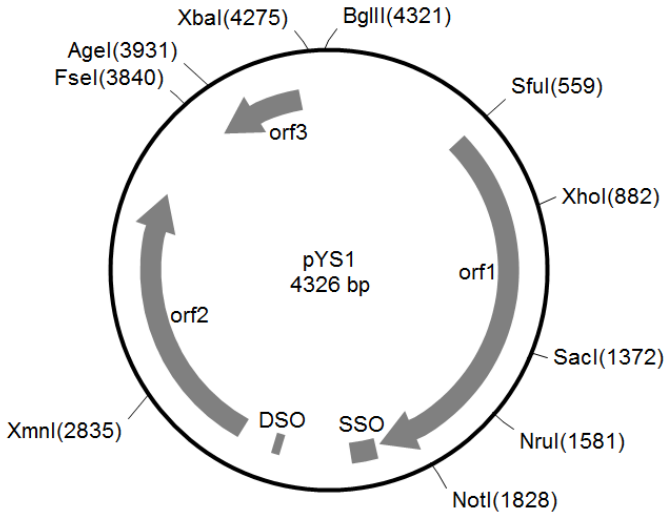



Figure 3.1. Schematic map of pYS1. Predicted ORFs and origins of replication are indicated by arrows and boxes, respectively. Some unique restriction sites and their positions are shown.

The exact translational start site for *orf1* was uncertain, as there were three potential start codons in frame at positions 447, 483, and 540, as predicted by Glimmer (v3.02) (Delcher et al., 1999), GeneMark (v2.4) (Lukashin and Borodovsky, 1998), and FGENESB (www.softberry.com). Although *Nocardia* ribosome binding sites have been hardly characterized, studies (Dale and Patiki, 1990; Dellagostin *et al*., 1995; Harth and Horwitz, 1997; Movahedzadeh *et al*., 1997; Mulder *et al*., 1997; Papavinasasundaram *et al*., 1997) have revealed those of mycobacteria to be purine-rich sequences, most of which have been detected 4-9 bp upstream of translational start sites. We therefore speculated position 540 to be the most probable candidate for *orf1* translational start due to the presence of the sequence AGGAGGA, resembling a mycobacterial ribosome binding site (Dellagostin *et al*., 1995), 4 bp upstream of it (Table 3.2). Consequently, *orf1* spans positions 540 to 1970 and encodes a protein of 477 amino acids. From sequence similarity, *orf1* is related to the transfer proteins of conjugative *Streptomyces* plasmids such as pIJ101 (Kendall *et al*., 1988) and pSN22 (Kataoka *et al*., 1994), as well as the FtsK/SpoIIIE family of proteins (Sherratt *at al*., 2010) from various bacteria, which are involved in DNA translocation during cell division and sporulation.

Orf2 of pYS1 also had multiple potential translational start sites, at positions 2073, 2313, 2505, and 2799 (Table 3.2). Of these, position 2505 has the closest (17 bp) upstream sequence of six consecutive purines, GAGAGG, resembling a mycobacterial ribosome binding site (Movahedzadeh *et al*., 1997;

Papavinasasundaram *et al*., 1997). Assuming this position to be its translational start site, *orf2* spans positions 2505 to 3548 and encodes a protein of 348 amino acids that is similar to those responsible for the initiation of rolling-circle replication of plasmids belonging to the pIJ101/pJV1 family (Ilyina and Koonin, 1992).

The protein encoded by *orf3* is composed of either 113 or 239 amino acids, as there are two potential translational start sites, at positions 4213 and 4114, each of which is complemented by a conceivable ribosome binding sequence AAGGAG (Harth and Horwitz, 1997) 7 bp away and AGGAAGG (Dale and Patiki, 1990) 4 bp away, respectively, in the upstream direction (Table 3.2). In addition to this uncertainty, Orf3 sequence had no significant similarity to anything in the database.

Table 3.2. Prediction of translational start sites by detecting potential ribosome-binding-sites

| ORF | Potential translational start position[a] | Translational end position[b] | Length in nucleotides | Length in amino acids | Start codon | Purine-rich sequence within 40 bp of start codon | Bp between purine-rich sequence and start codon | Has this sequence been reported as a ribosome-binding-site in *Mycobacterium*?[c] | Reference |
|---|---|---|---|---|---|---|---|---|---|
| orf1 | 447 | 1970 | 1524 | 508 | GTG | GAGGGAGGG | 0 | | |
| | | | | | | GGCAAG | 20 | | |
| | | | | | | GAAGA | 31 | | |
| | 483 | 1970 | 1488 | 496 | ATG | GAGGGAGGGG | 35 | | |
| | 540 | 1970 | 1431 | 477 | GTG | AGGAGAAA | 4 | yes (AGGAGA) | Dellagostin *et al*., 1995 |
| | | | | | | GGTAA | 34 | | |
| orf2 | 2073 | 3548 | 1476 | 492 | ATG | none | | | |
| | 2313 | 3548 | 1236 | 412 | GTG | AGCAGGAGCG | 0 | | |
| | | | | | | GAGCGGG | 14 | | |
| | | | | | | GGAGGG | 33 | | |
| | 2505 | 3548 | 1044 | 348 | GTG | GAGAGG | 17 | yes (GAGAGG) | Movahedzadeh *et al*., 1997 |
| | 2799 | 3548 | 750 | 250 | GTG | GAATGAGGCG | 9 | no | |
| | | | | | | AAGGAC | 33 | | |
| orf3 | 4213 | 3875 | 339 | 113 | ATG | GCAAGGAGATA | 4 | yes (AAGGAG) | Harth *et al*., 1997 |
| | 4114 | 3875 | 240 | 80 | GTG | AGGAAGGG | 3 | yes (AGGAAGG) | Dale and Patki, 1997 |
| | | | | | | GCAGAA | 17 | | |
| | | | | | | AAGACAG | 26 | | |

[a,b] Translational start and end positions were predicted by Glimmer (v3.02) (Delcher et al., 1999), GeneMark (v2.4) (Lukashin and Borodovsky, 1998), and FGENESB (www.softberry.com).

[c] Where a purine-rich sequence was detected 1-9 bp upstream of the potential start codon, literature search was conducted for previous reports describing such sequence as a ribosome-binding-site in mycobacteria. When such a report could not be found, then the next closest sequence was investigated.

### 3.1.3 pYS1 is a rolling-circle-type plasmid

#### 3.1.3.1 Orf2 is a rolling-circle-replication protein

The resemblance of Orf2 protein of pYS1 to replication initiation proteins of the pIJ101/pJV1 family (Ilyina and Koonin, 1992) suggested the likelihood of pYS1 replicating by a rolling-circle mechanism (Khan, 2005). Multiple alignment of pYS1 Orf2 amino acid sequence with those of Rep from related plasmids revealed the presence of five motifs in Orf2 that are usually conserved among proteins which initiate rolling-circle replication (Ilyina and Koonin, 1992; Nakashima and Tamura, 2004) (Fig. 3.2). Motif IV is a heavy-metal-binding motif rich in cysteine residues (Brasch and Cohen, 1995; Mendiola and de la Cruz, 1992; Nakashima and Tamura, 2004). Motif II has histidine residues which bind $Mg^{2+}$ and $Mn^{2+}$ (Ilyina and Koonin, 1992). Motif III has a tyrosine residue that nicks the leading strand to initiate replication and also forms a covalent link to the 5′ end of the nicked DNA (Ilyina and Koonin, 1992). Motif V, also known as the C-terminal motif, has tyrosine and/or glutamic acid residues which are believed to be involved in nucleophilic attack of the newly synthesized leading strand near the replication termination site (Billington *et al*., 1998: Fernandez-Gonzalez *et al*., 1994; Noirot-Gross *et al*., 1994).

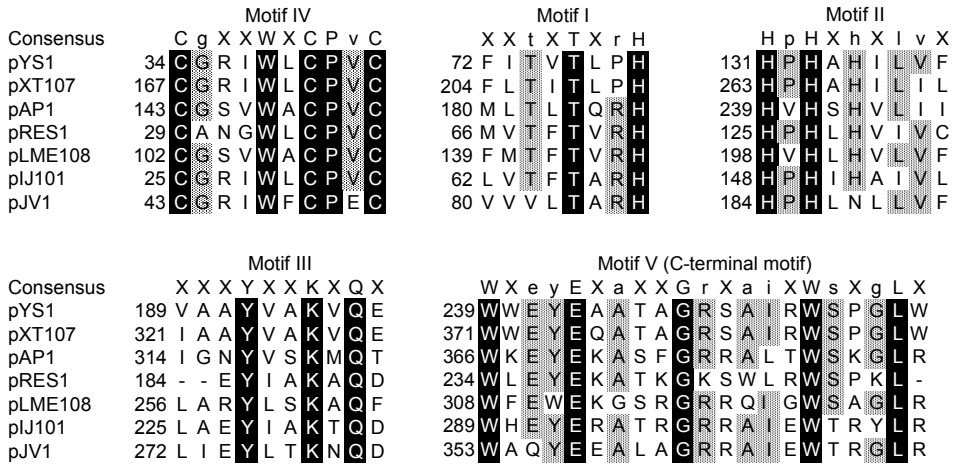


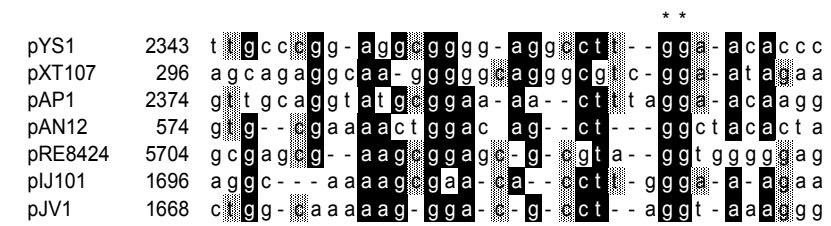

Figure 3.2. Multiple alignment of the pYS1 Orf2 sequence with Rep sequences from pXT107 (Xia *et al*., 2007), pAP1 (Billington *et al*., 1998), pRES1 (Ishikawa *et al*., 1996), pLME108 (Brede *et al*., 2005), pIJ101 (Kendall and Cohen, 1988), and pJV1 (Servin-Gonzalez *et al*., 1995). Residues with 100% identity are indicated with a black

background and those with >70% identity with grey. Numbers indicate positions of the first amino acids shown.

### 3.1.3.2 Prediction of the double- and single-stranded origins of replication

In addition to the five motifs in their replication proteins, rolling-circle plasmids are characterized by the presence of double- and single-stranded origins (DSO and SSO) of replication. These *cis*-acting elements were indeed predicted in the pYS1 sequence, in the intergenic region upstream of *orf2* on the same strand (Fig. 3.1). As DSOs are usually conserved among plasmids of the same family (Khan, 2005), multiple alignment of the pYS1 sequence with DSO sequences from related plasmids (Fig. 3.3A) revealed their potential counterpart in pYS1, which contained the GG dinucleotide shown in pIJ101 (Servin-Gonzalez, 1993), pJV1 (Servin-Gonzalez, 1993), and pSN22 (Suzuki *et al*., 1997) to be the nick site for Rep to initiate leading strand replication. SSOs, on the other hand, are not as well conserved but are characterized by inverted repeats that form extensive secondary structures, occurring in proximity to the DSOs (Khan, 2005). The pYS1 SSO secondary structure predicted by the mfold program (version 3.2) (Zuker, 2003) formed a hairpin with two terminal loops (Fig. 3.3B), one of which contained a sequence highly similar to a conserved CS-6 sequence TAGCGA (Khan, 2004; Kramer *et al*., 1997), shown to act as a terminator of RNA primer synthesis during lagging strand replication (Kramer *et al*. 1998a; Kramer *et al*., 1998b).

A



B

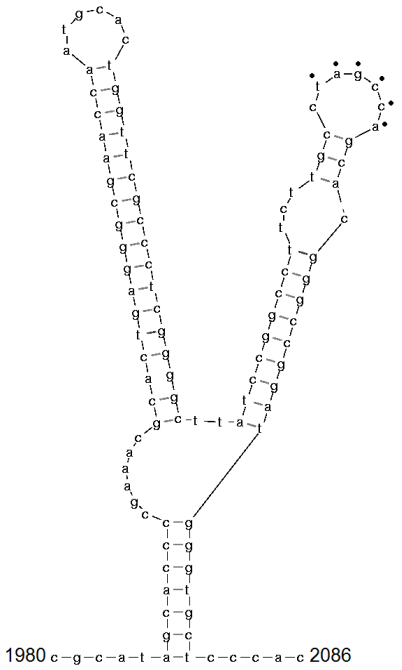



Figure 3.3. DSO and SSO of pYS1. (A) Multiple alignment of the predicted pYS1 DSO with those from related plasmids. The conserved GG dinucleotide, shown to be the nicking site for Rep in pIJ101 (Servin-Gonzalez, 1993), pJV1 (Servin-Gonzalez, 1993), and pSN22 (Suzuki *et al*., 1997), is indicated by the stars. Numbers indicate positions of the first nucleotides shown. Nucleotides with >70% identity are indicated with a black background and those with 70%-60% identity with grey. (B) Secondary structure of the SSO as predicted by the mfold program (version 3.2) (Zuker, 2003). Dots indicate a six base-pair sequence that is highly similar to a conserved CS-6 sequence TAGCGA in the SSO of rolling-circle plasmids (Kramer *et al*., 1997; Kramer *et al*., 1998a; Kramer *et al*., 1998b), which serves as the terminator of RNA primer synthesis for lagging strand replication.

### 3.1.3.3 Detection of the single-stranded intermediate

To obtain experimental evidence on the mode of replication of pYS1, a blot of total DNA from *N. aobensis* IFM 10795 was exposed to a double-stranded 1400 bp biotinylated probe that could hybridize to a sequence within *orf2* of the plasmid. A faint band which migrated faster than the double stranded plasmid was observed, which was shown to be a single-stranded form as the band disappeared upon treatment with S1 nuclease, whose activity is specific to single-stranded nucleic acids, prior to gel electrophoresis (Fig. 3.4). The presence of single-stranded intermediates being the hallmark of rolling-circle replication (Khan, 2005), its detection in pYS1, along with the five motifs in Orf2, and

the prediction of DSO and SSO all together serve as strong evidence that the plasmid employs this mode of replication.

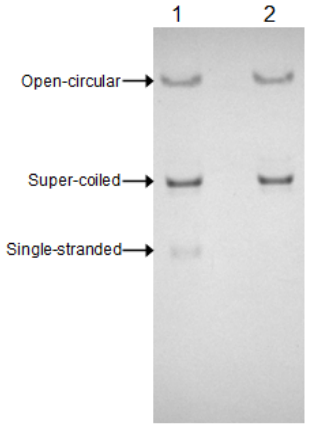



Figure 3.4. Detection of single-stranded intermediate of pYS1 by Southern analysis of total DNA from *N. aobensis* IFM 10795. Lane 1, total DNA untreated; lane 2, total DNA treated with S1 nuclease.

### 3.1.4 Predicted DSO and SSO are essential in pYS1 replication

#### 3.1.4.1 DSO mutagenesis

To investigate the importance of the predicted DSO and SSO, a 2 kb region of pYS1 incorporating these elements along with *orf2* was cloned into the *Nhe*I site of *E. coli* vector pK18 (Pridmore, 1987), yielding pYS1R1. This hybrid plasmid was then transformed into *N. nova*, followed by verification of plasmid maintenance by Southern analysis of DNA extracted from the transformants and also by *E. coli* transformation using the extracted DNA. Plasmid pYS1R1 was able to transform *N. nova* at a relatively high efficiency (3e6 cfu/µg DNA) and replicate in this host (Fig. 3.5), suggesting that all elements essential for replication in *Nocardia* are within the 2 kb region. The sequence containing the GG dinucleotide in the predicted DSO in pYS1R1 was then altered from TTGGAA to AGATCT by *in-vitro* mutagenesis. This plasmid, pYS1R1D1, was no longer able to be maintained in *N. nova* (Fig. 3.5), showing that the DSO, as expected, is essential in the replication process.

### 3.1.4.2 SSO mutagenesis and deletions

Deletion of the SSO in most rolling-circle plasmids causes instability in replication, leading to an accumulation of single-stranded intermediates (Khan, 2005). Plasmid replication is not all together abolished, however, probably due to the presence of secondary signals for single-strand to double-strand conversion (Khan, 2005). To investigate whether this situation applied to pYS1R1, the conserved sequence in the second terminal loop of the predicted SSO was modified from TAGCCA to TGCAGA, yielding pYS1R1S1. Replication of this construct in *Nocardia*, however, was not affected (Fig. 3.5), and no accumulation of single-stranded intermediate of the plasmid could be detected as judged by band intensity upon Southern analysis (Fig. 3.6). No phenotypic difference could be observed with pYS1R1S1 compared to pYS1R1, suggesting that the mutated region is not important for replication. To further investigate whether the predicted SSO region is of any importance, plasmids with a series of deletions to this region were constructed (Fig. 3.5). This was achieved by amplifying different lengths of pYS1 DNA and cloning them into the *Nhe*I site of pK18, which were then used to transform *N. nova*. Initially, constructs pYS1R6 and pYS1R7, both fully missing the SSO, were checked for plasmid maintenance in *Nocardia*. Both pYS1R6 and pYS1R7 transformed *N. nova* at a low efficiency (6e2 and 1e2 cfu/μg DNA, respectively) (Fig. 3.5), and intact plasmids could not be isolated from these transformants, suggesting that the predicted SSO is essential in plasmid maintenance. No accumulation of single-stranded intermediates could be detected in these transformants either (Fig. 3.6). The most probable explanation for these transformants was that the plasmid had integrated into the host chromosome, giving them resistance to kanamycin on the selective plates. To investigate the extent of deletion the SSO was able to tolerate to allow plasmid replication, pYS1R3 (missing the lower stem of the hairpin), pYS1R4 (containing only half of the second stem-loop), and pYS1R5 (missing the entire stem-loop but less-so than pYS1R6 and pYS1R7) were tested (Fig. 3.5). When the lower stem of the hairpin was missing (pYS1R3), high transformation efficiency of *N. nova* and plasmid maintenance were both sustained, showing that the missing DNA was expendable for these purposes. In pYS1R4, however, high transformation efficiency of *N. nova* and plasmid maintenance were both lost. Similar observations were made for pYS1R5. No accumulation of single-

stranded intermediates was observed in any of these transformants (Fig. 3.6). These results confirmed that the predicted SSO is indeed essential for pYS1 replication and that it can only tolerate minor changes to its structure as there seems to be no secondary signals that can compensate for its deficiency.

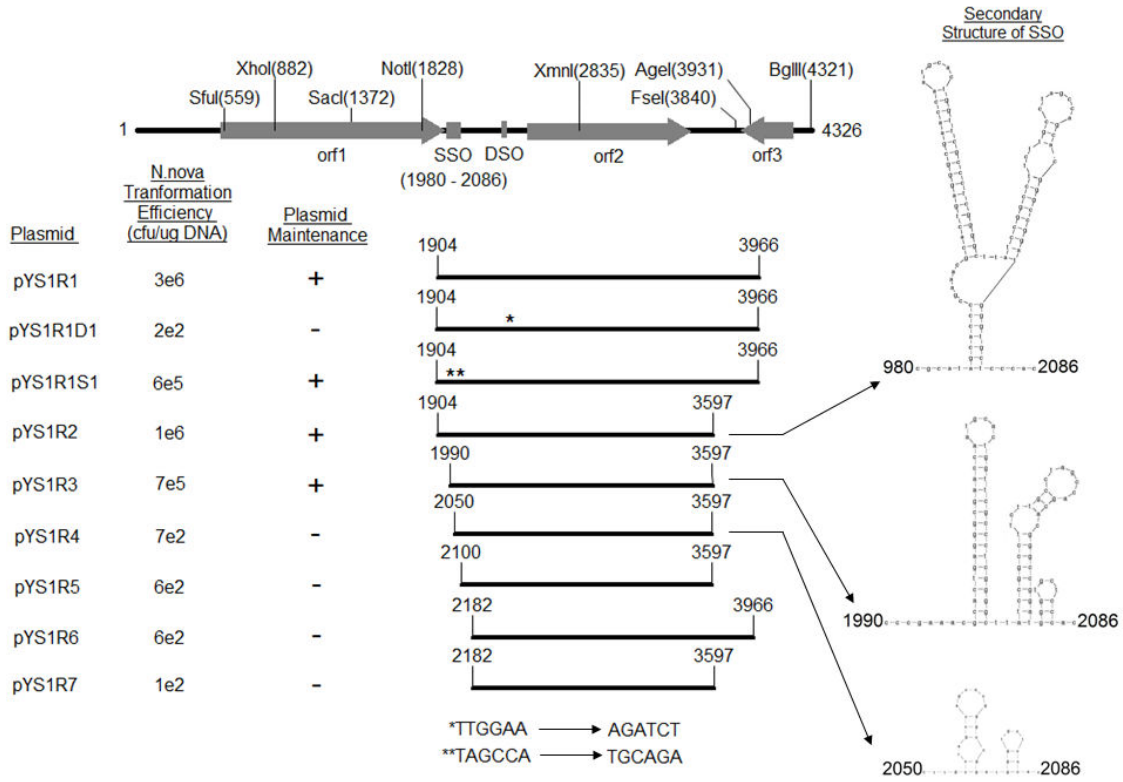



Figure 3.5. Effect of DSO and SSO mutations and SSO deletions on pYS1 replication. Plasmids were constructed by amplifying the indicated regions of pYS1 using *Nhe*I-site-containing primers and cloning them into the *Nhe*I site of pK18 (Pridmore, 1987). Names of constructed plasmids and their efficiencies to transform *N. nova* are indicated in the first and the second columns, respectively. Presence or absence of plasmids from these *N. nova* transformants is indicated by +/- in the third column. In pYS1R1S1 and pYS1R1D1, sequences in the second exposed loop of the predicted SSO and the potential nicking site in the predicted DSO, respectively, were modified as indicated by the stars. For the SSO deletion experiment, cloned regions differed in the extent to which the SSO was included. Position of the SSO is indicated in parentheses on the linear map of pYS1 at the top. For pYS1R2, pYS1R3, and pYS1R4, secondary structures of the SSO region predicted by mfold (version 3.2) (Zuker, 2003) are shown in the last column. Plasmids pYS1R2 had the SSO in its intact form, whereas pYS1R3 had the lower part of the stem missing and therefore a shorter hairpin. Plasmid pYS1R4 had most of the SSO missing.

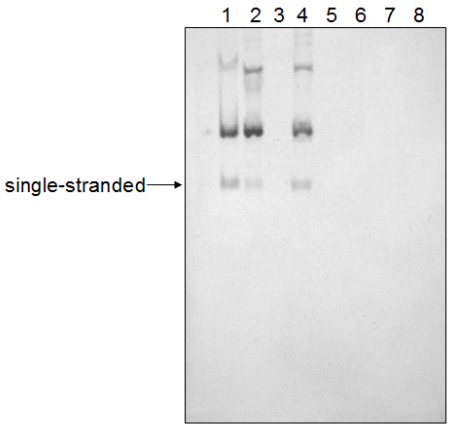



Figure 3.6. Southern analysis of constructs with mutated DSO or SSO, or deleted SSO. Lane 1, pYS1R1; lane 2, pYS1R1S1; lane 3, pYS1R1D1; lane 4, pYS1R3; lane 5, pYS1R4; lane 6, pYS1R5; lane 7, pYS1R6; lane 8, pYS1R7.

### 3.1.5 *Nocardia-E. coli* shuttle vector pYS1R1

The minimal replicon of pYS1 was revealed in the above experiment to be within a 2 kb region containing *orf2*, DSO, and SSO. As pYS1R1 (Fig. 3.5 and Fig. 3.7) included all three elements and was stably maintained in both *Nocardia* and *E. coli*, this plasmid serves as a *Nocardia* – *E. coli* shuttle vector. Further manipulations, such as increasing the number of unique cloning sites, may be made to improve its utility.

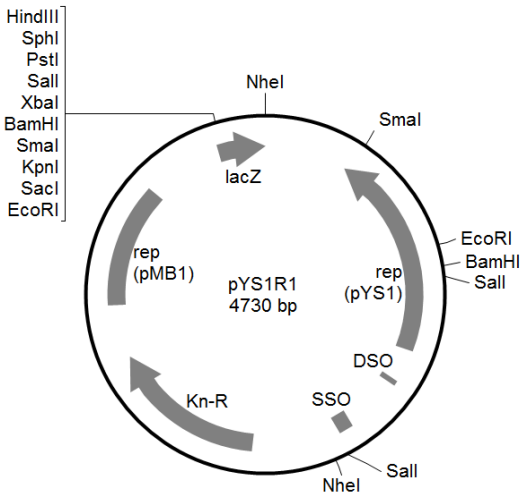



Figure 3.7. Schematic map of pYS1R1, which was constructed by ligating a 2063 bp region of pYS1 containing the *rep* (*orf2*), DSO, and SSO into the *Nhe*I site of *E. coli* vector pK18.

### 3.1.6 pYS1 replicon is compatible with pNV18/19

Although a few *Nocardia-E. coli* shuttle vectors have been described previously (Yao *et al*., 1994; Liu *et al*., 2000), plasmid pNV18/19 (Chiba *et al*., 2007), to our knowledge, has been the most useful. To check whether the pYS1 replicon is compatible with pNV18/19, which uses a replicon from a mycobacterial plasmid pAL5000 (Stolt and Stoker, 1998), *N. aobensis* IFM 10795 was transformed with pNV19. Stable coexistence of the two plasmids was confirmed, as a band corresponding to each plasmid was detected from a transformant by gel electrophoresis (Fig. 3.8). This could be a useful feature of the pYS1 replicon in experiments requiring introduction of two separate plasmids into *Nocardia*.

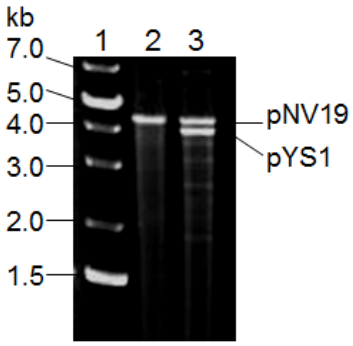



Figure 3.8. Gel electrophoresis of a plasmid extraction from *N. aobensis* IFM 10795 transformed with pNV19 (Chiba *et al*., 2007). Lane 1, molecular weight marker; lane2, *Xba*I digest, which produced co-migrating bands of two linearized plasmids; lane 3, *Pst*I digest, which linearized pNV19 (upper band) and reduced the size of the major pYS1 band (lower band).

### 3.1.7 pYS1 copy number

Real-time PCR, due to its sensitivity, is becoming an increasingly popular method for determination of plasmid copy number (Lee *et al*., 2006; Werbowy *et al*., 2009). With this method, copy number is calculated as a ratio of frequency of a certain DNA sequence (target) in a plasmid to that which occurs as a single copy (reference) in chromosomal DNA. Initially, real-time PCR was attempted by using *rep* as the target gene and *gyrA*, encoding for the α-subunit of the gyrase, as the reference gene. Primers for *gyrA* were designed according to its sequence in the *N. farcinica* genome (*Ishikawa et al*., 2004), as the *N. aobensis* genome has not been sequenced. However, its amplification from the *N.*

*aobensis* genome caused several unspecific bands to appear following conventional PCR, which translated to numerous unspecific peaks during melting-curve analysis in real-time PCR (data not shown). An alternative method that does not rely on the host genome sequence, therefore, was selected.

By comparing plasmid and chromosomal DNA peak area values (Pushnova *et al*., 2000) after gel electrophoresis of total DNA (Fig. 3.9 and Table 3.3) from *N. aobensis* IFM 10795 at late logarithmic phase, copy number of pYS1 was estimated under the assumption that the genome size of the host is the same as that for *N. farcinica* IFM 10152 (6,292,345 bp) (Ishikawa *et al*., 2004). Copy number values obtained from experiments conducted in triplicate were 13, 7, and 10. Consequently, the average copy number was 10, with a standard deviation of 3.

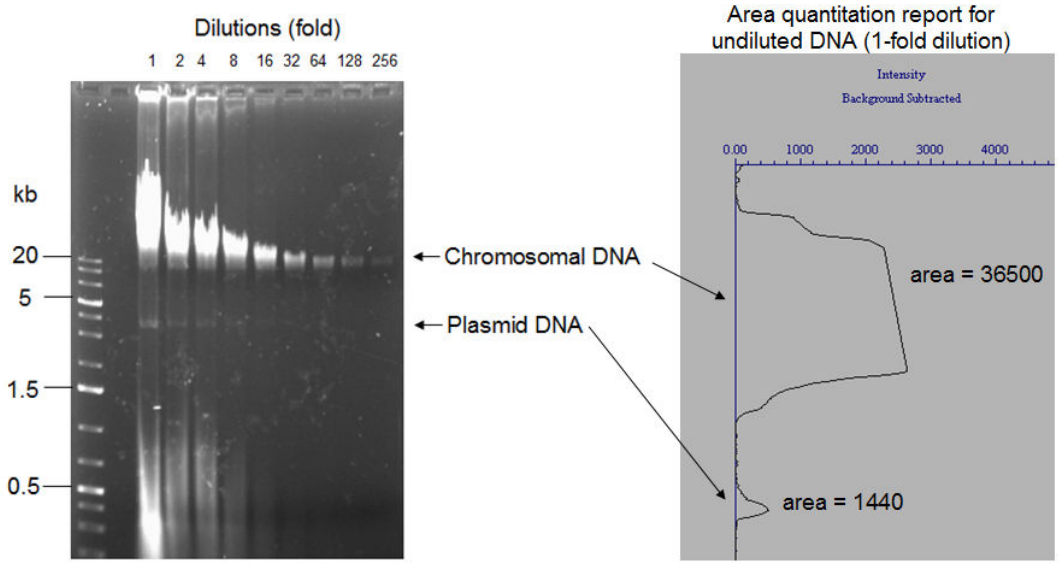



Figure 3.9. Gel electrophoresis of serial dilutions of total DNA from *N. aobensis* IFM 10795 (left) and a representative area quantitation report (right). Dilutions of DNA run on each lane are indicated above the gel. Chromosomal and plasmid bands are indicated. Area quantitation report was generated using Quantity One software (version 4.6.7) (Bio-Rad) after subtraction of background fluorescence. For estimation of plasmid copy number, such a report was generated for every single lane of the gel.

Table 3.3. Plasmid copy number calculation form

| Dilution (fold) | 1 | 2 | 4 | 8 | 16 | 32 | 64 | 128 | 256 |
|---|---|---|---|---|---|---|---|---|---|
| Chromosomal DNA area [a] | 36500 | 29400 | 28000 | 19300 | 113300 | 5490 | 2580 | 964 | 528 |
| Adjusted area [b] | 36500 | 58800 | 112000 | 154400 | 212800 | 175680 | 165120 | 123392 | 135168 |
| % Difference [c] | | 61% | 90% | 38% | 38% | -17% | -6% | -25% | 9.5% |
| Data used [d] | | | | | | 175680 | 165120 | | |
| Chromosomal Average | 170400 | | | | | | | | |
| Plasmid DNA area [e] | 1440 | 335 | 371 | 185 | 100 | | | | |
| Adjusted area [f] | 1440 | 670 | 1484 | 1480 | 1600 | | | | |
| % Difference [g] | | -53% | 121% | -0.3% | 8.1% | | | | |
| Data used [h] | | | | 1480 | 1600 | | | | |
| Plasmid Average | 1540 | | | | | | | | |
| Plasmid copy number [i] | 13 | | | | | | | | |

[a,e] For each dilution, chromosomal and plasmid peak area values were obtained using Quantity One software (v4.6.7) (Bio-Rad).

[b,f] Adjusted areas were calculated by multiplying the area value by the appropriate dilution factor.

[c,g] % difference was calculated by the following formula:

$$\%\ \text{difference} = \frac{\text{adjusted area} - \text{adjusted previous area}}{\text{adjusted previous area}} \times 100$$

[d,h] Only where the % difference was between -20% and 20%, the adjusted area values were used to calculate the average value, as they best reflected the amount of DNA present.

[i] Plasmid copy number was calculated using the following formula:

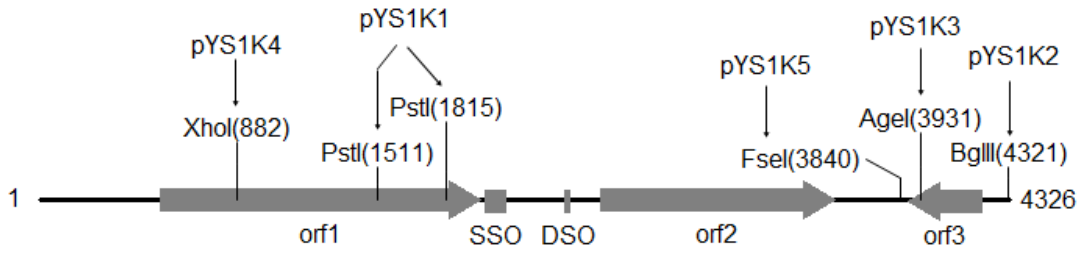
$$\text{plasmid copy number per genome} = \frac{\text{size of chromosomal DNA (bp)} \times \text{average plasmid DNA area}}{\text{size of plasmid DNA (bp)} \times \text{average chromosomal DNA area}}$$

### 3.1.8 Construction of insertion mutants of *orf1*, *orf3*, and the surrounding regions

To investigate the functions of *orf1*, *orf3*, and the flanking intergenic regions, insertion mutants of pYS1 were constructed where an aminoglycoside-3′phosphotransferase (*aph*) gene was separately inserted into five different sites (Fig. 3.10A). Two of these were in *orf1* (for pYS1K1 and pYS1K4), one in *orf3* (for pYS1K3), one in the intergenic region between *orf2* and *orf3* (for pYS1K5), and one in the intergenic region between *orf3* and *orf1* (for pYS1K2). For pYS1K1, the 304 bp *Pst*I fragment of pYS1 was replaced with an *aph* gene. Construction of pYS1K5 involved insertion at the *Fse*I site. This site, however, was resistant to digestion by *Fse*I, an enzyme produced by *Frankia* species that recognizes the sequence GGCCGGCC. When a region containing this restriction site was amplified by PCR, however, the amplicon was successfully digested by *Fse*I, suggesting

presence of modification on pYS1 DNA preventing enzyme activity. Although this restriction enzyme is known to be blocked by overlapping Dcm methylation (New England Biolabs, 2009), which recognizes the sequence CCWGG, the closest such sequence was found 60 bp away from the restriction site. There is no previous report of Dcm methylation blocking restriction at such a distance. It is therefore possible that the blockage is due to another modification, which has not been characterized, that recognizes a sequence within or overlapping the *Fse*I site. To construct pYS1K5 (Fig. 3.10B), 2903 bp *Sac*I-*Xba*I fragment containing the *Fse*I site was first cloned into pUC18 in *E. coli*, thereby yielding pYSA1 (Fig. 3.10B). Only when this construct was cloned in *E. coli* strain of GM2929, incapable of Dam and Dcm methylation, was the site digested by *Fse*I. An *aph* gene was then inserted into this site, yielding pYSA2 (Fig. 3.10B), followed by removal of the region originating from pUC18 and ligation with the missing region of pYS1, the 1423 bp *Sac*I-*Xba*I fragment. Construction of pYS1K3 involved insertion of *aph* gene at the *Age*I site, which lies within the C-terminus of *orf3*. When this site was disrupted, the only plasmid obtainable was spontaneously deleted in *orf1* (see section 3.1.11.3).

A



B

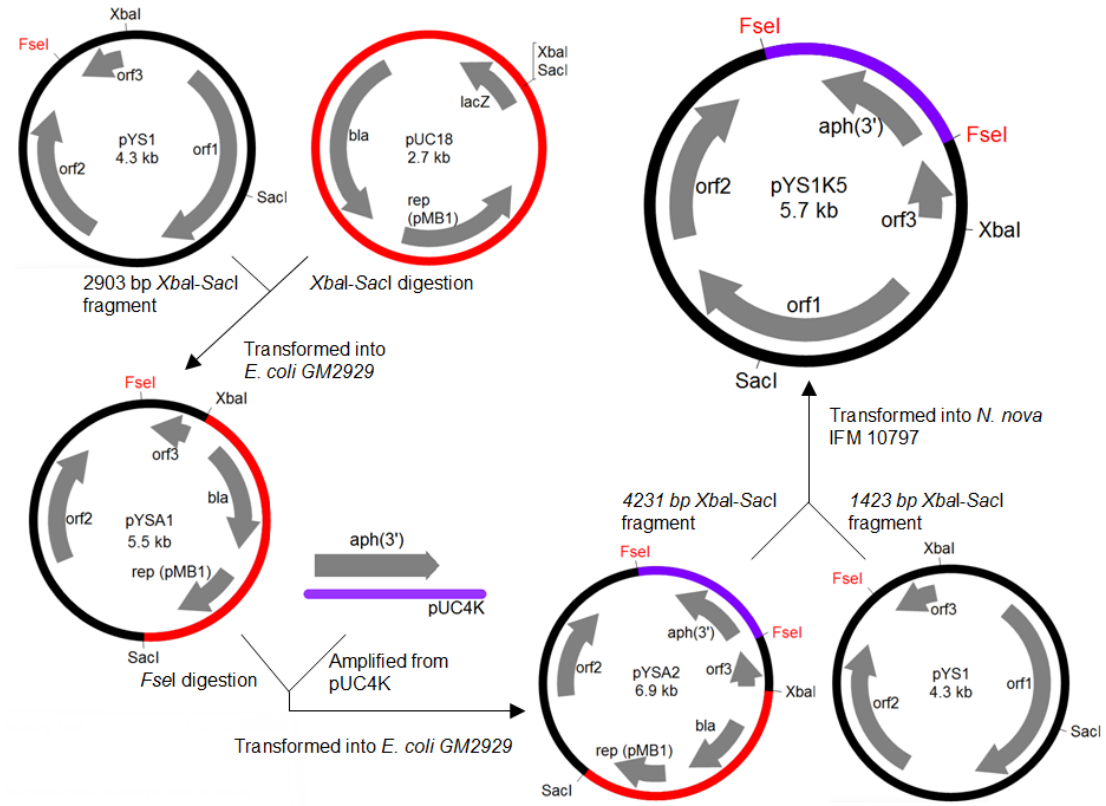



Figure 3.10. Construction of insertion mutants. (A) Locations of insertions. Restriction sites pointed by the arrows are those used for the construction of plasmids pYS1K1, pYS1K2, pYS1K3, pYS1K4, and pYS1K5. In case of pYS1K1, the 304 bp *Pst*I fragment was replaced with the *aph* gene. In all other cases, the *aph* gene was inserted into the respective restriction sites. Insertion into the *Age*I site, located within the C-terminus of *orf3*, correlated with spontaneous deletion of *orf1*. (B) Construction of pYS1K5. As the *Fse*I site was resistant to digestion by *Fse*I, the region containing this site was first cloned into pUC18 in *E. coli* GM2929, followed by insertion of the *aph* gene and ligation with the rest of DNA from pYS1.

### 3.1.9 pYS1 replicon is broad-host-range

Using pYS1K1 and pYS1K2, ten different species from four genera were successfully transformed to give a kanamycin resistance phenotype (Table 3.4). Plasmid of indistinguishable size could be detected in each of these transformants (data not shown). The replicon of pYS1 is therefore broad-host-range.

Table 3.4. Host range of the pYS1 replicon

| Species and Strain | Transformation Efficiency [a] (cfu/μg DNA) |
|---|---|
| *Nocardia aobensis* IFM 10795 | 3e2 |
| *Nocardia aobensis* IFM 10795c | 5e4 |
| *Nocardia nova* IFM 10797 | 4e5 |
| *Nocardia brasiliensis* IFM 10745 | 1e5 |
| *Nocardia farcinica* IFM 10757 | 1e5 |
| *Mycobacterium smegmatis* $mc^2155$ | 4e5 |
| *Mycobacterium parafortuitum* 490 | 4e4 |
| *Gordonia rubropertincta* ATCC 25593 | 1e3 |
| *Rhodococcus erythropolis* ATCC 4277 | 2e5 |
| *Rhodococcus rhodochrous* 01 | 5e6 |
| *Rhodococcus equi* ATCC 14887 | 1e5 |
| *Arthrobacter oxydans* C7 | < 3e1 |
| *Streptomyces lividans* TK23 | < 3e1 |
| *Escherichia coli* MM294 | < 3e1 |

[a] Transformation efficiency is the average between those for pYS1K1 and pYS1K2, as similar values were observed for both. Average values from three independent transformation experiments are listed. Where the transformation efficiency is < 3e1, no single transformant was observed using 33 ng of plasmid DNA.

### 3.1.10 Curing of *N. aobensis* IFM 10795

After 90 generations, *N. aobensis* IFM 10795 stably harboured pYS1 as judged by band intensity following gel electrophoresis of plasmid extractions from this strain (Figure 3.11).

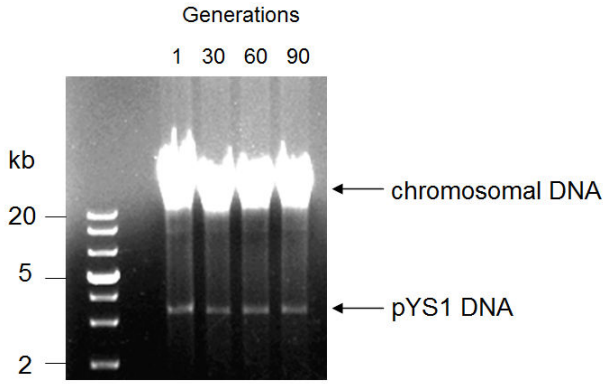



Figure 3.11. Gel electrophoresis of total DNA from *N. aobensis* IFM 10795 at generations 1, 30, 60, and 90.

To cure the strain of the plasmid, growth-retardatory concentrations of ethidium bromide (256 μg/ml) or mitomycin C (0.2 μg/ml) were first tested. These concentrations were empirically determined by growing the cells in series of concentrations. Following growth in media containing each of these potential curing agents, none of 20 screened

colonies had lost the plasmid. We therefore attempted using pYS1K2 as a competitor to pYS1 (Fig. 3.12). IFM 10795 was first transformed with pYS1K2, followed by growth of the transformant in the presence of kanamycin for 30 generations. Except for the presence of the *aph* gene in pYS1K2, both plasmids were identical and therefore should not have been compatible to one another. As kanamycin should have forced pYS1K2 to remain in the cell, pYS1 was lost either at the point of transformation or during growth in the presence of the antibiotic. The resulting cell, which then only harboured pYS1K2, was then grown in the absence of kanamycin for 30 generations, leading to loss of this plasmid as well. Of ten colonies screened, three had lost both pYS1 and pYS1K2, as no plasmid could be detected by both Southern hybridization and PCR (Fig. 3.13). Moreover, transformation efficiency of the cured strain, IFM 10795c, increased by two orders of magnitude as compared to the original strain when transformed with pYS1K1 or pYS1K2 (Table 3.4), confirming absence of a resident plasmid that would otherwise compete with the incoming plasmid. However, the relative ease at which pYS1K2 was lost from the cell in the absence of a selective marker should be noted, as this is in contradiction with the stability of pYS1. The site of *aph* gene insertion in pYS1K2 was the *Bgl*II site in the intergenic region between *orf3* and *orf1*, which may therefore contain an element contributing to pYS1 stability.

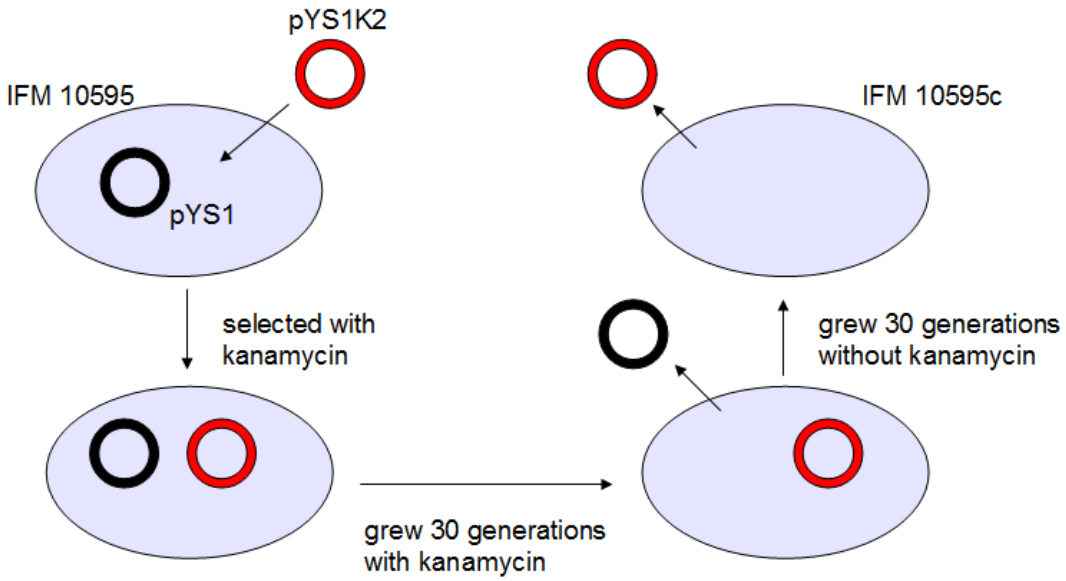



Figure 3.12. Process of curing using a competitor plasmid.

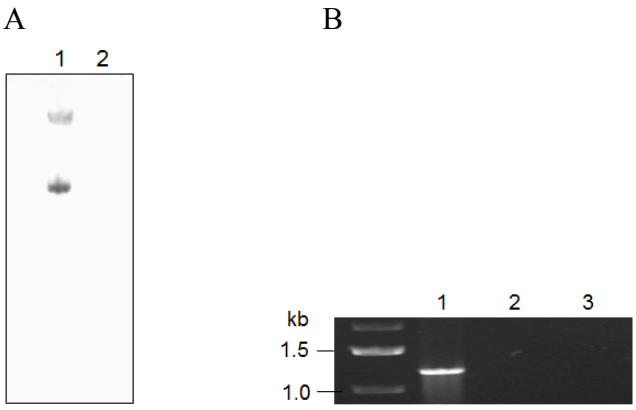


Figure 3.13. Verification of plasmid loss from the cured strain. (A) Southern analysis of total DNA using a biotinylated probe hybridizing to a 1.3 kb region of *orf2*. Lane 1, IFM 10795; lane 2, IFM 10795c. (B) Gel electrophoresis of colony PCR to amplify a 1.3 kb region of *orf2*. Lane 1, IFM 10795; lane 2, IFM 10795c; lane 3, no colony control.

### 3.1.11 Analysis of *orf3-orf1* pair

#### 3.1.11.1 *orf1* is involved in stability

When various species were transformed with pYS1K1 (disrupted in *orf1*) or pYS1K2 (disrupted in the *orf3-orf1* intergenic region), it took significantly longer for transformants of pYS1K1 to grow up than those of pYS1K2 (Table 3.5). This observation was made in eight out of the eleven transformed species, with the greatest difference in time observed for *N. farcinica* IFM 10757, which took 76% longer for pYS1K1 transformants to become visible on selective plates. However, both groups of transformants streaked onto fresh media grew at the same rate (data not shown), leading to a speculation that *orf1* may play a role in pYS1 stability only when the plasmid is establishing itself in new hosts.

Table 3.5. Differences in time taken for transformants of pYS1K1 and pYS1K2 to appear

| Species and Strain | Hrs taken for transformants to appear [b] | |
|---|---|---|
| | pYS1K1 (*orf1*-disrupted) | pYS1K2 (*orf1*-non-disrupted) |
| *N. aobensis* IFM 10795 | 116 | 116 |
| *N. aobensis* IFM 10795c | 112 | 112 |
| *N. nova* IFM 10797 | 67 | 67 |
| *N. brasiliensis* IFM 10745 | ND | ND |
| *N. farcinica* IFM 10757 | 37 | 21 |
| *Mycobacterium smegmatis* $mc^2155$ | 66 | 55 |
| *Mycobacterium parafortuitum* 490 | 181 | 118 |
| *Gordonia rubropertincta* ATCC 25593 | 57 | 42 |
| *Rhodococcus erythropolis* ATCC 4277 | 52 | 32 |
| *Rhodococcus rhodochrous* 01 | 46 | 38 |
| *Rhodococcus equi* ATCC 14887 | 70 | 43 |

[b] These columns show the number of hours taken for pYSK1 and pYS1K2 transformants to become visible on selective agar plates. Average values from three independent transformation experiments are listed.
ND, number of hours not determined.

### 3.1.11.2 Orf1 does not mediate pYS1 transfer

The similarity of the Orf1 sequence to those of plasmid transfer proteins such as Tra of pIJ101 (Kendall and Cohen, 1988) or TraB of pSN22 (Kosono *et al*., 1996) suggested the possibility of its involvement in conjugation. These sequences also fell into the same group as the FtsK/SpoIIIE family proteins (Sherratt *et al*., 2010), which are involved in DNA translocation during cell division and sporulation. A multiple alignment of Orf1 amino acid sequence with related sequences revealed the presence of Walker type A and B motifs (Walker *et al*., 1982) (Fig. 3.14), both of which have been shown in pSN22 to be essential in plasmid transfer (Kosono *et al*., 1996).

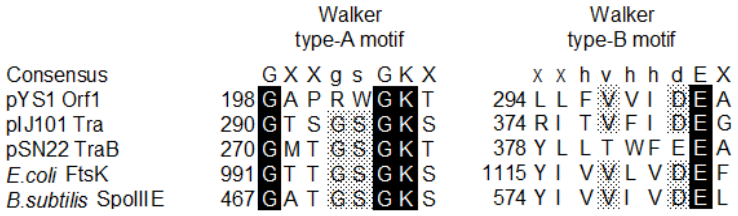



Figure 3.14. Multiple alignment of pYS1 Orf3 amino acid sequence with those from the transfer proteins of pIJ101 (Kendall and Cohen, 1988) and pSN22 (Kosono *et al*., 1996), FtsK from *E. coli* (Begg *et al*., 1995), and SpoIIIE from *B. subtilis* (Wu *et al*., 1995). Walker type-A and B motifs (Walker *et al*., 1982) are shown. Numbers indicate positions of the first amino acid residue shown. Residues with 100% identity are indicated with a black background whereas those with 80% in grey.

To investigate whether *orf1* played a role in pYS1 transfer, *N. aobensis* IFM 10795c transformed separately with the five insertion mutants of pYS1 were used as donors and *N. nova* IFM 10797 as recipients in a mating experiment. All insertion mutants were transferred at a similar frequency of ~5e-4 (transconjugant per recipient), suggesting lack of involvement of *orf1* in conjugation.

### 3.1.11.3 *orf3-orf1* intergenic region has promoter activity which is controlled by a transcriptional repressor encoded by *orf3*

As mentioned, the only plasmid that could be obtained following insertion of *aph* at the *Age*I site (within *orf3*) of pYS1, was one that was spontaneously deleted in *orf1* (Fig. 3.15), suggesting a lethal function of *orf1* that is suppressed by *orf3*. This observation was in line with the sequence similarity of Orf1 to Tra of pIJ101 (Kendall and Cohen, 1988) or TraB of pSN22 (Kataoka *et al*., 1994), which are lethal to the host unless negatively regulated by Gnt-R type transcriptional repressors KorA (pIJ101) or TraR (pSN22). Genes encoding these repressors occur just upstream to the lethal genes, separated by a short intergenic region, and are transcribed in the opposite orientation (Fig. 3.15). Although *orf3* bore no sequence similarity to these transcriptional repressors, its location and direction with respect to *orf1*, in addition to its potential involvement in "kil-override" (Stein *et al*., 1989), presented a similar situation as those *Streptomyces* plasmids.

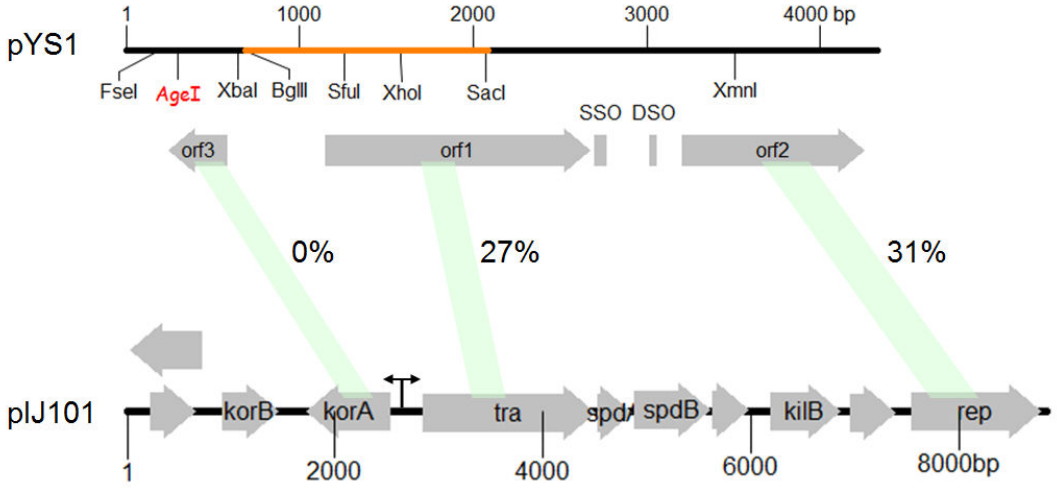
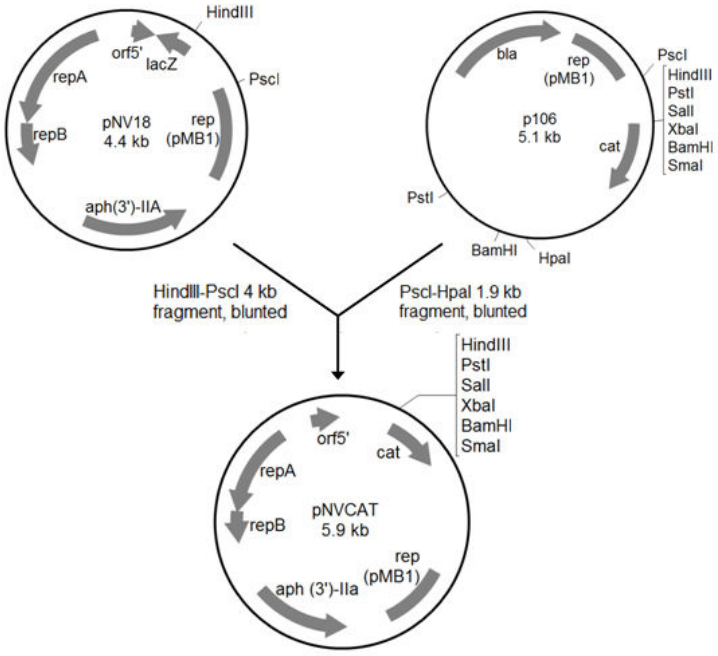



Figure 3.15. Similarity of pYS1 to pIJ101. In pYS1, insertion of *aph* into the *Age*I site correlated with spontaneous deletion of the region shown in orange. Sequence identities of proteins encoded by the respective genes in the two plasmids are indicated. The nucleotide positions of pYS1 on this map are different from those in previous figures, as *orf*3 and *orf*1 needed to be shown as a continuous unit. The two maps are not to scale.

To investigate whether *orf3* encoded a novel transcriptional repressor, a reporter assay vector pNVCAT (Fig. 3.16A) was constructed, which bears a multiple cloning site upstream of a promoter-less chloramphenicol acetyl-transferase (*cat*) gene in a pNV18 backbone. In a phenotypic CAT-assay, promoter activity of DNA inserted upstream of the *cat* gene was measured by observing for increase in MIC of chloramphenicol for *N. nova* IFM 10797 transformed with constructs depicted in Figure 3.16B. First, the intergenic region between *orf3* and *orf1* was inserted in both orientations. Two different lengths of this region were inserted – one between positions 4214 and 426 and another between positions 4115 and 539. For the latter, due to reason which could not be determined, the fragment could be cloned in *E. coli* only in the orientation allowing testing of promoter activity towards *orf1*, thus preventing testing of that towards *orf3*. Insertion of these DNA in *orf1* orientation resulted in increase in MIC for *N. nova* from 42 µg/ml (pNVCAT only) to 126 µg/ml (pNVCATP1) and 140 µg/ml (pNVCATP3) (Fig. 3.16B), suggesting promoter activity towards *orf1*. In the *orf3* orientation (pNVCATP2), however, no significant increase in MIC was observed. To investigate whether the presence of *orf3* reduced the promoter activity detected above, the open reading frame was inserted together with the intergenic region as a single unit upstream of the *cat* gene

(pNVCATP4 and pNVCATP5), such that the organization of *orf3* with respect to *cat* was identical to that of *orf3* with respect to *orf1* in pYS1 (Fig. 3.16B). The presence of *orf3* in these constructs led to a reduction in MIC to 76 µg/ml (pNVCATP4) and 77 µg/ml (pNVCATP5). It should be noted that these reduced values were higher than that for pNVCAT alone, suggesting some level of promoter activity. Because it was presumed that *orf3* suppressed *orf1* lethality, and its presence in the above constructs led to a reduction in promoter activity, it is rational to conclude that its encoded product acts as a transcriptional repressor that negatively regulates promoter activity detected in the *orf3*-*orf1* intergenic region.

A



B

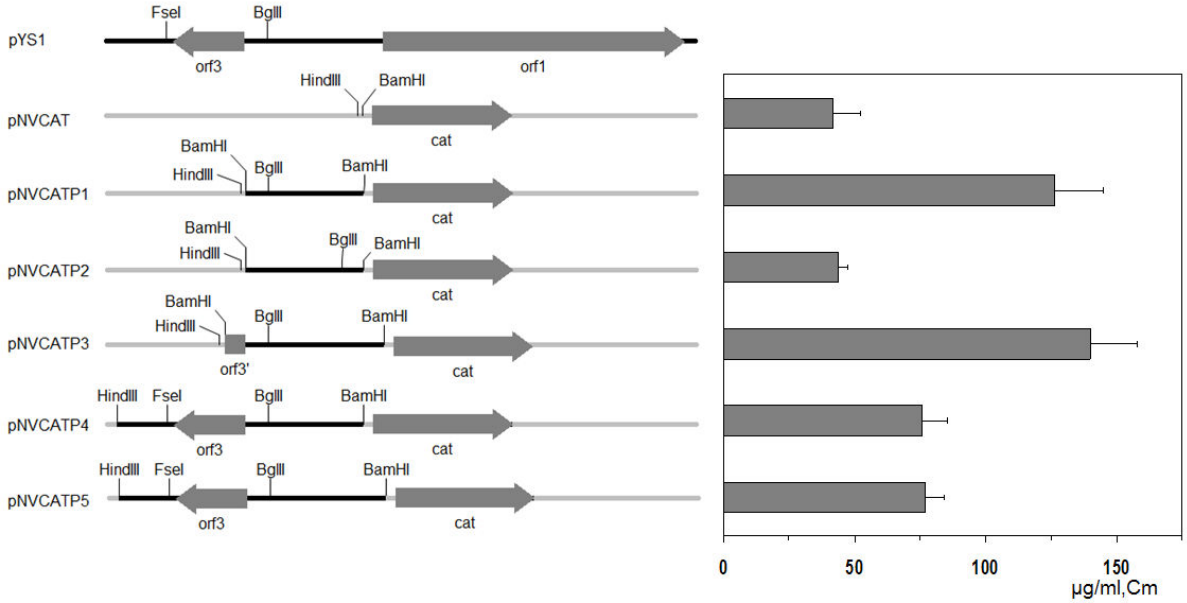



Figure 3.16. Phenotypic CAT-assay in *Nocardia* by using the reporter assay vector pNVCAT. (A) Construction of pNVCAT and its schematic map. A 1.9 kb *Psc*I-*Hpa*I fragment from p106 (Atchison *et al*., 1989), which contained a promoter-less *cat* gene downstream of a multiple cloning site, and a 4 kb *Hind*III-*Psc*I fragment from pNV18 (Chiba *et al*., 2007) were both filled-in and ligated to form pNVCAT. (B) Phenotypic CAT-assay to detect promoter activity in the *orf3*-*orf1* intergenic region and reduction of promoter activity in the presence of *orf3*. Parts of pYS1 as indicated were inserted upstream of the promoter-less *cat* gene in pNVCAT, and MIC of chloramphenicol (Cm) for *N. nova* transformed with these constructs are shown. Region of pYS1 containing *orf3* and *orf1* is shown at the top, with the intergenic regions in black lines. Below pYS1 is pNVCAT, in grey, in which nothing is inserted upstream of *cat*. In pNVCATP1, pNVCATP2, and pNVCATP3, the *orf3*-*orf1* intergenic region from pYS1 was inserted in the *Bam*HI site upstream of *cat*. The inserted DNA, whose orientations can be deduced from the position of the *Bgl*II site, is shown in black in these constructs. Plasmid pNVCATP1 contains the pYS1 *orf3*-*orf1* intergenic region in *orf1* orientation with respect to *cat*, whereas plasmid pNVCATP2 has the exact same region in *orf3* orientation. Plasmid pNVCATP3 is similar to pNVCATP1, except it has a longer section of the pYS1 *orf3*-*orf1* intergenic region towards *orf1* and the N-terminus of *orf3*. In plasmids pNVCATP4 and pNVCATP5, a larger section from pYS1 spanning the entire *orf3* gene and the *orf3*-*orf1* intergenic region was inserted upstream of *cat* such that the orientation of this section with respect to *cat* was identical to that with respect to *orf1* in pYS1. Plasmid pNVCATP5 contained a longer section of the *orf3*-*orf1* intergenic region towards *orf1* than pNVCATP4.

## 3.2 Construction of an inducible expression system for the analysis of antimicrobial genes from *Rhodococcus* phages

### 3.2.1 Testing of antimicrobial DNA from *Rhodococcus* phages on *Mycobacterium*

In my previous study (Shibayama, 2007), plasmid clones from genomic libraries of *Rhodococcus* phages FND1 and PGS1 were screened for inhibitory activity in *R. erythropolis*. In nine inhibitory clones that were analyzed, 18 ORFs were detected, of which 13 had no sequence similarity match to anything in the database, suggesting a huge pool of novel antibacterial mechanisms (Shibayama, 2007). As *Mycobacterium* is the actinomycete against which new antimicrobials are most urgently needed (Cox *et al*., 2010), those phage DNA were re-cloned into a mycobacterial vector pOLYG (O'Gaora *et al*., 1997) and screened for inhibitory activity in *M. smegmatis* (Table 3.6). Of the nine clones, five had transformation efficiencies of at least two orders of magnitude less than the vector-only control. To further study the inhibition exerted by these clones, construction of an inducible expression system in *Mycobacterium* was attempted.

Table 3.6. Inhibitory activity of *Rhodococcus* phage DNA towards *M. smegmatis*

| | Transformation efficiency (transformants × $10^4$/μg DNA) | | Transformation efficiency (transformants × $10^4$/μg DNA) |
|---|---|---|---|
| Clone | *R. erythropolis* SQ1 [a] | Clone | *M. smegmatis* $mc^2155$ |
| pFB7 | 0.02 | pFBM7 | <0.002 |
| pFB8 | <0.002 | pFBM8 | 0.64 |
| pFB10 | <0.002 | pFBM10 | 0.003 |
| pFB19 | <0.002 | pFBM19 | 0.067 |
| pFP8 | 0.006 | pFPM8 | 0.6 |
| pFP14 | <0.002 | pFPM14 | 0.008 |
| pFP16 | 0.006 | pFPM16 | 0.053 |
| pPH1 | 0.004 | NC | ND |
| pPH7 | 0.009 | pPHM7 | 7.9 |
| pDA71 only | 2.6 | pOLG only | 16 |

[a] Inhibitory activity towards *R. erythropolis* was determined by Shibayama (2007).
Clones in the same row have the same insert DNA in different vectors (pDA71 or pOLYG). NC, not constructed; ND, not determined.

### 3.2.2 Temperature-inducible expression system in *E. coli* using the λ switch

#### 3.2.2.1 Construction of a heat-inducible expression system in *E. coli*

Plasmid pEcoR251 (Zabeau and Stanley, 1982) (Fig. 3.19) is an *E. coli* vector with a suicide gene encoding the restriction endonuclease *Eco*RI, whose expression is driven off a λ promoter $P_R$. Binding of a λ-encoded cI repressor to two operators, $O_R1$ and $O_R2$, adjacent to the $P_R$ prevents the association of the RNA polymerase with the promoter, thereby repressing transcription (Ptashne, 2004). Although the suicide gene was meant to allow the positive selection of transformants with recombinant plasmids (Zabeau and Stanley, 1982), it could be also used as a model lethal gene for constructing an effective switch. To utilize the λ switch, the gene encoding a temperature-sensitive mutant of the repressor, cI857 (Sussman and Jacob, 1962), was cloned into pEcoR251. This was achieved by shotgun cloning a *Sau*3AI partial-digest of λ (*cI857*) DNA into the unique *Bam*HI site of pEcoR251 in *E. coli* at 30 °C. As the suicide gene of the vector should kill all transformants except where the *cI857* gene was successfully ligated into the vector, any transformant appearing on the selective plate was likely to be the desired clone. Fifteen transformants that grew on the selective plate were tested for growth at both permissive (30 °C) and non-permissive (42 °C) temperatures. Plasmid from a clone which grew at only the permissive temperature and had the shortest insert (~1.4 kb) was named pYS13 (Fig. 3.19). Sequencing of the insert confirmed the presence of the *cI857* gene. A deletion mutant pYS13* (Fig. 3.19), which was no longer lethal to the host, was also constructed by removing the ~900 bp *Eco*RI-*Sfu*I fragment containing most of the *ecoR*.

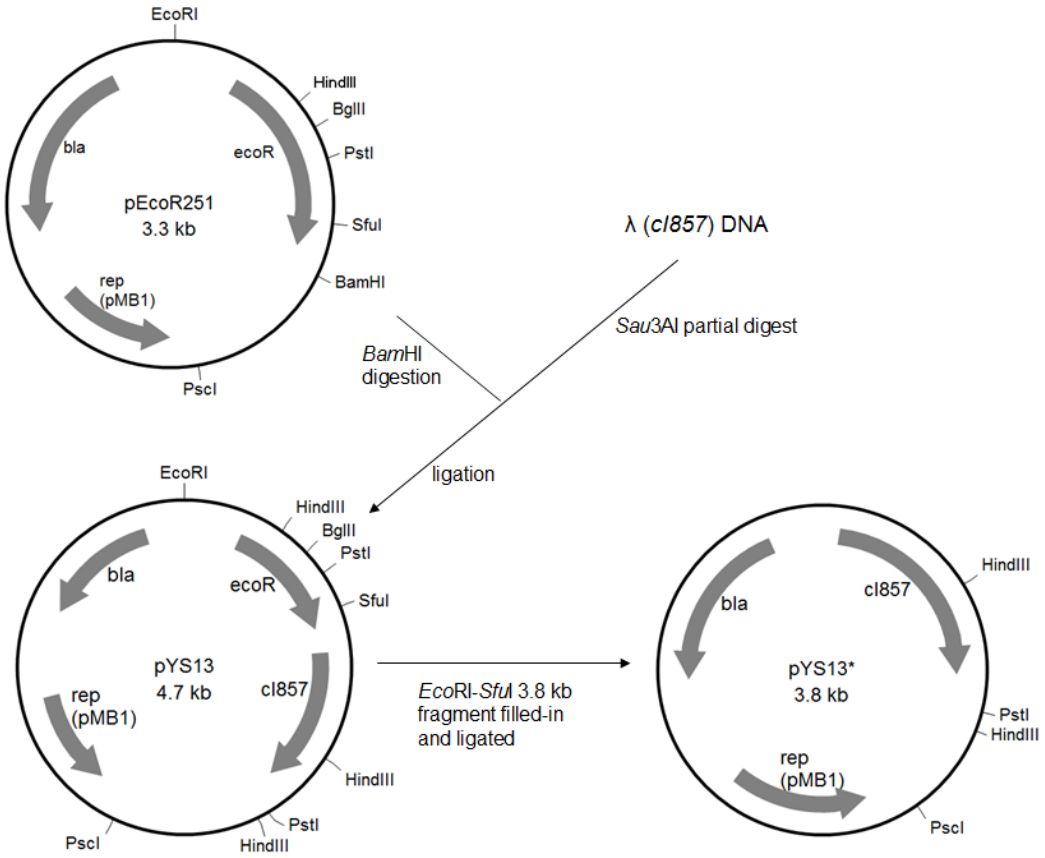



Figure. 3.17. Construction of pYS13 and pYS13*.

### 3.2.2.2 Survival curve studies using pYS13

Studies on inhibitory genes require a tight switch that would keep background transcription levels to a minimum and would allow a rapid increase in transcription under controlled situations. Using the *ecoR* as a model bactericidal gene, its inducibility in pYS13 was assessed in a survival curve study (Fig. 3.20), whereby the temperature of a culture of *E. coli* harbouring pYS13 was transferred from 30 °C to 42 °C at late logarithmic phase, during which colony forming units were measured every 30 minutes. At 4 hours after heat-induction, cell count decreased by more than five orders of magnitude, suggesting that this switch could tightly repress the transcription of a lethal gene and its expression could be induced effectively upon heat-induction.

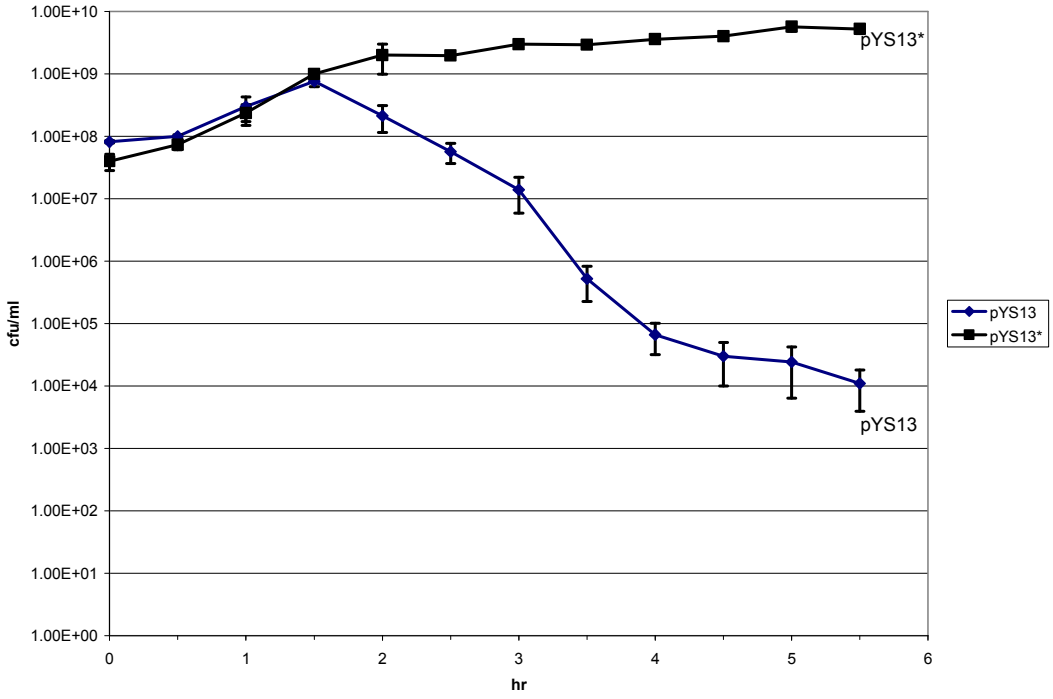



Figure 3.18. Phenotypic assay for heat-induction of *ecoR* expression in *E. coli* MM294::pYS13. The culture was transferred from 30 °C to 42 °C at 1.5 hrs. The values are from three independent assays.

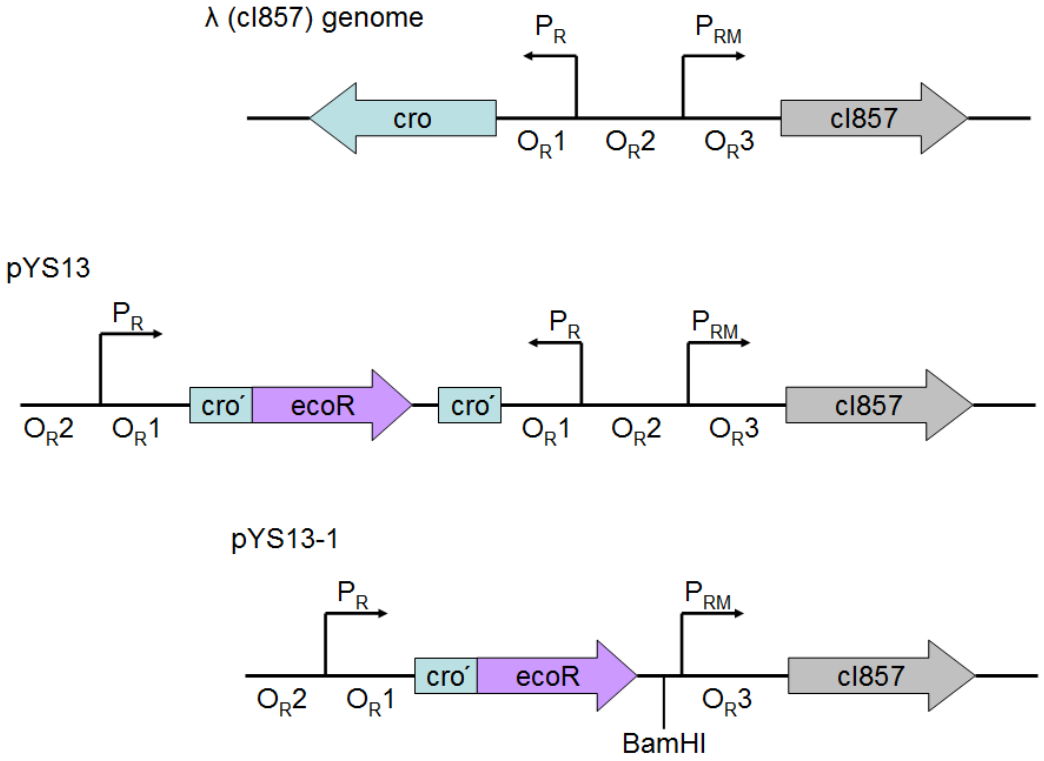
### 3.2.2.3 Construction and analysis of pYS13-1

Because pYS13 was constructed by randomly ligating a *Sau*3AI partial-digest of λ DNA into pEcoR251, the piece of phage DNA containing *cI857* had extra regions with undesirable regulatory elements (Fig. 3.21A). The entire intergenic region upstream of *cI857* and the N-terminus of *cro* was present in pYS13, although the only necessary element for the expression of *cI857* should have been the $P_{RM}$ promoter. Operator $O_R3$ may have been of additional benefit as it should negatively regulate *cI857* transcription when its cellular concentration is high (Ptashne, 2004). Because the region containing $P_R$, $O_R1$, $O_R2$, and the N-terminus of *cro*, as well as the downstream region of *cI857*, seemed unnecessary, a plasmid incorporating only the minimum necessary region containing $P_{RM}$, $O_R3$, and *cI857* was reconstructed. This was achieved by amplifying this region (~770 bp) from pYS13 and ligating into the unique *Bam*HI site of pEcoR251. A *Bam*HI site was incorporated into the forward primer and a *Bgl*II site into the reverse primer used for amplification, such that the unique *Bam*HI site was maintained upstream of *cI857* after

ligation into pEcoR251, which will allow for insertion of additional regulatory elements in the future. The new plasmid was named pYS13-1 (Fig. 3.21B). To construct a non-lethal version of the plasmid, pYS13-1* (Fig. 3.21B), the ~190 bp *Bgl*II-*Pst*I fragment from within *ecoR* was removed.

To test the switch in pYS13-1, a survival curve study was conducted (Fig. 3.22). A reduction in cell count of approximately five orders of magnitude was achieved, showing that this plasmid contains necessary elements to effectively induce the expression of *ecoR*.

A



B

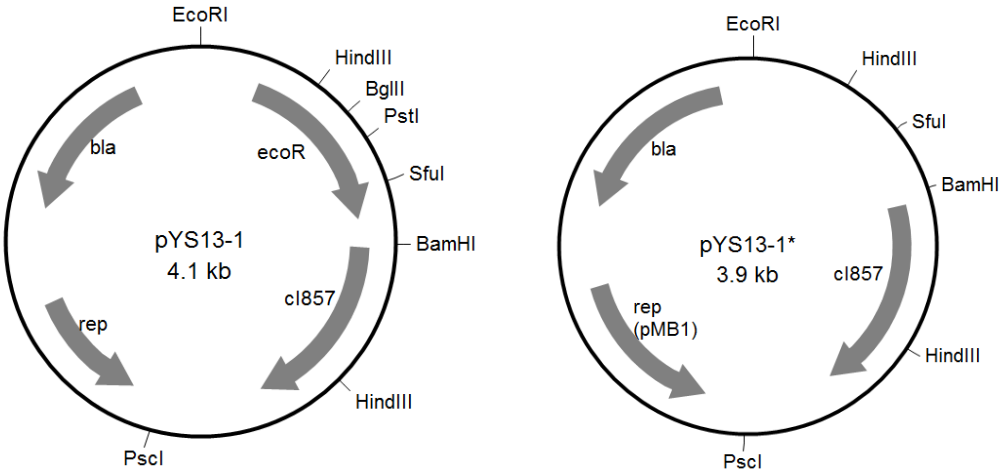



Figure 3.19. Construction of pYS13-1. (A) Organization of $P_R$, $P_{RM}$, $O_R1$, $O_R2$, and $O_R3$ in the λ genome, pYS13, and pYS13-1. The length of the ORFs are not to scale to the intergenic regions. (B) Schematic map of pYS13-1 and pYS13-1*.

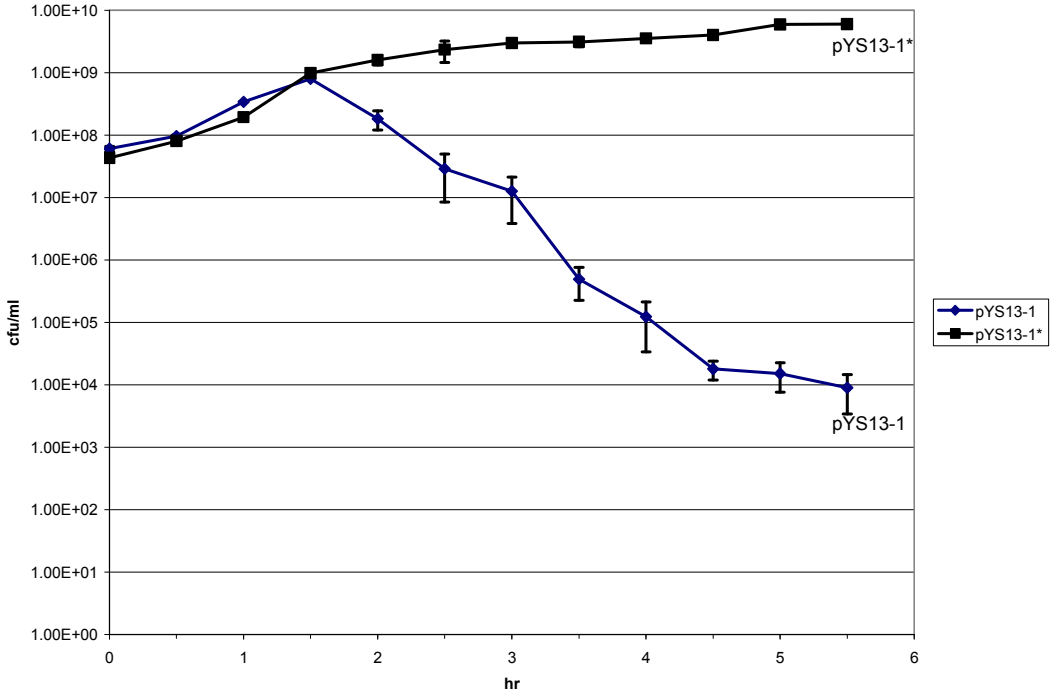



Figure 3.20. Phenotypic assay for heat-induction of *ecoR* expression in *E. coli* MM294::pYS13-1. The culture was transferred from 30 °C to 42 °C at 1.5 hrs. The values are from three independent assays.

### 3.2.3 Construction and analysis of a heat-inducible expression system in *M. smegmatis*

To test the effectiveness of the above heat-inducible λ switch in *Mycobacterium*, it was first necessary to transfer the components that make up the switch from pYS13-1 into a mycobacterial replicon. This was done by ligating the ~2400 bp *Eco*RI-*Psc*I fragment containing *ecoR* and *cI857* from pYS13-1 to the backbone of pNV19, the ~4200 bp *Eco*RI-*Psc*I fragment (Fig. 3.23). Similar treatment was made for pYS13-1*. The newly constructed plasmids were respectively named pYSM13-1 and pYSM13-1*.

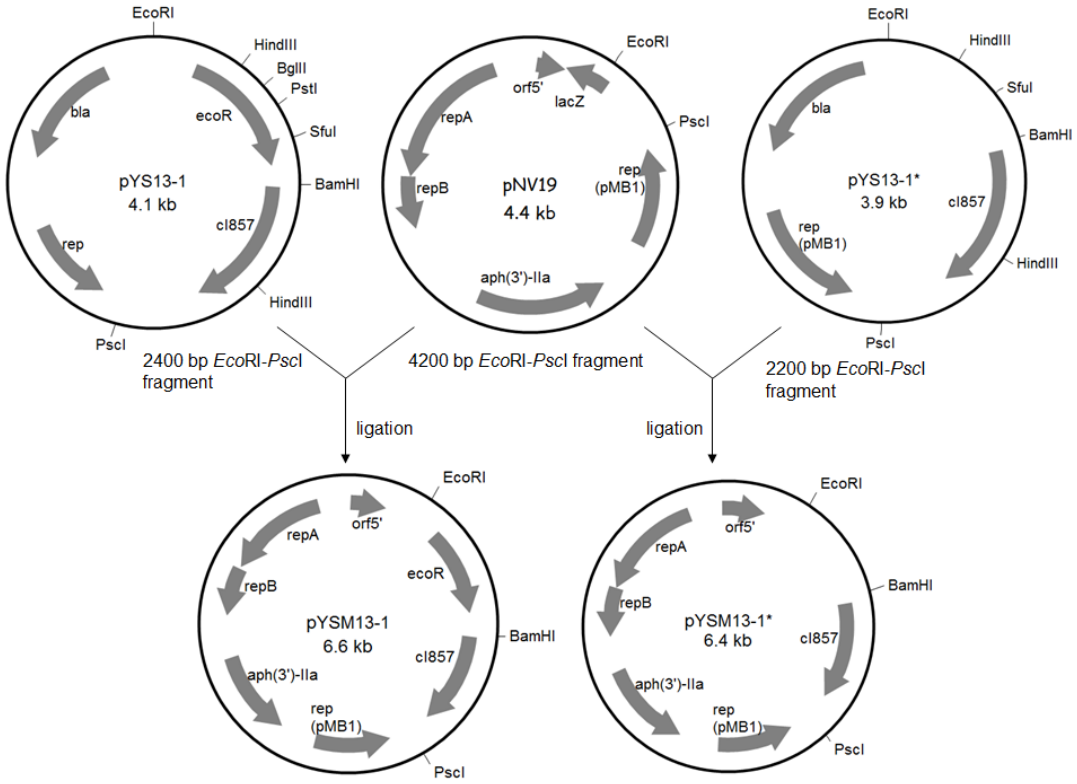



Figure 3.21. Construction of pYSM13-1 and pYSM13-1*.

The temperature-sensitive switch was first tested in *Mycobacterium* on agar plates. *M. smegmatis* $mc^2155$ transformants of pYSM13-1 only grew at 30 °C while transformants of pYSM13-1* grew at both 30 °C and 42 °C, suggesting that both *ecoR* and *cI857* were functional in this organism under these conditions. The switch was then tested in broth by

a survival curve assay (Fig. 3.24). However, no reduction in cell count was observed following heat induction.

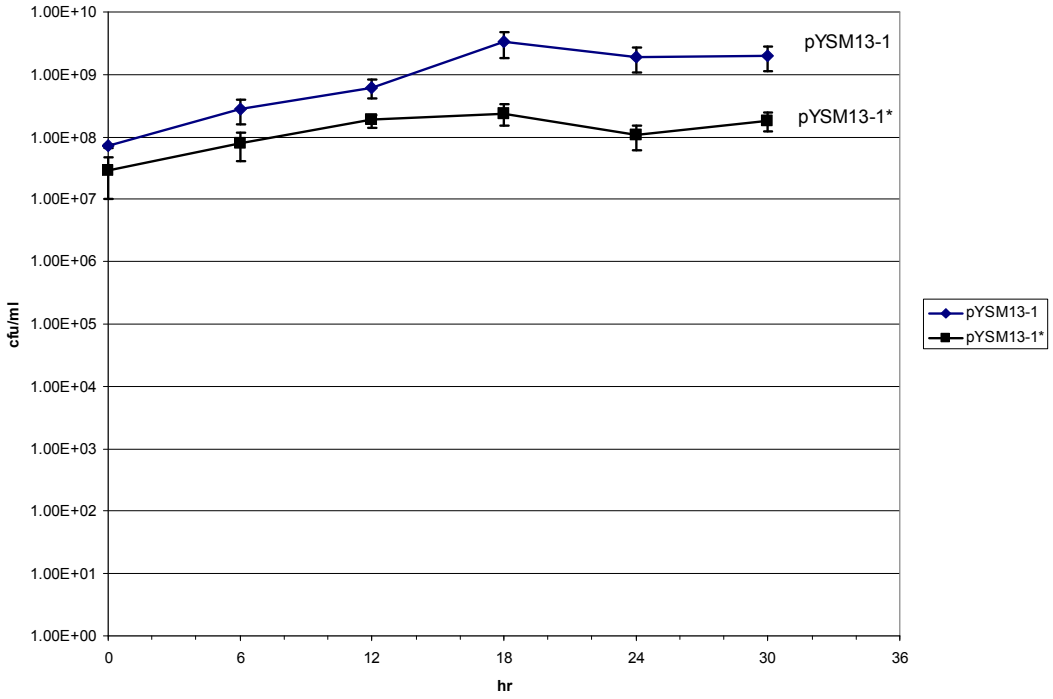



Figure 3.22. Phenotypic assay for heat-induction of *ecoR* expression in *M. smegmatis* mc$^2$155::pYSM13-1. The culture was transferred from 30 °C to 42 °C at 6 hrs. The values are from three independent assays.

### 3.2.4 Isolation of bacteriophages plaquing on *M. smegmatis*.

It was unclear why the temperature-sensitive switch in *M. smegmatis* was functional on solid medium but not in broth. One possibility was inefficient expression of *ecoR* and *cI857* in the latter condition. To investigate whether insertion of transcriptional promoters upstream of either or both of these genes would allow the switch to function in broth, bacteriophages plaquing on *M. smegmatis* were isolated from various soil samples (Table 3.7), under the assumption that these phages will provide a strong promoter for the mycobacterial host.

Table 3.7. Isolation of phages plaquing on *M. smegmatis*

| Soil location | No. plaques |
|---|---|
| Rosebank, South Africa | 0 |
| Melville, South Africa | 0 |
| Sandton, South Africa | ~10 |
| Wits, South Africa | 2 |
| Parys, South Africa | 0 |
| Yeoville Park, South Africa | 0 |
| Palm Ridge, South Africa | 0 |
| Palm Ridge 2, South Africa | 6 |
| Emmarentia, South Africa | 2 |
| El Vecino, Mozambique | 1 |
| Ciudad de Rodrigo, Mozambique | 1 |
| Maputo, Mozambique | 3 |
| Bilene, Mozambique | 10 |
| Chiba, Japan | 5 |
| Hakone, Japan | 4 |
| Pasadena, USA | 2 |
| Caltech, USA | confluent lysis |
| Los Robles, USA | confluent lysis |
| Berlin, Germany | 0 |

### 3.2.5 Construction of promoter-probe vector for *Mycobacterium*

To isolate a promoter from phages, a promoter-probe vector to supplement pNVCAT (Fig. 3.16) was constructed. For this vector, the green fluorescent protein (*gfp*) gene contained within a 800 bp *Hind*III-*Bam*HI fragment from pJM608 was ligated into pNV18 (Fig. 3.25). The resulting vector, pNVGFP, had a unique *Hind*III site upstream of a promoter-less *gfp* gene in a pNV18 backbone.

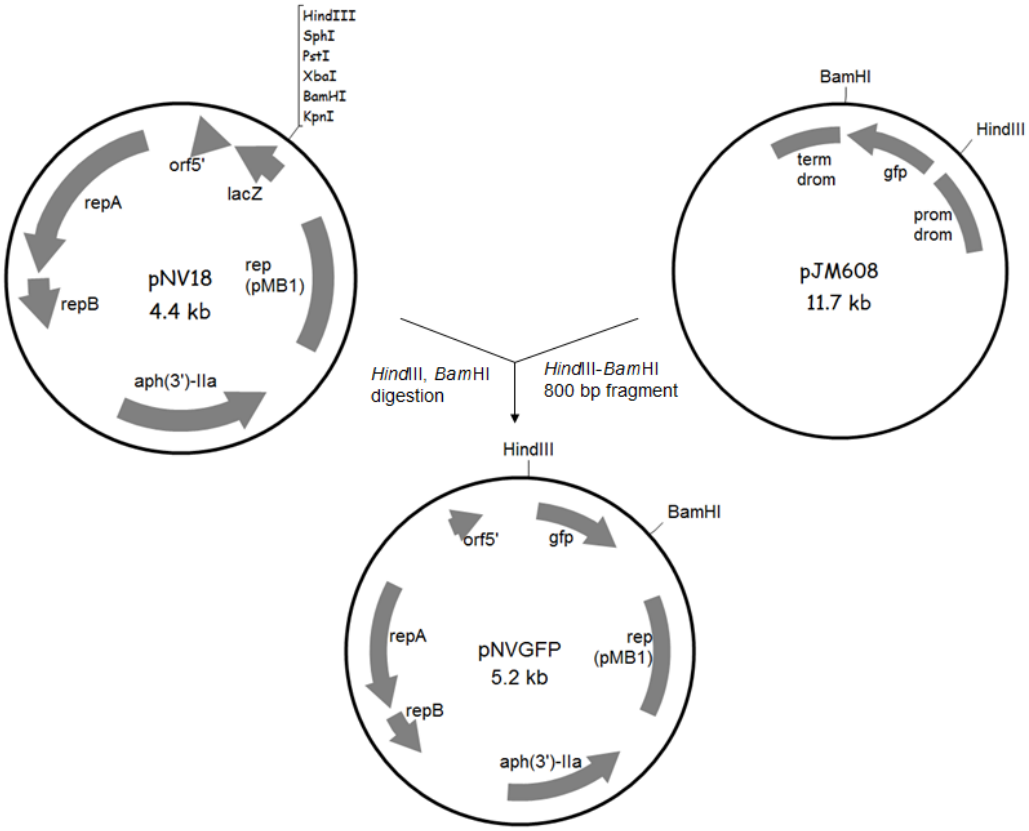



Figure 3.23. Construction of pNVGFP.

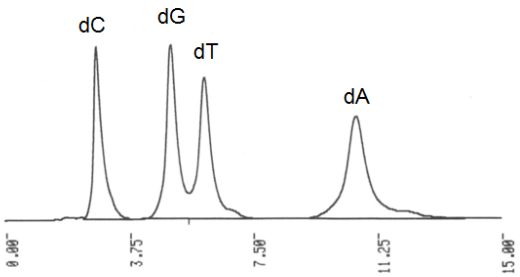
## 3.3 Analysis of nucleotide bases in *Rhodococcus* phage nucleic acid

### 3.3.1 Potential presence of modifications or unusual bases in phage nucleic acid

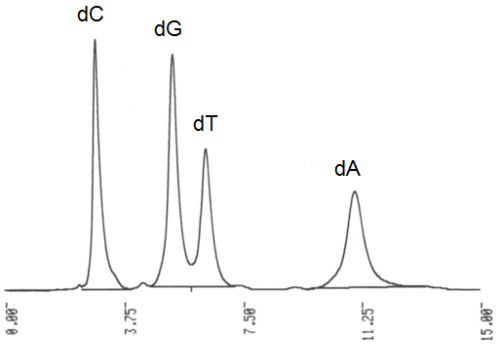
In a previous study (Shibayama, 2007), the genomic DNA from *Rhodococcus* phage WTS1 was resistant to restriction by *Hind*III, *Bgl*II, *Bcl*I, *Pst*I, *Nsi*I, *Bmy*I, *Sfu*I, and *Cla*I. Following digestion by *Xho*II, the DNA could not be cloned into the *Bgl*II site of pDA71 (Quan and Dabbs, 1993). DNA from phage KZA1 was resistant to *Acc*I in addition to all of the enzymes above. Cloning of its *Bam*HI-digested DNA into the *Bgl*II site of pDA71 was also unsuccessful. As there was a possibility that these results were due to the potential presence of modifications or unusual bases in the phage DNA, it was necessary to gain further evidence to support this hypothesis. Methods chosen were HPLC and TLC.

### 3.3.2 Analysis by HPLC

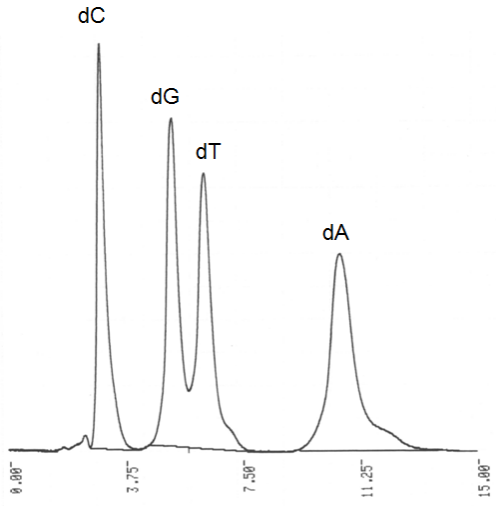
Nucleosides from each phage were separated by HPLC. Several important controls were included. These were DNA from *E. coli* phages λ (no unusual base) and T4 (hydroxymethylcytosine) and *B. subtilis* phages PBS1 (uracil) and SPO1 (hydroxymethyluracil). The HPLC profiles obtained for all samples are illustrated in Figure 3.17. The identities of peaks for the regular bases were according to previous studies (Kaneko *et al*., 1986; Katayama-Fujimura *et al*., 1984; Tamaoka and Komagata, 1984), and the retention time of each base was as follows: deoxycytidine, 2.76 minutes; deoxyguanidine, 5.03 minutes; deoxythymidine, 6.06 minutes; deoxyadenosine, 10.65 minutes. For *E. coli* MM294-4 and phage λ, similar peaks were obtained. The profile for phage T4 contained a new peak with a retention time of 1.78 minutes, which probably corresponded to hydroxymethyldeoxycytidine, while the deoxycytidine peak had disappeared. For phage PBS1, the profile reflected the complete absence of deoxythymidine and the presence of different base with a retention time of 3.71 minutes, probably corresponding to deoxyuridine. The profile for phage SPO1 also had the deoxythymidine peak replaced by one with a retention time of 3.71 minutes, probably corresponding to hydroxymethyldeoxyuridine. This assumption was supported by the fact that the commercial sample of this unusual base gave a similar peak. It should be noted that deoxyuridine and hydroxymethyldeoxyuridine had the same retention time. For *R. erythropolis* and its phages WTS1, KZA1, and PGS1, the profiles were similar to one another, and nothing irregular was observed, suggesting absence of unusual bases. For *S. aureus* phage KG3 (Green, 2007), no unusual peak could be observed either.



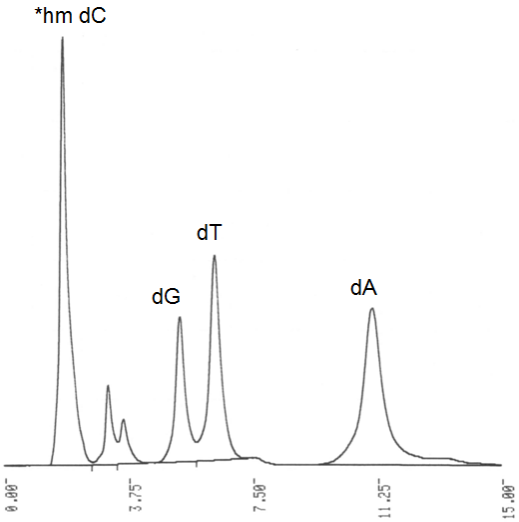

Calf thymus DNA



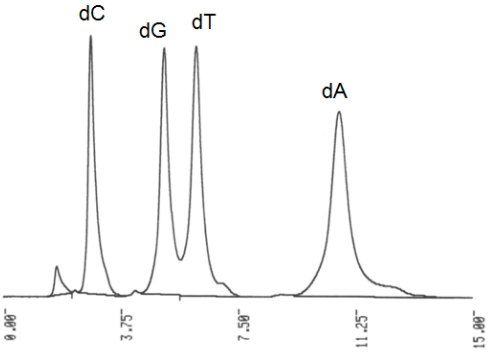

*E. coli* MM294-4



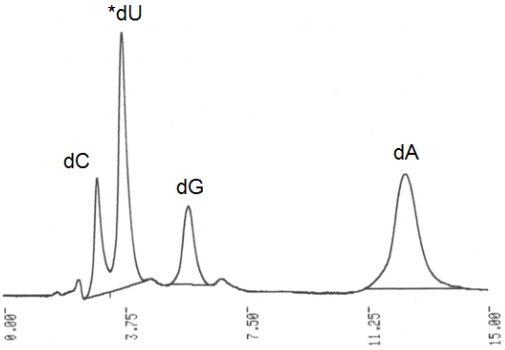

*E. coli* phage λ



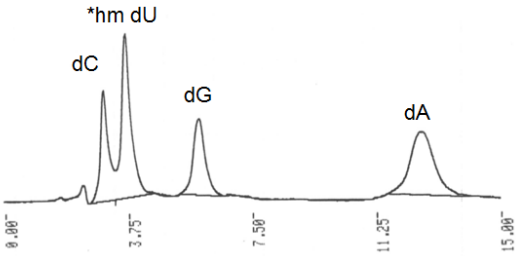

*E. coli* phage T4



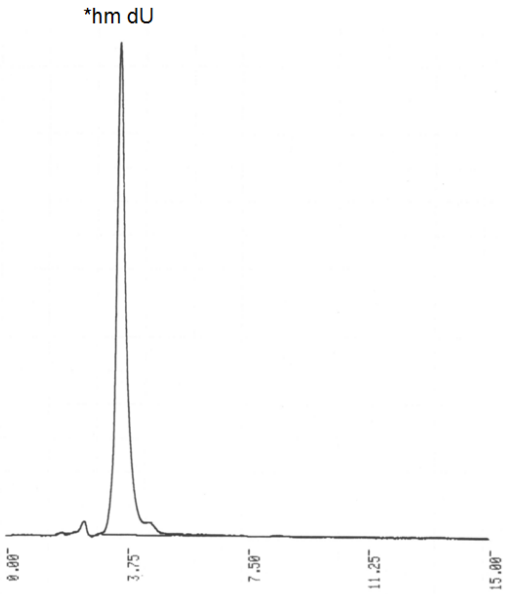

*B. subltilis* IA3



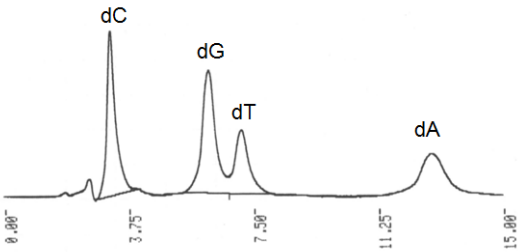

*B. subltilis* phage PBS1



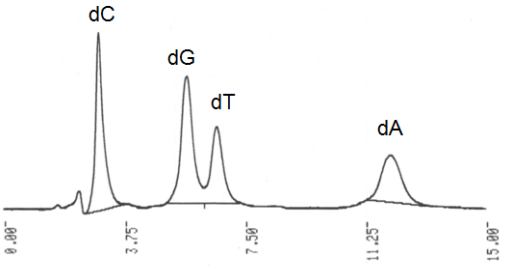

*B. subltilis* phage SPO1



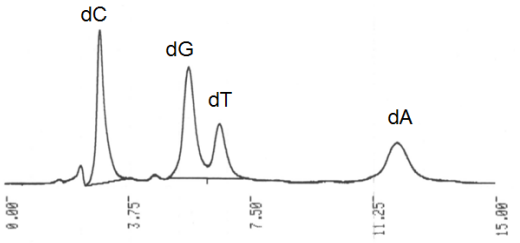

hydoxymethyluracil



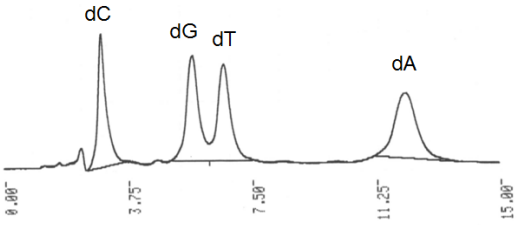

*R. erythropolis* SQ1



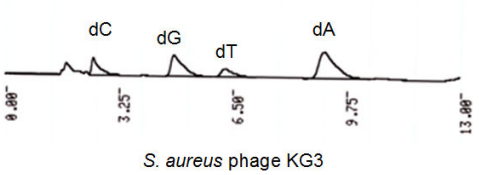

*R. erythropolis* phage WTS1




*R. erythropolis* phage KZA1




*R. erythropolis* phage PGS1




*S. aureus* phage KG3

Figure 3.24. HPLC profiles of nucleosides from phages and their hosts. Calf thymus DNA, phage λ DNA, and hydroxymethyluracil were commercially obtained. The identities of peaks for the regular bases were according to previous studies (Katayama-Fujimura *et al*., 1984; Tamaoka and Komagata, 1984; Kaneko *et al*., 1986). *These peaks were speculated to represent the unusual bases, as they were not present in their hosts. dC, deoxycytidine; dG, deoxyguanidine; dT, deoxythymidine; dA, deoxyadenosine; dU, deoxyuridine; hm dC, hydroxymethyldeoxycytidine; hm dU, hydroxymethyldeoxyuridine.

### 3.3.3 Analysis by TLC and Mass spectrometry (MS)

Further analysis was undertaken by TLC (Fig. 3.25). Deoxyribonucleoside monophosphates from commercially obtained herring sperm DNA gave four spots. DNA from two bacteria included in this analysis, *Nocardia mexicana* and *Staphylococcus aureus*, both gave spots at similar positions as the herring sperm DNA. Analysis by MS (undertaken by Dr. Yoshihisa Sei, Kagawa School of Pharmaceutical Sciences, Tokushima Bunri University, Kagawa, Japan) revealed the identity of the four spots from the *S. aureus* DNA, as indicated in Figure 3.25. Spots from phage T4 reflected the missing cytosine, but one corresponding to hydroxymethyldeoxycytosine monophosphate was also missing, probably due to co-migration with one of the three other bases. Similarly, for phages PBS1 and SPO1, spots for deoxythymidine monophosphate was missing, but those for the unusual bases had probably co-migrated with one of the regular ones, thus only three spots were observed. For phage λ, an extra spot was observed between those for deoxythymidine monophosphate and deoxyguanidine monophosphate. Further purification by HPLC revealed the presence of two molecules from this spot, one of which was revealed to be adenosine monophosphate, and the other remains unidentified. For *Rhodococcus* phage FND1, two unusual spots were present. One of these was at a similar position as the extra spot from λ, and MS analysis revealed it to be adenosine monophosphate (i.e. from RNA). From the other extra spot, occurring between deoxyguanidine monophosphate and deoxyadenosine monophosphate, two different molecules were detected, one of which was identified again as adenosine monophosphate while the other could not be identified. No irregular spots could be observed with phage

WTS1. For *S. aureus* phage KG3, two unusual spots were observed at similar positions as phage FND1. However, identity of the top one could not be revealed by MS analysis. Two molecules were detected from the bottom spot, whereby one turned out to be guanidine monophosphate and the other adenosine monophosphate.

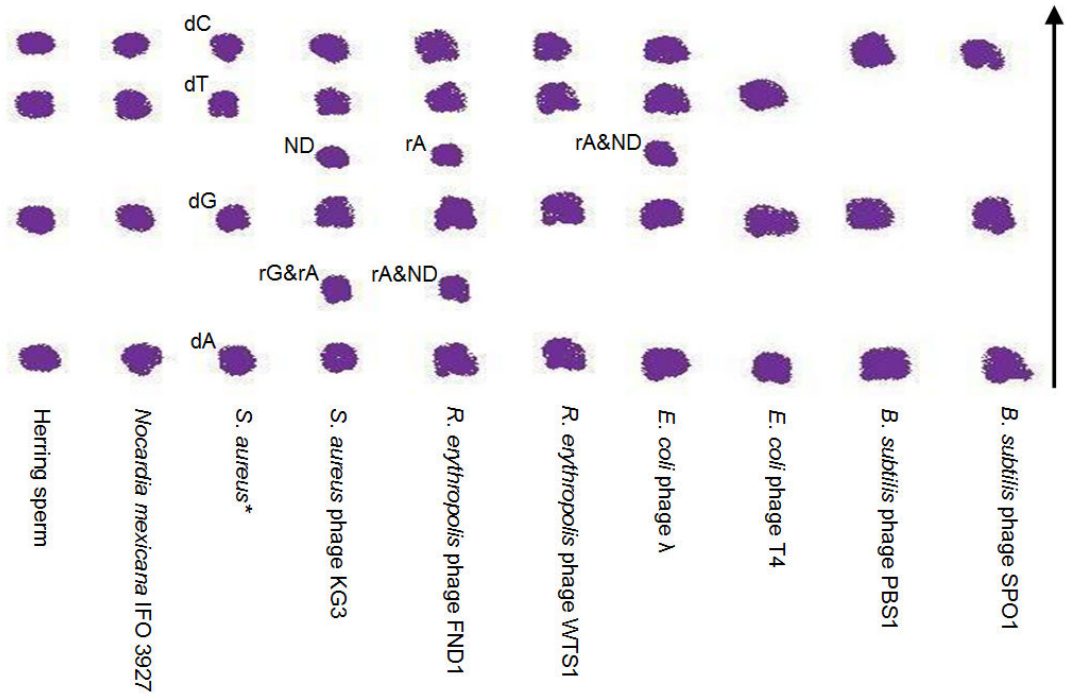



Figure 3.25. TLC profiles of nucleotides from bacteria and phages. Profiles were drawn (due to lack of appropriate photographic equipment) according to observation of the TLC plates under the UV light. Arrow indicates the direction of migration on the TLC plate. Identities of spots as determined by MS are shown. dCM, deoxycytidine monophosphate; dTM, deoxythymidine monophosphate; dGM, deoxyguanidine monophosphate; dAM, deoxyadenosine monophosphate; rGM, guanidine monophosphate; rAM, adenosine monophosphate; ND, identity could not be determined. *strain unknown.

### 3.3.4 Ribosomal contamination of phage preparations

Out of the five irregular spots observed above, at least four of them contained ribonucleoside monophosphates, which raised the suspicion that those phage preparations were contaminated with host RNA. Phage preparation generally involved precipitation of phage particles from a high-titre lysate using PEG 6000 followed by purification in caesium chloride, which could be either a step or an equilibrium gradient. Yamamoto *et al*. (1970) reported the co-precipitation of ribosomes with phage particles by PEG and

their subsequent co-purification in a caesium chloride gradient, as ribosomes possess similar buoyant density as phage particles. For example, the buoyant density of the 70S ribosome is 1.62 g $cm^{-3}$ (Fenwick, 1971), while that of phages is generally ~1.5 g $cm^{-3}$ (Yamamoto *et al*., 1970). To investigate whether host RNA contamination was the cause of the presence of RNA in phage nucleic acid preparations, phage KG3 was re-prepared under more stringent conditions. Previously, its lysate was not treated with DNAse and RNAse prior to precipitation by PEG, and the purification was done in a step gradient of caesium chloride. This time, the lysate was first extensively treated with both RNAse and DNAse prior to precipitation using PEG 6000. Following precipitation, the phages were purified twice in an equilibrium gradient. When the nucleotides from this newly prepared KG3 were separated by TLC, only the four regular spots could be observed (Fig. 3.19), suggesting that the two extra spots previously observed had originated from the host bacteria, most probably from the ribosomes. This observation was supported by the TLC profile of WTS1, which had been purified in an equilibrium gradient and did not produce any extra spots either.

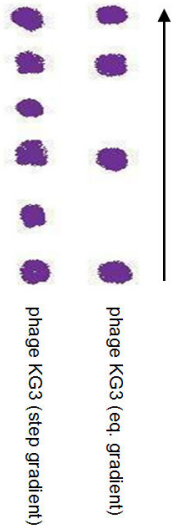



Figure. 3.26. TLC profiles of nucleotides from phage KG3 purified by step or equilibrium (eq.) gradient.

# 4. DISCUSSION

## 4.1 Functional analysis of plasmid pYS1 from *Nocardia*

### 4.1.1 Detection of plasmids from *Nocardia*

Plasmids across the whole prokaryotic spectrum are diverse in terms of size, mode of replication and transfer, host range, and the set of genes they carry (Thomas, 2000). For *Nocardia*, information of this nature is limited, which was a principal reason for my characterizing pYS1 from *N. aobensis*. I detected plasmids in pathogenic strains of *Nocardia* at a similar frequency to previous reports for this genus (Kasweck *et al*., 1981; Provost *et al*., 1996; Qasem *et al*., 1999). Those studies could not attribute any phenotypic traits to plasmids harboured, at least not with respect to the properties tested, viz. virulence or antibiotic resistance. Although these phenotypes were not dealt with in this study, the data support the notion that harbouring of plasmids is not a common feature among pathogenic *Nocardia*. It remains unknown whether pYS1 or any other plasmid from *Nocardia* contains genes advantageous to its virulence.

### 4.1.2 Rolling-circle-replication of pYS1

From both sequence analyses and experimental data, it is clear that pYS1 is a rolling-circle-type plasmid belonging to the pIJ101/pJV1 (Kendall and Cohen, 1988, Servin-Gonzalez *et al*., 1995) family. Motifs usually associated with replication initiation proteins of rolling-circle-type plasmids were all identified in the Orf2 amino acid sequence.

In the current model of rolling-circle-replication (Fig. 4.1) (Khan, 2005), leading strand synthesis is initiated when the Rep protein makes a nick at the GG dinucleotide at the DSO, which was conserved in pYS1 as shown by the multiple alignment of its nucleotide sequence to DSO sequences from related plasmids. As expected, the DSO in pYS1 was demonstrated to be indispensable for replication when pYS1R1D1, a plasmid containing the mutated form of the DSO, was unable to replicate in *Nocardia*.

As replication proceeds, synthesis of the leading strand causes the old leading strand to be displaced and to form a single-stranded intermediate (Khan, 2005), which I detected for

pYS1 by Southern hybridization as a band migrating faster than the double-stranded form and disappearing upon treatment with S1 nuclease. On the single-stranded intermediate, lagging strand synthesis begins because the SSO serves as a signal for RNA primer production (Khan, 2005). The SSO is characterized by inverted repeats forming an extensive secondary structure (Khan, 2005), which was predicted on pYS1. On one of the exposed loop regions, a sequence similar to a conserved CS-6 sequence TAGCGA (Kramer *et al*., 1998a; Khan, 2004) was detected, strengthening the accuracy of my prediction. Although overall SSO sequences are not well-conserved among plasmids of the same family, predicted secondary structures for two related plasmids, pRE8424 (Nakashima and Tamura, 2004) of *R. erythropolis* and pXT107 (Xia et al., 2006) of *Nocardia* sp. 107, are depicted for comparison (Fig. 4.2).

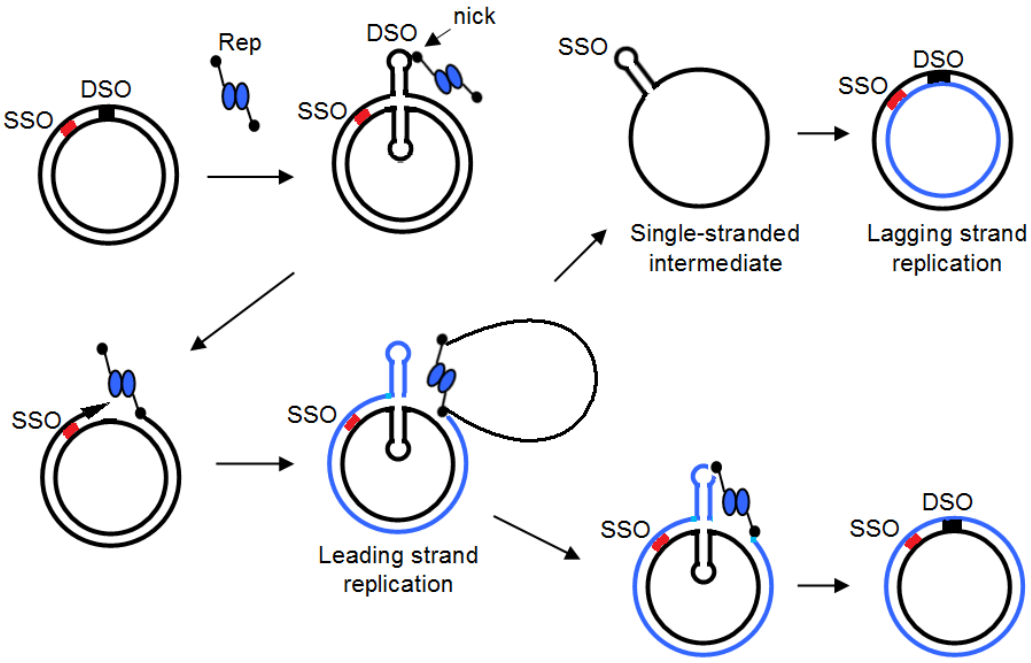
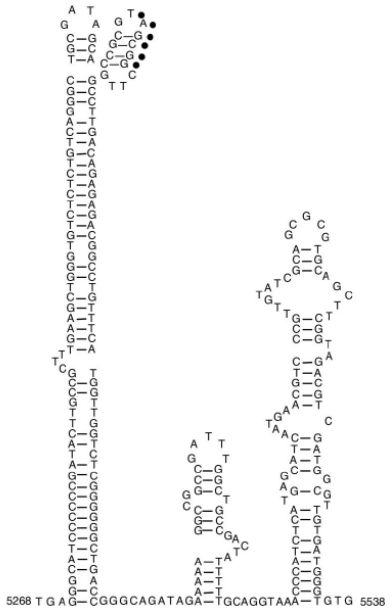



Figure 4.1. Current model of plasmid rolling-circle replication. Newly synthesized strands are indicated in blue. Modified from Khan, 2005.

A




B

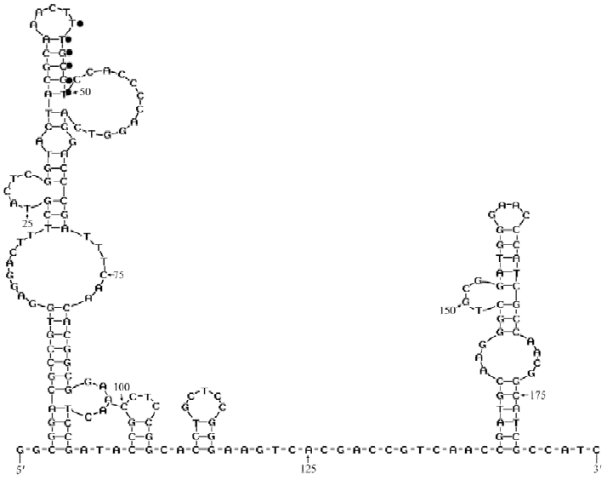



Figure 4.2. Predicted SSO structures of (A) pRE8424 (Nakashima and Tamura, 2004) and (B) pXT107 (Xia *et al*., 2006). Sequences similar to conserved CS-6 sequences are indicated by dots.

Four types of SSO have so far been described based on their CS-6 sequences (Table 4.1). Of these, the pYS1 SSO is most similar to type A, exemplified by pE194 of *S. aureus* (Kramer *et al*., 1998a). However, this type of SSO is known to function efficiently only in their native host (Khan, 2004), which is contradictory to the broad host-range of pYS1. It is therefore clear that these four types of SSO do not make up an exhaustive list and

that additional work is necessary to further characterize the pYS1 SSO. Analysis on new plasmids should further reveal a greater variety of this element.

Table 4.1. SSO types and their CS-6 sequences

| SSO type | CS-6 sequence | Plasmid example | Host | Reference |
|---|---|---|---|---|
| A | TAGCGT | pLS1 | *Streptococcus pneumoniae* | del Solar *et al*., 1993 |
| A | TAGCGA | pE194 | *Staphylococcus aureus* | Kramer *et al*., 1998a |
| W | TAGCGT | pWV01 | *Lactococcus lactis* | Seegers *et al*., 1995 |
| T | CAGCGC | pBAA1 | *Bacillus* sp. | Devine *et al*., 1989 |
| U | AAGGGT & GATCTT | pUB110 | *Staphylococcus aureus* | Boe *et al*., 1989 |

In most rolling-circle-type plasmids, mutating or deleting the SSO leads to an accumulation of the single-stranded intermediate (Khan, 2005). This is because plasmid replication can begin as long as the DSO serves as a signal for replication initiation (leading strand synthesis). However, the single-stranded intermediate produced would have difficulty initiating lagging strand replication in the absence of a correct SSO sequence. As a result, this intermediate form would remain single-stranded, without any increase in the number of double-stranded plasmid. The cell would therefore repeatedly initiate plasmid replication at the DSO, eventually leading to an exponential increase in the number of the single-stranded intermediate (Khan, 2005).

In pYS1 mutation of the conserved sequence in the SSO, however, did not lead to accumulation of single-stranded intermediates (Fig. 3.6). Although this was not the case in a previous study concerning pRE8424 (Nakashima and Tamura, 2004), whose mutation in a similar region did lead to accumulation of the single-stranded form, it is understandable that tampering with a sequence involved in termination of RNA primer synthesis (i.e. CS-6) (Kramer *et al*., 1997; Kramer *et al*., 1998b) might not necessarily cause much instability to plasmid replication.

On the other hand, the SSO deletion experiment showed the indispensability of the SSO region in pYS1 replication. Maximum change tolerable to the secondary structure was deletion of the lower stem consisting of 7 bp, leading to a shorter hairpin structure (Fig.

3.5, pYS1R3 clone). Further deletions abolished plasmid maintenance, indicating that pYS1 does not carry secondary signals for single-strand to double-strand conversion as do many rolling-circle plasmids. There is room for future studies on the interaction between the RNA primase and different parts of the secondary structure of the SSO. It would also be interesting to mutate the first terminal loop of the predicted SSO structure while maintaining the same secondary structure. This experiment, however, is not easy as the loop is flanked by a perfect inverted repeat whereby primers designed to bind to this region would readily undergo both intramolecular and intermolecular annealing.

### 4.1.3 *Nocardia- E. coli* shuttle vector pYS1R1

As the SSO deletion experiment revealed the minimal replicon of pYS1 to be within a 2 kb region containing the *rep* (*orf2*), DSO, and SSO, pYS1R1 (Fig. 3.5) can serve as a *Nocardia-E. coli* shuttle vector, which, with further manipulation, can result in a more useful vector. Potential further manipulations include increasing the number of unique cloning sites. Currently, restriction sites for *Bam*HI, *Eco*RI, *Sal*I, and *Hinc*II, which are in the multiple cloning site, also exist within *rep*, limiting their use as cloning sites. Silent mutations may be introduced to these sites to eliminate them while keeping the Rep amino acid sequence intact.

To my knowledge, pNV18/19 (Chiba *et al*., 2007), which uses the replicon from a mycobacterial plasmid pAL5000 (Stolt and Stoker, 1996), has been the most useful *Nocardia-E. coli* shuttle vector available. Compatibility of the pYS1 and pAL5000 replicons may serve as an advantage where introduction of two separate plasmids is required, not only into *Nocardia* but also into related bacteria such as *Mycobacterium* whose significant number of vectors rely on the pAL5000 replicon (Pashley and Stoker, 2000). As a broad-host-range replicon, vectors based on pYS1 may also benefit studies on *Rhodococcus* and *Gordonia*.

### 4.1.4 pYS1 copy number

Under the assumption that the *N. aobensis* genome size was not very different from that of *N. farcinica*, pYS1 copy number was estimated to be approximately ten per cell. It was expected that pYS1 copy number would be similar to that of pAL5000, as the gel

indicating coexistence of pYS1 and pNV19, which uses the pAL5000 replicon, showed a comparable intensity of fluorescence, implying similar copy number. As the copy number of pAL5000 in *M. smegmatis* is reported to be about five (Ranes *et al*., 1990), it can be concluded that the method of comparing fluorescence intensities between chromosomal and plasmid DNA on an agarose gel is suitable for deriving a rough estimate of copy number.

### 4.1.5 *orf1* and *orf3* of pYS1

In the well-described *Streptomyces* plasmid pIJ101 (Kendall and Cohen, 1987), plasmid transfer is mediated by the product of the *tra* gene, whose expression is controlled by a Gnt-R type transcriptional repressor KorA. Without this negative regulation, the gene is lethal to the host. A similar situation applies to another *Streptomyces* plasmid pSN22 (Kataoka *et al*., 1991), where expression of *traB* is controlled by the *traR* product. As TraB was shown to be localized to the cell membrane (Kosono *et al*., 1996), it has been speculated that its uncontrolled production leads to interference with membrane integrity. In both pIJ101 and pSN22, the gene encoding this transcriptional repressor occurs just upstream of the transfer gene, separated by a short intergenic region, and is transcribed in the opposite direction (Kendall *et al*., 1987; Kataoka *et al*., 1991) (Fig. 4.3). Expression of both genes is regulated by a bidirectional promoter occurring in this intergenic region (Kendall *et al*., 1987; Kataoka *et al*., 1991).

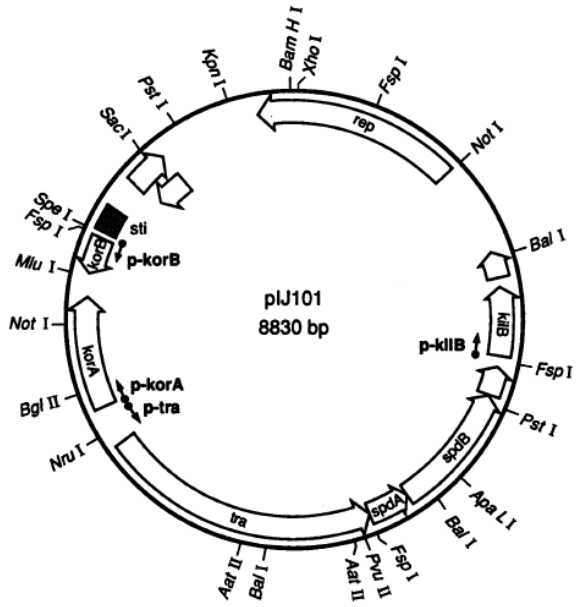



Figure 4.3. Schematic map of pIJ101 showing the positions of *tra and korA* (Stein *et al*., 1989). p-*korA* and p-*tra* refer to the bidirectional promoter in the intergenic region.

In pYS1, Orf1 presented a curious case, as its lethality was consistent with its sequence resemblance to Tra, but its lack of obvious involvement in plasmid transfer, as judged by the transfer frequencies of the insertion mutants, was inconsistent with it. It should be noted, however, that there is a possibility that each of the five locations disrupted upon construction of the insertion mutants, was involved in conjugation, in which case the transfer frequency would have been reduced in all mutants. This situation would be true if both *orf3*-*orf1* intergenic region (disrupted for pYS1K2) and *orf2*-*orf3* intergenic region (disrupted for pYS1K5) play a role in plasmid transfer as much as *orf1*. Otherwise, similar conjugation frequencies arising from mutants disrupted within *orf1* or outside it indicate its lack of involvement in plasmid transfer and the presence of other signals which mediate this function.

Data from transformations suggested that *orf1* may instead be involved in plasmid stability, as it took significantly longer for colonies to arise when various species were transformed with an *orf1*-disrupted plasmid, as compared to those transformed with a plasmid disrupted in the *orf3*-*orf1* intergenic region. This retardation was observed for just one of the four *Nocardia* species but for all of the six *Mycobacterium*, *Gordonia*, and *Rhodococcus* species (Table 3.5). The effect was not simply due to slower growth of

these transformants, and while it may be premature to draw conclusions from these observations, it is tempting to speculate that *orf1* is involved in plasmid stability when establishing itself in a new host, probably through a cell-membrane-associated mechanism.

The phenotyptic CAT-assay suggested the presence of a promoter for *orf1* in the *orf3-orf1* intergenic region of at least ~650 bp, but no promoter activity could be detected for *orf3* in the same region. There may be an *orf3* promoter in this region which escaped detection, or it may be further upstream in a region overlapping *orf1*. Orf3, despite its lack of sequence similarity to anything in the database, was shown to be a novel transcriptional repressor and is analogous to KorA of pIJ101 or TraR of pSN22 in that it acts as a "kil-override" (Stein *et al*., 1989). This is an interesting situation, as the *orf1-orf3* pair in pYS1 is similar to the *tra-korA* pair in pIJ101 in terms of gene organization and both pairs having a related lethal gene, and yet the transcriptional repressor from each pair is distinctly different from one another. Further studies are necessary to better understand the repressor encoded by *orf3* and to narrow down the region to which where it binds. The *orf3-orf1* intergenic region must also be involved in plasmid stability, as pYS1K2, an insertion mutant disrupted within this region, was readily lost from the host cell during curing, while original pYS1 seemed stable after 90 generations.

### 4.1.6 Concluding remarks on pYS1

In summary, I have shown pYS1 is a cryptic, broad-host-range, rolling-circle-type plasmid from *N. aobensis*. While similarities with other Gram-positive plasmids exist, pYS1 has major differences from some well-described plasmids in that it has a more stringent control for lagging strand synthesis and the lethal function of a gene which seems to be involved in stability is controlled by a newly described transcriptional repressor. This study highlights the diversity of plasmids, not only with respect to pYS1 and those previously described, but also the nearly-infinite variety which may exist across the whole prokaryotic spectrum.

## 4.2 Construction of an inducible expression system for the analysis of antimicrobial genes from *Rhodococcus* phages

Tuberculosis causes two million human deaths annually, and its causative agent, *M. tuberculosis*, is one of the most successful bacterial pathogens (Okada and Kobayashi, 2007). The emergence and prevalence of multi-drug resistant strains (Cox *et al*., 2010) call for identification of novel drug targets. At least five out of nine previously isolated *Rhodococcus* phage genes were inhibitory to *M. smegmatis*, providing hope that products encoded by these genes may lead to identification of such targets. To study these genes, it is necessary to have a system where their expression can be induced, which will allow comparison of the host cell before and after induction. Therefore an attempt was made to develop an inducible expression system in *M. smegmatis*.

For the construction and testing of such a system, a gene encoding a known inhibitory product which would serve as a model was required. For this purpose *ecoR*, encoding a restriction endonuclease, was selected. Because pEcoR251 contained this gene under the control of the λ $P_R$ promoter, the plasmid was selected as the backbone of the inducible system. In a λ lysogen, the cI repressor binds to this promoter and prevents expression of *cro* in order to maintain the lysogenic state (Ptashne, 2004). As expression of *cro* will cause the phage to enter the lytic cycle, the cI repressor should not allow any leaky expression to occur. For this reason, a switch consisting of $P_R$ and cI seemed ideal for inducing expression of toxic genes. A temperature-sensitive mutant cI857 was chosen, as this would allow the repressor to lose functionality upon transfer from 30 °C to 42 °C.

Although the heat-inducible switch using pYS13 and pYS13-1 in *E. coli* worked effectively, pYSM13-1 did not allow induced killing of *M. smegmatis* cells in broth. As the switch was functional on solid media, differences in growth conditions between solid and liquid media and their effect on cell physiology need to be considered to identify the potential cause. As one such cause may be the lack of expression of *ecoR* and *cI857* in broth, phages plaquing on *M. smegmatis* were isolated as a source of strong mycobacterial promoters. Future work includes using either pNVCAT or pNVGFP to isolate a phage promoter and test its effect on the inducible system.

### 4.2.1 Concluding remarks on phage genes inhibiting mycobacteria

The existing diversity of phages and their genes (Breitbart *et al*., 2002; Hatfull *et al*., 2010; Kwan *et al*., 2005; Kwan *et al*., 2006) present a potentially enormous pool from which a large number of novel bactericidal mechanisms may be identified. The *Rhodococcus* phage genes which proved inhibitory to *M. smegmatis* may provide us with new susceptibility targets. Further work is necessary to construct the *cI857*-$P_R$ system in *M. smegmatis.* An inducible expression system with a tight genetic switch functional in this organism will allow further characterization of the products encoded by the phage genes.

## 4.3 Analysis of nucleotide bases in *Rhodococcus* phage nucleic acid

While the deoxyribonucleic acids of *Rhodococcus* phages WTS1 and KZN1 were resistant to restriction by many enzymes and attempts to clone their DNA in *E. coli* were unsuccessful in a previous study (Shibayama, 2007), more direct evidence was required to suggest an involvement of modifications or unusual bases. HPLC and TLC were chosen for this purpose, as they could separate individual bases according to their chemical properties and sizes. Molecules behaving differently to the regular bases were then analysed by mass spectrometry to elucidate their chemical structures.

For HPLC, the differences in peak times between a sample containing only the canonical bases (e.g. *E. coli* MM294-4) and those containing hydroxymethylcytosine (T4), uracil (PBS1), or hydroxymethyluracil (SPO1) suggested the credibility of this method in the detection of such bases. However, the times at which peaks for the two latter unusual bases appeared were identical, suggesting that different bases could not always be separated under the conditions used. For the *Rhodococcus* phages, nothing unusual could be detected in their HPLC profiles, which, therefore, was in contradiction with the previous study.

For TLC, spots for the unusual bases from T4, PBS1, and SPO1 were missing, suggesting that they had probably co-migrated with one of the regular bases. The absence of one of the regular bases from each of these phages, however, was clearly reflected by a missing

spot. Therefore, under the conditions used, TLC may only be useful when running a sample in which one of the regular bases is completely or mostly absent.

The detection of ribonucleotides from phages FND1, KG3, and λ indicated the stringency required when purifying phages for subsequent analysis of their nucleic acids. As these three phages were purified in a step gradient of caesium chloride, host RNA contamination was the most likely explanation, as Yamamoto *et al*. (1970) reported the co-precipitation of ribosomes by PEG and their subsequent co-purification with phage particles in caesium chloride. Phage WTS1, which was purified in an equilibrium gradient, did not show any extra spots, suggesting that finer gradients prevent this contamination. The choice of gradient was confirmed to be a crucial factor when nucleic acid from KG3, re-purified in an equilibrium gradient, did not show the extra spots seen earlier.

### 4.3.1 Concluding remarks on phage unusual/modified bases

As shown for T4, PBS1, and SPO1, irregularities in their DNA were reflected by both HPLC and TLC. Therefore a combination of these methods was an effective way to detect unusual bases, although TLC required as much as 20 μg of DNA per run and therefore not efficient. No unusual bases could be detected in the *Rhodococcus* and *Staphylococcus* phages by these methods. The resistance of DNA from WTS1 and KZN1 to restriction digestion and difficulty in cloning them in my previous study remain to be explained: it may be due to a dearth of palindromic sequences recognized by the restriction endonucleases used and/or it is of technical origin.

# 5. APPENDIX

## 5.1. Media

**LB**
3 g tryptone
1.5 g sodium chloride
1.5 g yeast extract
300 ml distilled water

**LB agar**
3 g tryptone
1.5 g sodium chloride
1.5 g yeast extract
4.5 g agar
300 ml distilled water

**Sloppy agar**
1 g tryptone
0.5 g sodium chloride
0.5 g yeast extract
0.75 g agar
100 ml distilled water

**BHI**
11.1 g brain heart infusion
300 ml distilled water

**BHI agar**
11.1 g brain heart infusion
4.5 g agar
300 ml distilled water

***Rhodococcus* regeneration medium**
3 g tryptone
1.5 g yeast extract

0.9 g sodium chloride

30.9 g sucrose

1 g glucose

1 g magnesium chloride hexahydrate

250 ml distilled water

Solution was microwaved to dissolve sucrose. Then, 5.5 g agar was added, autoclaved, and cooled to 60°C, after which the following were added:

10 ml 0.25 M TES pH 7.2

6 ml 1 M calcium chloride

3 ml 0.5% potassium dihydrogen orthophosphate

1.5 ml 10 mg/ml nystatin

0.9 ml 10 mg/ml rifampicin

Plates of constant volume of 25 ml were poured on a horizontal surface.

## 5.2. SOLUTIONS FOR PHAGE PURIFICATION

**SC buffer** 0.58 g sodium chloride

0.2 g calcium chloride hexahydrate

5 ml 1 M tris-hydrochloride pH 8.0

0.5 ml 2% gelatin

95 ml distilled water

**2% Gelatin** 0.02 g gelatin

1 ml distilled water

**Caesium chloride solutions for gradient centrifugation**

(Density values are relative to that of water at 20 °C)

**1.45 g/ml** 3.53 g caesium chloride

5 ml SC buffer

**1.50 g/ml** 4.09 g caesium chloride
5 ml SC buffer

**1.70 g/ml** 6.33 g caesium chloride
5 ml SC buffer

## 5.3. SOLUTIONS FOR DNA PREPARATION

**Solution I** 0.9 g glucose
2.5 ml 1 M tris-hydrochloride pH 8.0
2 ml 0.5 M EDTA
95.5 ml distilled water

**Solution II** 0.8 g sodium hydroxide
1 g SDS
100 ml distilled water

**Solution III** 60 ml 5 M potassium acetate
11.5 ml glacial acetic acid
28.5 ml distilled water

**1 M tris-hydrochloride pH 8.0**

12.1 g tris base, dissolved in distilled water
pH adjusted to 8.0 with hydrochloric acid
Distilled water added to 100 ml

**0.5 M EDTA pH 8.0**

18.61 g EDTA, dissolved in distilled water
pH adjusted to 8.0 with sodium hydroxide
Distilled water added to 100 ml

**TE** 1 ml 1M tris-hydrochloride pH 8.0
2 ml 0.5 M EDTA pH 8.0
100 ml distilled water

**10% SDS in water**

1 g SDS
10 ml distilled water

**10% SDS in TE**

1 g SDS
10 ml TE buffer

**4.5 M potassium acetate pH 6.0**

3.7 g sodium acetate
pH adjusted to 6.0 with acetic acid
Distilled water added to 10 ml

**TE-saturated phenol**

10 g phenol
10 ml TE

**1 M sodium chloride**

5.8 g sodium chloride
100 ml distilled water

**1% Ribonuclease**

0.01 g ribonuclease
1 ml distilled water
Boiled for 10 min before use

**Dialysis buffer**

0.58 g sodium chloride
2.2 g calcium chloride hexahydrate
50 ml 1 M tris-hydrochloride pH 8.0
950 ml distilled water

## 5.4. SOLUTIONS FOR ELECTROPHORESIS

**5× TBE** 54 g tris base
27.5 g boric acid
20 ml 0.5 M EDTA pH 8.0
Distilled water added to 1000 ml

**Agarose gel** 0.8 g (for 0.4%), 1.6 g (for 0.8%), or 2.8 g (for 1.4%) agarose
20 ml 5× TBE
180 ml distilled water

**Running buffer**

20 ml 5× TBE
180 ml distilled water
20 µl 1% EtBr

**1% EtBr** 0.1 g EtBr
10 ml distilled water

**Tracking dye** 0.01 g bromophenol blue
0.01 g xylene cyanol
10 ml 30% glycerol in TE

## 5.5. SOLUTIONS FOR TRANSFORMATION

**Calcium chloride transformation buffer**

4.38 g calcium chloride hexahydrate
2 ml 1 M tris-hydrochloride
198 ml distilled water

**Electroporation buffer**

20.5 g sucrose
30 ml glycerol
170 ml distilled water

**Basal (B) buffer**

10.3 g sucrose
0.025 g potassium sulfate
0.202 g magnesium chloride hexahydrate
10 ml 0.25 M TES pH 7.2

**Protoplast (P) buffer**

2 ml B buffer
20 µl 0.5% potassium dihydrogen orthophosphate
50 µl 1 M calcium chloride

**P-PEG buffer**

0.5 g PEG 6000
1 ml P buffer
PEG sterilized by UV irradiation for 10 min and vortexed vigorously to dissolve.

**20% glucose** 4 g glucose
20 ml distilled water

**0.25 M TES**

17.2 g TES, dissolved in distilled water
pH adjusted to 7.2 with sodium hydroxide
Distilled water added to 300 ml

**1 M calcium chloride**

21.9 g calcium chloride hexahydrate
100 ml distilled water

**0.5% potassium dihydrogen orthophosphate**

0.5 g potassium dihydrogen orthophosphate
100 ml distilled water

**5.6. Solutions for Southern blot and hybridization**

**0.25 M hydrochloric acid**

1.6 ml hydrochloric acid
198.4 ml distilled water

**0.5 N sodium hydroxide**

4 g sodium hydroxide
200 ml distilled water

**20× SSC** 175.3 g sodium chloride
88.2 g sodium citrate
pH adjusted to 7.0 with hydrochloric acid
Distilled water added to 1000 ml

Solution was diluted accordingly with distilled water to give 10×, 5×, 2×, and 0.1×.

**Hybridization solution**

0.02 g N-lauroylsarcosine

0.004 g SDS

0.2 g blocking reagent (Roche)

20 ml 5× SSC

Dissolved at 60 °C for 1 hr.

**0.1% SDS in 2× SSC**

0.1 g SDS

100 ml 2× SSC

**0.1% SDS in 0.1× SSC**

0.1 g SDS

100 ml 0.1× SSC

## 5.7. Solutions for chromatography

**3 M ammonium sulphate**

198.2 g ammonium sulphate

500 ml distilled water

**1 M sodium acetate**

8.2 g sodium acetate

100 ml distilled water

**20 mM ammonium dihydrogen phosphate**

2.3 g ammonium dihydrogen phosphate

1000 ml distilled water

## 5.8. ANTIMICROBIAL AGENTS

Antimicrobial agents are listed in Table 5.1.

**Table 5.1** Antimicrobial agents

| Agent | Stock concentration (mg/ml) | Solvent |
|---|---|---|
| Ampicillin | 100 | 70% ethanol, 30% distilled water |
| Chloramphenicol | 4 or 20 | ethanol |
| Kanamycin | 100 | distilled water |
| Nalidixic acid | 10 | 70% ethanol, 30% distilled water |
| Nystatin | 10 | methanol |
| Rifampicin | 10 | methanol |
| Streptomycin | 20 | distilled water |

# 6. REFERENCES

**Abedon, S. T., T. D. Herschler, and D. Stopar.** 2001. Bacteriophage latent-period evolution as a response to resource availability. Appl. Environ. Microbiol. **67:**4233-4241.

**Ambrosioni, J., D. Lew, and J. Gabrino.** 2010. Nocardiosis: updated clinical review and experience at a tertiary center. Infection **38:**89-97.

**Assaduzzaman, S. M., and K. Sonomoto.** 2009. Lantibiotics: diverse activities and unique modes of action. J. Biosci. Bioeng. **107:**475-487.

**Atchison, M. L., O. Meyuhas, and R. P. Perry.** 1989. Localization of transcriptional regulatory elements and nuclear factor binding sites in mouse ribosomal protein gene rpL32. Mol. Cell. Biol. **9:**2067-2074.

**Bandow, J. E., H. Brotz, L. I. Leichert, H. Labischinski, and M. Hecker.** 2003. Proteomics approach to understanding antibiotic action. Antimicrob. Agents Chemother. **47:**948-955.

**Beaman, B. L., and L. Beaman.** 1994. *Nocardia*: host-parasite relationships. Clin. Microbiol. Rev. **7:**213-264.

**Begg, K. J., S. J. Dewar, and W. D. Donachie.** 1995. A new *Escherichia coli* cell division gene, *ftsK*. J. Bacteriol. **177:**6211-6222.

**Bentley, S. D., K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barell, J. Parkhill, and D. A. Hopwood.** 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature **417:**141-147.

**Bernhardt, T. G., W. D. Roof, and R. Young.** 2000. Genetic evidence that the bacteriophage ϕX174 lysis protein inhibits cell wall synthesis. Proc. Natl. Acad. Sci. USA **97:**4297-4302.

**Billington, S., B. H. Jost, and J. G. Songer.** 1998. The *Arcanobacterium* (*Actinomyces*) *pyogenes* plasmid pAP1 is a member of the pIJ101/pJV1 family of rolling circle replication plasmids. J. Bacteriol. **180:**3233-3236.

**Bingham, R., S. I. Ekunwe, S. Falk, L. Snyder, and C. Kleanthous.** 2000. The major head protein of bacteriophage T4 binds specifically to elongation factor Tu. J. Biol. Chem. **275:**23219-23226.

**Blokpoel, M. C. J., H. N. Murphy, R. O'Toole, S. Wiles, E. S. C. Runn, G. R. Stewart, D. B. Young, and B. D. Robertson.** 2005. Tetracycline-inducible gene regulation in mycobacteria. Nucleic Acids Res. **33:**e22.

**Boe, L., M. F. Gros, H. te Riele, S. D. Ehrlich, and A. Gruss.** 1989. Replication origins of single-stranded-DNA plasmid pUB110. J. Bacteriol. **171:**3366-3372.

**Boman, H. G.** 1995. Peptide antibiotics and their role in innate immunity. Annu. Rev. Immunol. **13:**61-92.

**Bott, K., and B. Strauss.** 1965. The carrier state of *Bacillus subtilis* infected with the transducing bacteriophage SP10. Virology **25:**212-225.

**Bradbury, J.** 2004. My enemy's enemy is my friend. Lancet **363:**624-625.

**Brandon, C., P. M. Gallop, J. Marmur, H. Hayashi, and K. Nakanishi.** 1972. Structure of a new pyrimidine from *Bacillus subtilis* phage SP-15 nucleic acid. Nat. New Biol. **239:**70-71.

**Brasch, M. A., and S. N. Cohen.** 1995. Sequences essential for replication of plasmid pIJ101 in *Streptomyces lividans*. Plasmid **33:**191-197.

**Braun, V., H. Pilsl, and P. Gross.** 1994. Colicins: structures, modes of action, transfer through membranes, and evolution. Arch. Microbiol. **161:**199-206.

**Brede, D. A., T. Faye, M. P. Stierli, G. Dasen, A. Theiler, I. F. Nes, L. Meile, and H. Holo.** 2005. Heterologous production of antimicrobial peptides in *Propionibacterium freudenreichii*. Appl. Environ. Microbiol. **71:**8077-8084.

**Breitbart M., P. Salamon, B. Andersen, J. M. Mahaffy, A. M. Segall, D. Mead, F. Azam, and F. Rohwer.** 2002. Genomic analysis of uncultured marine viral communities. Proc. Natl. Acad. Sci. USA **99:**14250-14255.

**Brown-Elliot, B. A., J. M. Brown, P. S. Conville, and R. J. Wallace JR.** 2006. Clinical and laboratory features of the *Nocardia* spp. based on current molecular taxonomy. Clin. Microbiol. Rev. **19:**259-282.

**Burrus, V., and M. K. Waldor.** 2004. Shaping bacterial genomes with integrative and conjugative elements. Res. Microbiol. **155:**376-386.

**Bush, K.** 2004. Antibacterial drug discovery in the 21st century. Clin. Microbiol. Infect. **10:**S10-S17.

**Carroll, P., D. G. N. Muttucumaru, and T. Parish.** 2005. Use of a tetracycline-inducible system for conditional expression in *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*. Appl. Environ. Microbiol. **71:**3077-3084.

**Centers for Disease Control and Prevention.** 2006. Emergence of *Mycobacterium tuberculosis* with extensive resistance to second-line drugs – worldwide, 2000-2004. MMWR Morb. Mortal. Wkly. Rep. **55:**301 – 305.

**Cerdeno-Tarraga, A. M., A. Efstratiou, L. G. Dover, M. T. Holden, M. Pallen, S. D. Bentley, G. S. Besra, C. Churcher, K. D. James, A. De Zoysa, T. Chillingworth, A. Cronin, L. Dowd, T. Feltwell, N. Hamlin, S. Holroyd, K. Jagels, S. Moule, M. A. Quail, E. Rabbinowitsch, K. M. Rutherford, N. R. Thomson, L. Unwin, S. Whitehead, B. G. Barrell, and J. Parkhill.** 2003. The complete genome sequence and analysis of *Corynebacterium diphtheriae* NCTC13129. Nucleic Acids Res. **31:**6516-6523.

**Chevalier, B. S., and B. L. Stoddard.** 2001. Homing endonucleases: structural and functional insight into the catalysis of intron/intein mobility. Nucleic Acids Res. **29:**3757-3774.

**Chiba, K., Y. Hoshino, K. Ishino, T. Kogure, Y. Mikami, Y. Uehara, and J. Ishikawa.** 2007. Construction of a pair of practical *Nocardia – Escherichia coli* shuttle vectors. Jpn. J. Infect. Dis. **60:**45-47.

**Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeiser, S. Gas, C. E. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. -A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead, and B. G. Barrell.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393:**537-544.

**Corti, M. E., and M. F. Villafañe-Fioti.** 2003. Nocardiosis: a review. Int. J. Infect. Dis. **7:**243-250.

**Cox, H. S., C. McDermid, V. Azevedo, O. Muller, D. Coetzee, J. Simpson, M. Barnard, G. Coetzee, G. van Cutsem, and E. Goemaere.** 2010. Epidemic levels of drug resistant tuberculosis (MDR and XDR-TB) in a high HIV prevalence setting in Khayelitsha, South Africa. PloS One **5:**e13901.

**Cregg, J. M., A. H. Nguyen, and J. Ito.** 1980. DNA modification induced during infection of *Bacillus subtilis* by phage ϕ3T. Gene **12:**17-24.

**Curtis, T. P., W. T. Sloan, and J. W. Scannell.** 2002. Estimating prokaryotic diversity and its limits. Proc. Natl. Acad. Sci. USA **99:**10494-10499.

**Dabbs, E. R., and G. J. Sole.** 1988. Plasmid-borne resistance to arsenate, arsenite, cadmium, and chloramphenicol in a *Rhodococcus* species. Mol. Gen. Genet. **211:**148-154.

**Dale, J. W. and A. Patiki.** 1990. Mycobacterial gene expression and regulation, p. 173-198. *In* J. McFadden (ed.), Molecular biology of the mycobacteria. Surrey University Press, London.

**Datta, I., S. Sau, A. K. Sil, N. C. Mandal.** 2005. The bacteriophage λ DNA replication protein P inhibits the *ori*C DNA- and ATP-binding functions of the DNA replication initiator protein DnaA of *Escherichia coli*. J. Biochem. Mol. Biol. **38:**97-103.

**del Solar, G., G. Kramer, S. Ballester, and M. Espinosa.** 1993. Replication of the promiscuous plasmid pLS1: a region encompassing the minus origin of replication is associated with stable plasmid inheritance. Mol. Gen. Genet. **241:**97-105.

**Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27:**4636-4641.

**Dellagostin, O. A., G. Esposito, L. J. Eales, J. W. Dale, and J. McFadden.** 1995. Activity of mycobacterial promoters during intracellular and extracellular growth. Microbiology **141:**1785-1792.

**Destoumieux-Garzón, D., J. Peduzzi, and S. Rebuffat.** 2002. Focus on modified microcins: structural features and mechanisms of action. Biochimie **84:**511-519.

**Devine, K. M., S. T. Hogan, D. G. Higgins, and D. J. McConnell.** 1989. Replication and segregational stability of *Bacillus* plasmid pBAA1. J. Bacteriol. **171:**1166-1172.

**Ducati, R. G., A. Ruffino-Netto, L. A. Basso, and D. S. Santos.** 2006. The resumption of consumption – a review on tuberculosis. Mem. Inst. Oswaldo Cruz **101:**697-714.

**Duport, C., C. Baysse, and Y. Michel-Briand.** 1995. Molecular characterization of pyocin S3, a novel S-type pyocin from *Pseudomonas aeruginosa*. J. Biol. Chem. **270:**8920 – 8927.

**Ehrlich, M., K. Ehrlich, J. A. Mayo.** 1975. Unusual properties of the DNA from *Xanthomonas* phage XP-12 in which 5-methylcytosine completely replaces cytosine. Biochim. Biophys. Acta. **395:**109-119.

**Ehrt, S., X. V. Guo, C. M. Hickey, M. Ryou, M. Monteleone, L. W. Riley, and D. Schnappinger.** 2005. Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor. Nucleic Acids Res. **33:**e21.

**Euzéby, J. P.** 15 December 2010, accession date. List of prokaryotic names with standing in nomenclature. [Online.] http://www.bacterio.cict.fr/n/nocardia.html.

**Fenwich, M. L.** 1971. The density of ribosomes bearing messenger RNA in phage-infected and normal bacteria. J. Cell Sci. **8:**649-658.

**Fernandez-Gonzalez, C., R. F. Cadenas, M. F. Noirot-Gros, J. F. Martin, and J. A. Gil.** 1994. Characterization of a region of plasmid pBL1 of *Brevibacterium lactofermentum* involved in replication via the rolling circle model. J. Bacteriol. **176:**3154-3161.

**Filice, G. A.** 2005. Nocardiosis in persons with human immunodeficiency virus infection, transplant recipients, and large, geographically defined populations. J. Lab. Clin. Med. **145:**156-162.

**Garcia-Villegas, M. R., F. M. de la Vega, J. M. Galindo, M. Segura, R. H. Buckingham, and G. Guarneros.** 1991. Peptidyl-tRNA hydrolase is involved in λ inhibition of host protein synthesis. EMBO J. **10:**3549-3555.

**Gathogo, E. W. N., A. C. W. Waugh, N. Perić-Concha, M. B. Redpath, and P. F. Long.** 2004. Rapid detection of 'rare' actinomycetes in environmental samples. Biotechnol. Lett. **26:**897-900.

**Ghandi, N. R., A. Moll, A. W. Sturm, R. Pawinski, T. Govender, U. Lallo, K. Zeller, J. Andrews, and G. Friedland.** 2006. Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. Lancet **368:**1575-1580.

**Gillor, O., L. M. Nigro, and M. A. Riley.** 2005. Genetically engineered bacteriocins and their potential as the next generation of antimicrobials. Curr. Pharm. Des. **11:**1067-1075.

**Glupczynski, Y., C. Berhin, M. Janssens, and G. Wauters.** 2006. Determination of antimicrobial susceptibility patterns of *Nocardia* spp. from clinical specimens by Etest. Clin. Microbiol. Infect. **12:**905-912.

**Gold, M., M. Gefter, R. Hausmann, and J. Hurwitz.** 1966. Methylation of DNA. J. Gen. Physiol. **49:**5-28.

**Gowan, B., and E. R. Dabbs.** 1994. Identification of DNA involved in *Rhodococcus* chromosomal conjugation and self-incompatibility. FEMS Microbiol. Lett. **115:**45-50.

**Green, K.** 2007. Screening staphylococcal phage genomes for novel antimicrobials. Honours research report. University of the Witwatersrand, Johannesburg.

**Gupta, N., P. K. Roychoudhury, and J. K. Deb.** 2005. Biotechnology of desulfurization of diesel: prospects and challenges. Appl. Microbiol. Biotechnol. **66:**356-366.

**Hamad, B.** 2010. The antibiotics market. Nat. Rev. Drug Discov. **9:**675-676.

**Harth, G., and M. A. Horwitz.** 1997. Expression and efficient export of enzymatically active *Mycobacterium tuberculosis* glutamine synthetase in *Mycobacterium smegmatis* and evidence that the information for export is contained within the protein. J. Biol. Chem. **272:**22728-22735.

**Hashimoto, M., T. Komori, and T. Kamiya.** 1976. Letter: nocardicin A and B, novel monocyclic beta-lactam antibiotics from a *Nocardia* species. J. Am. Chem. Soc. **12:**3023-3025.

**Hatfull, G. F., M. L. Pedulla, D. Jacobs-Sera, P. M. Cichon, A. Foley, M. E. Ford, R. M. Gonda, J. M. Houtz, A. J. Hryckowian, V. A. Kelchner, S. Namburi, K. V. Pajcini, M. G. Popovich, D. T. Schleicher, B. Z. Simanek, A. L. Smith, G. M. Zdanowicz, V. Kumar, C. L. Peebles, W. R. Jacobs Jr., J. G. Lawrence, and R. W. Hendrix.** 2006. Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. PLoS Genetics **2:**1-13.

**Hatfull, G. F., D. Jacobs-Sera, J. G. Lawrence, W. H. Pope, D. A. Russel, C. -C. Ko, R. J. Weber, M. C. Patel, K. L. Germaine, R. H. Edgar, N. N. Hoyte, C. A. Bowman, A. T. Tantoco, E. C. Paladin, M. S. Myers, A. L. Smith, M. S. Grace, T. T. Pham, M. B. O'Brien, A. M. Vogelsberger, A. J. Hryckowian, J. L. Wynalek, H. Donis-Keller, M. W. Bogel, C. L. Peebles, S. G. Cresawn, and R. W. Hendrix.** Comparative genomic analysis of 60 mycobacteriophage genomes: genome clustering, gene acquisition, and gene size. J. Mol. Biol. **397:**119-143.

**Hattman, S.** 1980. Specificity of the bacteriophage MU $mom^{+}$-controlled DNA modification. J. Virol. **34:**277-279.

**Hernandez-Abanto, S. M., S. C. Woolwine, S. K. Jain, and W. R. Bishai.** 2006. Tetracycline-inducible gene expression in mycobacteria within an animal host using modified *Streptomyces tcp830* regulatory elements. Arch. Microbiol. **186:**459-464.

**Hintermann, G., M. Zatchej, and R. Hütter.** 1985. Cloning and expression of the genetically unstable tyrosine structural gene from *Sreptomyces glaucescens*. Mol. Gen. Genet. **200:**422-432.

**Holt, J. G., N. R. Krieg, P. H. A. Sneath, J. T. Staley, and S. T. Williams.** 2000. Bergey's Manual of Determinative Bacteriology, 9th ed., Lippincott Williams & Wilkins, Philadelphia.

**Hsueh, P.R., C. C. Hung, L. J. Teng, M. C. Yu, Y. C. Chen, H. K. Wang, and K. T. Luth.** 1998. Report of invasive *Rhodococcus equi* infections in Taiwan, with an emphasis on the emergence of multidrug-resistant strains. Clin. Infect. Dis. **27:**370-375.

**Igarashi, K., N. Fujita, and A. Ishihama.** 1991. Identification of a subunit assembly domain in the alpha subunit of *Escherichia coli* RNA polymerase. J. Mol. Biol. **218:**1-6.

**Ikeda, H., J. Ishikawa, A. Hanamoto, M. Shinose, H. Kikuchi, T. Shiba, Y. Sakaki, M. Hattori, and S. Omura.** 2003. Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. Nat. Biotechnol. **21:**526-531.

**Ilyina, T. V., and E. V. Koonin.** 1992. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. Nucleic Acids Res. **20:**3279-3285.

**Ishikawa, J., Y. Niino, and K. Hotta.** 1996. Construction of pRES18 and pRES19, *Streptomyces – Escherichia coli* shuttle vectors carrying multiple cloning sites. FEMS Microbiol. Lett. **145:**113-116.

**Ishikawa, J., A. Yamashita, Y. Mikami, Y. Hoshino, H. Kurita, K. Hotta, T. Shiba, and M. Hattori.** 2004. The complete genomic sequence of *Nocardia farcinica* IFM 10152. Proc. Natl. Acad. Sci. USA **101:**14925-14930.

**Kalinowski J., B. Bathe, D. Bartels, N. Bischoff, M. Bott, A. Burkovski, N. Dusch, L. Eggeling, B. J. Eikmanns, L. Gaigalat, A. Goesmann, M. Hartmann, K. Huthmacher, R. Krämer, M. Linke, A. C. McHardy, F. Meyer, B. Möckel, W. Pfefferle, A. Pühler, D. A. Rey, C. Rückert, O. Rupp, H. Sahn, V. F. Wendisch, I Wiegräbe, and A Tauch.** 2003. The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. J. Biotechnol. **104:**5-25.

**Kalkus, J., M. Reh, and H. G. Schlegel.** 1990. Hydrogen autotrophy of *Nocardia opaca* strains is encoded by linear megaplasmids. J. Gen. Microbiol. **136:**1145-1151.

**Kallen, R. G., M. Simon, and J. Marmur.** 1962. The new occurrence of a new pyrimidine base replacing thymine in a bacteriophage DNA: 5-hydoxymethyl uracil. J. Mol. Biol. **5:**248-250.

**Kaneko, T., K. Katoh, M. Fujimoto, M. Kumagai, J. Tamaoka, and Y. Katayama-Fujimura.** 1986. Determination of the nucleotide composition of a deoxyribonucleic acid by high-performance liquid chromatography of its enzymatic hydrolysate: a review. J. Microbiol. Methods **4:**229-240.

**Kasweck, K. L., M. L. Little, and S. G. Bradley.** 1981. Characteristics of plasmids in *Nocardia asteroides*. Actinomycetes and Related Organisms **16:**57-63.

**Kataoka, M., T. Seki, and T. Yoshida.** 1991. Regulation and function of the *Streptomyces* plasmid pSN22 genes involved in pock formation and inviability. J. Bacteriol. **173:**7975-7981.

**Kataoka, M., Y.-M. Kiyose, Y. Michisuji, T. Horiguchi, T. Seki, and T. Yoshida.** 1994. Complete nucleotide sequence of the *Streptomyces* plasmid, pSN22; genetic organization and correlation with genetic properties. Plasmid **32:**55-69.

**Katayama-Fujimura, Y., Y. Komatsu, H. Kuraishi, and T. Kaneko.** 1984. Estimation of DNA base composition by high performance liquid chromatography of its nuclease P1 hydrolysate. Agric. Biol. Chem. **48:**3169-3172.

**Kawamura, F., T. Mizukami, H. Shimotsu, H. Anzai, H. Takahashi, and H. Saito.** 1981. Unusually infrequent cleavage with several endonucleases and physical map construction of *Bacillus subtilis* bacteriophage ϕ1 DNA. J. Virol. **37:**1099-1102.

**Kędzierska, B., M. Glinkowska, A. Iwanichi, M. Obuchowski, P. Sojka, M. S. Thomas, G. Węgrzyn.** 2003. Toxicity of the bacteriophage λ *cII* gene product to

*Escherichia coli* arises from inhibition of host cell DNA replication. Virology **313:**622-628.

**Kendall, K. J., and S. N. Cohen.** 1987. Plasmid transfer in *Streptomyces lividans*: Identification of a *kil-kor* system associated with the transfer region of pIJ101. J. Bacteriol. **169:**4177-4183.

**Kendall, K. J., and S. N. Cohen.** 1988. Complete nucleotide sequence of the *Streptomyces lividans* plasmid pIJ101 and correlation of the sequence with genetic properties. J. Bacteriol. **170:**4634-4651.

**Khan, S. A.** 2004. Rolling-circle replication, p. 63-78. *In* B. E. Funnel and G. J. Phillips, (ed.), Plasmid Biology, ASM Press, Washington, DC.

**Khan, S. A.** 2005. Plasmid rolling-circle replication: highlights of two decades of research. Plasmid **53:**126-136.

**Kirnos, M. D., I. Y. Khudyakov, N. I. Alexandruschinka, and B. F. Vanyushin.** 1977. 2-aminoadenine is an adenine substituting for a base in S-2L cyanophage DNA. Nature **270:**369-370.

**Kosono, S., M. Kataoka, T. Seki, and T. Yoshida.** 1996. The TraB protein, which mediates the intermycelial transfer of the *Streptomyces* plasmid pSN22, has functional NTP-binding motifs and is localized to the cytoplasmic membrane. Mol. Microbiol. **19:**397-405.

**Krajewska, E., and D. Shugar.** 1971. Photochemical transformation of 5-alkyluracils and their nucleosides. Science **173:**435-437.

**Kramer, M. G., S. A. Khan, and M. Espinosa.** 1997. Plasmid rolling circle replication: identification of the RNA polymerase-directed primer RNA and requirement of DNA polymerase I for lagging strand synthesis. EMBO J. **16:**5784-5795.

**Kramer, M. G., M. Espinosa, T. K. Misra, and S. A. Khan.** 1998a. Lagging strand replication of rolling-circle plasmids: specific recognition of the *ssoA*-type origins in different gram-positive bacteria. Proc. Natl. Acad. Sci. USA **95:**10505-10562.

**Kramer, M. G., S. A. Khan, and M. Espinosa.** 1998b. Lagging strand replication from *ssoA* origin of plasmid pMV158 in *Streptococcus pneumoniae*: *in vivo* and *in vitro* influences of mutations in two conserved *ssoA* regions. J. Bacteriol. **180:**83-89.

**Krüger, D. A., and T. A. Bickle.** 1983. Bacteriophage survival: multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. Microbiol. Rev. **47:**345-360.

**Kurtböke, D. I.** 2010. Exploitation of phage battery in the search for bioactive actinomycetes. Appl. Microbiol. Biotechnol. Dec 1 [Epub ahead of print].

**Kutter, E., B. Guttman, and K. Carlson.** 1994a. The transition from host to phage metabolism after T4 infection, p.343-346. *In* Karam, J., J. W. Drake, K. N. Kreuzer, G. Mosig, D. H. Hall, F. A. Eiserling, L. W. Black, E. K. Spicer, E. Kutter, K. Carlson, and S. Miller, (ed.), Molecular biology of bacteriophage T4, ASM Press, Washington, D.C.

**Kutter, E., E. Kellenberger, K. Carlson, S. Eddy, J. Neitzel, L. Messinger, J. North, and B. Guttman.** 1994b. Effects of bacterial growth conditions and physiology on T4 infection, p.406-418. *In* Karam, J., J. W. Drake, K. N. Kreuzer, G. Mosig, D. H. Hall, F. A. Eiserling, L. W. Black, E. K. Spicer, E. Kutter, K. Carlson, and S. Miller, (ed.), Molecular biology of bacteriophage T4, ASM Press, Washington, D.C.

**Kwan, T., J. Liu, M. DuBow, P. Gros, and J. Pelletier.** 2005. The complete genomes and proteomes of 27 *Staphylococcus aureus* bacteriophages. Proc. Natl. Acad. Sci. USA **102:**5174-5179.

**Kwan, T., J. Liu, M. DuBow, P. Gros, and J. Pelletier.** 2006. Comparative genome analysis of 18 *Pseudomonas aeruginosa* bacteriophages. J. Bacteriol. **188:**1184-1187.

**Lee, C., J. Kim, S. G. Shin, and S. Hwang.** 2006. Absolute and relative QPCR quantification of plasmid copy number in *Escherichia coli*. J. Biotechnol. **123:**273-280.

**Lehman, I. R., and E. A. Pratt.** 1960. On the structure of the glucosylated hydroxymethylcytosine nucleotides of coliphages T2, T4, and T6. J. Biol. Chem. **235:**3254-3259.

**Liu, J., M. Dehbi, G. Moeck, F. Arhin, P. Bauda, D. Bergeron, M. Callejo, V. Ferreti, N. Ha, T. Kwan, J. McCarty, R. Srikumar, D. Williams, J. Wu, P. Gros, J. Pelletier, and M. DuBow.** 2004. Antimicrobial drug discovery through bacteriophage genomics. Nat. Biotechnol. **22:**185-191.

**Liu, Y.-T., C.-M. Su, C.-H. Lee, M.-J. Sui, Y.-H. Chang, W.-P. Lin, W.-T. Wu, and C.-Y. Chen.** 2000. Cloning and characterization of the replicon of the *Nocardia italica* plasmid, pNI100. Plasmid **43:**223-229.

**Livermore, D. M.** 2004. The need for new antibiotics. Clin. Microbiol. Infect. **10:**S1-S9.

**Loeffler, J. M., D. Nelson, and V. A. Fischetti.** 2001. Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. Science **294:**2170-2172.

**Lukashin, A., and M. Borodovsky.** 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res. **26:**1107-1115.

**Marchand, I., A. W. Nicholson, and M. Dreyfus.** 2001. Bacteriophage T7 protein kinase phosphorylates RNase E and stabilizes mRNAs synthesized by T7 RNA polymerase. Mol. Microbiol. **42:**767-776.

**Mayer, J. E., and M. Schweiger.** 1983. RNase III is positively regulated by T7 protein kinase. J. Biol. Chem. **258:**5340-5343.

**Mendiola, M. V., and F. de la Cruz.** 1992. IS91 transposase is related to the rolling-circle-type replication proteins of the pUB110 family of plasmids. Nucleic Acids Res. **20:**3521.

**McNeil, M.M., and J. M. Brown.** 1994. The medically important aerobic actinomycetes: epidemiology and microbiology. Clin. Microbiol. Rev. **7:**357-417.

**Miller, E. S., E. Kutter, G. Mosig, F. Arisaka, T. Kunisawa, and W. Rüger.** 2003. Bacteriophage T4 genome. Microbiol. Mol. Biol. Rev. **67:**86-156.

**Movahedzadeh, F., M. J. Colston, and E. O. Davis.** 1997. Determination of DNA sequences required for regulated *Mycobacterium tuberculosis* RecA expression in response to DNA-damaging agents suggests that two modes of regulation exist. J. Bacteriol. **179:**3509-3518.

**Mulder, M. A., H. Zappe, and L. M. Steyn.** 1997. Mycobacterial promoters. Tuber. Lung Dis. **78:**211-223.

**Muth, G., M. Farr, V. Hartmann, and W. Wohlleben.** 1995. *Streptomyces ghanaensis* plasmid pSG5: nucleotide sequence analysis of the self-transmissible minimal replicon and characterization of the replication mode. Plasmid **33:**113-126.

**Nakashima, N., and T. Tamura.** 2004. Isolation and characterization of rolling-circle-type plasmid from *Rhodococcus erythropolis* and application of the plasmid to multiple-recombinant-protein expression. Appl. Environ. Microbiol. **70:**5557-5568.

**Nechaev, S., and K. Severinov.** 1999. Inhibition of *E. coli* RNA polymerase by bacteriophage T7 gene 2 protein. J. Mol. Biol. **289:**815-826.

**New England Biolabs.** 2009. Restriction endonucleases. New England Biolabs, Ipswich, MA.

**Nikolskaya, I. I., M. G. Tediashvili, M. B. Vasielieva, T. G. Chanishvili, and S. S. Debov.** 1979. Specificity and functions of guanine methylase of *Shigella sonnei* DDVI phage. Biochim. Biophys. Acta **561:**232-239.

**Noirot-Gross, M. F., Bidnenko, V., and S. D. Ehrlich.** 1994. Active site of the replication protein in the rolling circle plasmid pC194. EMBO J. **13:**4412-4420.

**Nordmann, P., E. Rouveix, M. Guenounou, and M. H. Nicolas.** 1992. Pulmonary abscess due to a rifampin and fluoroquinolone resistant *Rhodococcus equi* strain in a HIV infected patient. Eur. J. Clin. Microbiol. Infect. Dis. **11:**557-558.

**Normand, P., P. Lapierre, L. S. Tisa, J. P. Gogarten, N. Alloisio, E. Bagnarol, C. A. Bassi, A. M. Berry, D. M. Bickhart, N. Choisne, A. Couloux, B. Cournoyer, S. Cruveiller, V. Daubin, N. Demange, M. P. Francino, E. Goltsman, Y. Huang, O. R. Kopp, L. Labarre, A. Lapidus, C. Lavire, J. Marechal, M. Martinez, J. E.**

**Mastronunzio, B. C. Mullin, J. Niemann, P. Pujic, T. Rawnsley, Z. Rouy, C. Schenowitz, A. Sellstedt, F. Tavares, J. P. Tomkins, D. Vallenet, C. Valverde, L. G. Wall, Y. Wang, C. Medigue, and D. R. Benson.** 2007. Genome characteristics of facultatively symbiotic *Frankia* sp. strains reflect host range and host plant biogeography. Genome Res. **17:**7-15.

**O'Gaora, P., S. Barnini, C. Hayward, E. Filley, G. Rook, D. Young, and J. Thole.** 1997. Mycobacteria as immunogens: development of expression vectors for use in multiple mycobacterial species. Med. Princ. Prac. **6:**91-96.

**Okada, M., and K. Kobayashi.** 2007. Recent progress in mycobacteriology. Kekkaku **82:**783-799.

**Otsuka, Y., K. Ohkusu, Y. Kawamura, S. Baba, T. Ezaki, and S. Kimura.** 2006. Emergence of multidrug-resistant *Corynebacterium striatum* as a nosocomial pathogen in long-term hospitalized patients with underlying diseases. Diagn. Microbiol. Infect. Dis. **54:**109-114.

**Pag, U., and G. Sahl.** 2002. Multiple activities in lantibiotics – models for the design of novel antibiotics. Curr. Pharm. Des. **8:**815-833.

**Papagianni, M.** 2003. Ribosomally synthesized peptides with antimicrobial properties: biosynthesis, structure, function, and applications. Biotechnol. Adv. **21:**465-499.

**Papavinasasundaram, K. G., F. Movahedzadeh, J. T. Keer, N. G. Stoker, M. J. Colston, and E. O. Davis.** 1997. Mycobacterial *recA* is cotranscribed with a potential regulatory gene called *recX*. Mol. Microbiol. **24:**141-153.

**Parish, T., and N. G. Stoker.** 1998. Methods in Molecular Biology, vol. 101: Mycobacteria Protocols Humana Press, Totowa, N. J.

**Parret, A., and R. de Mot.** 2000. Novel bacteriocins with predicted tRNase and pore-forming activities in *Pseudomonas aeruginosa* PAO1. Mol. Microbiol. **35:**472-475.

**Pashley, C. and N. G. Stoker.** 2000. Plasmids in mycobacteria, p. 55-68. *In* G. F. Hatfull and W. R. Jacobs Jr. (ed.), Molecular genetics of mycobacteria. ASM Press, Washington D.C.

**Pedulla, M. L., M. E. Ford, J. M. Houtz, T. Karthikeyan, C. Wadsworth, J. A. Lewis, D. Jacobs-Sera, J. Falbo, J. A. Gross, N. R. Pannunzio, W. Brucker, V. Kumar, J. Kandasamy, L. Keenan, S. Bardoarov, J. Kriakov, J. G. Lawrence, W. R. Jacobs Jr., R. W. Hendrix, and G. F. Hatfull.** 2003. Origins of highly mosaic mycobacteriophage genomes. Cell **113:**171-182.

**Pirisi, A.** 2000. Phage therapy – advantage over antibiotics? Lancet **356:**1418.

**Pons, A.-M., I. Lanneluc, G. Cottenceau, and S. Sable.** 2002. New developments in non-post translationally modified microcins. Biochimie **84:**531-537.

**Powers, J. H.** 2004. Antimicrobial drug development – the past, the present, and the future. Clin. Microbiol. Infect. **10:**S23-S31.

**Pridmore, R. D.** 1987. New and versatile cloning vectors with kanamycin-resistance marker. Gene **56:**309-312.

**Projan, S. J., and D. M. Shlaes.** 2004. Antibacterial drug discovery: is it all downhill from here? Clin. Microbiol. Infect. **10:**S18-S22.

**Provost, F., M. V. Blanc, B. L. Beaman, and P. Boiron.** 1996. Occurrence of plasmids in pathogenic strains of *Nocardia*. J. Med. Microbiol. **45:**344-348.

**Ptashne, M.** 2004. The master elements of control, p.11-29. *In* M. Ptashne (ed.), A genetic switch, 3rd ed., Phage lambda revisited. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

**Pushnova, E. A., M. Geier, and Y. S. Zhu.** 2000. An easy and accurate agarose gel assay for quantitation of bacterial plasmid copy numbers. Anal. Biochem. **284:**70-76.

**Putman, M., H. W. van Veen, and W. N. Konings.** 2000. Molecular properties of bacterial multidrug transporters. Microbiol. Mol. Biol. Rev. **64:**672-693.

**Qasem, J. A., Z. U. Khan, A. S. Mustafa, and T. D. Chugh.** 1999. Polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) analysis and plasmid profile of soil and clinical isolates of *Nocardia*. Microbiol. Res. **154:**157-165.

**Quan, S., and E. R. Dabbs.** 1993. Nocardioform arsenic resistance plasmid characterization and improved *Rhodococcus* cloning vectors. Plasmid **29:**74-79.

**Quan, S., H. Venter, E. R. Dabbs.** 1997. Ribosylative inactivation of rifampin by *Mycobacterium smegmatis* is a principal contributor to its low susceptibility to this antibiotic. Antimicrob. Agents Chemother. **41:**2456-2460.

**Ranes, M. G., J. Rauzier, M. Lagranderie, M. Gheorghiu, and B. Gicquel.** 1990. Functional analysis of pAL5000, a plasmid from *Mycobacterium fortuitum*: construction of a "mini" *Mycobacterium-Escherichia coli* shuttle vector. J. Bacteriol. **172:**2793-2797.

**Reeve, J. N., E. Amann, R. Tailor, U. Günthert, K. Scholz, and T. A. Trautner.** 1980. Unusual behaviour of SPO1 DNA with respect to restriction and modification enzymes recognizing the sequence 5'-G-G-C-C. Mol. Gen. Genet. **178:**229-231.

**Robertson, E. S., L. A. Aggison, and A. W. Nicholson.** 1994. Phosphorylation of elongation factor G and ribosomal protein S6 in bacteriophage T7-infected *Escherichia coli*. Mol. Microbiol. **11:**1045-1057.

**Rohwer, F.** 2003. Global phage diversity. Cell **113:**141.

**Salinas-Carmona, M. C.** 2000. *Nocardia brasiliensis*: from microbe to human and experimental infections. Microbes Infect. **2:**1373-381.

**Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Laboratory, Cold Spring Harbour, N.Y.

**Sampath, A., and C. R. Stewart.** 2004. Roles of genes 44, 50, and 51 in regulating gene expression and host takeover during infection of *Bacillus subtilis* by bacteriophage SPO1. J. Bacteriol. **186:**1785-1792.

**Seegers, J. F. M. L., A. C. Zhao, W. J. J. Meijer, S. A. Khan, G. Venema, and S. Bron.** 1995. Structural and functional analysis of the single-strand origin of replication from the lactococcal plasmid pWV01. Mol. Gen. Genet. **249:**43-50.

**Sensfuss, C., M. Reh, and H. G. Schlegel.** 1986. No correlation exists between the conjugative transfer of the autotrophic character and that of plasmids in *Nocardia opaca* strains. J. Gen. Microbiol. **132:**997-1007.

**Sergueev, K., D. Yu, S. Austin, and D. Court.** 2001. Cell toxicity caused by products of the *pL* operon of bacteriophage lambda. Gene **272:**227-235.

**Servin-Gonzalez, L.** 1993. Relationship between the replication functions of *Streptomyces* plasmids pJV1 and pIJ101. Plasmid **30:**131-140.

**Servin-Gonzalez, L., A. Shampieri III, J. Cabello, L. Galvan, V. Juarez, and C. Castro.** 1995. Sequence and functional analysis of the *Streptomyces phaeochromogenes* plasmid pJV1 reveals a modular organization of *Streptomyces* plasmids that replicate by rolling circle. Microbiology **141:**2499-2510.

**Severinova, E., and K. Severinov.** 2006. Localization of the *Escherichia coli* RNA polymerase β' subunit residue phosphorylated by bacteriophage T7 kinase gp0.7. J. Bacteriol. **188:**3470-3476.

**Schuch, R., D. Nelson, and V. A. Fischetti.** 2002. A bacteriolytic agent that detects and kills *Bacillus anthracis*. Nature **418:**825-826.

**Sharma, M. S., R. L. Ellis, and D. M. Hinton.** 1992. Identification of a family of bacteriophage T4 genes encoding proteins similar to those present in group I introns of fungi and phage. Proc. Natl. Acad. Sci. USA **89:**6658-6662.

**Sharp, P.M., and W. H. Li.** 1987. The codon adaptation index – a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. **15:**1281-1295.

**Shen, M., P. Fang, D. Xu, Y. Zhang, W. Cao, Y. Zhu, J. Zhao, and Z. Qin.** 2006. Replication and inheritance of *Nocardia* plasmid pC1. FEMS Microbiol. Lett. **261:**47-52.

**Sherratt, D. J., L. K., Arciszewska, E. Crozat, J. E. Graham, and I. Grainge.** 2010. The *Escherichia coli* DNA translocase FtsK. Biochem. Soc. Trans. **38:**395-398.

**Shibayama, Y.** 2004. Isolation and characterization of novel bacteriophages as potential biocontrol agents against *Rhodococcus equi*. Honours research report. University of the Witwatersrand, Johannesburg.

**Shibayama, Y.** 2006. Rapid screening for antimicrobial genes in novel nocardiophages. MSc dissertation. University of the Witwatersrand, Johannesburg.

**Solyanikova, I. P., V. M. Travkin, D. O. Rybkina, E. G. Plotnikova, and L. A. Golovleva.** 2008. Variability of enzyme system of *Nocardioform* bacteria as a basis of their metabolic activity. J. Environ. Sci. Health B **43:**241-252.

**Snapper, S. B., R. E. Melton, S. Mustafa, T. Kieser, W. R. Jacobs, Jr.** 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol. Microbiol. **4:**1911-1919.

**Stein, D. S., K. J. Kendall, and S. N. Cohen.** 1989. Identification and analysis of transcriptional regulatory signals for the *kil* and *kor* loci of *Streptomyces* plasmid pIJ101. J. Bacteriol. **171:**5768-5775.

**Stewart, C. R., I. Gaslightwala, K. Hinata, K. Krolikowski, D. S. Needleman, A. S. -Y. Peng, M. A. Peterman, A. Tobias, and P. Wei.** 1998. Genes and regulatory sites of the "host-takeover module" in the terminal redundancy of *Bacillus subtilis* bacteriophage SPO1. Virology **246:**329-340.

**Stolt, P., and N. G. Stoker.** 1996. Functional definition of regions necessary for replication and incompatibility in the *Mycobacterium fortuitum* plasmid pAL5000. Microbiology **142:**2795-2802.

**Stone, R.** 2002. Stalin's forgotten cure. Science **298:**728-732.

**Sussman, R., and F. Jacob. 1962.** Sur un systeme de repression thermosensible chez le bacteriophage d'*Escherichia coli*. Compt. Rend. Acad. Sci. **254:**1517-1526.

**Suzuki, I., T. Seki, and T. Yoshida.** 1997. Nucleotide sequence of a nicking site of the *Streptomyces* plasmid pSN22 replicating by the rolling circle mechanism. FEMS Microbiol. Lett. **150:**283-288.

**Takahashi, I., and J. Marmur.** 1963. Replacement of thymidylic acid by deoxyuridylic acid in the deoxyribonucleic acid of a transducing phage for *Bacillus subtilis*. Nature **197:**794-795.

**Takahashi, Y., and S. Ōmura.** 2003. Isolation of new actinomycete strains for the screening of new bioactive compounds. J. Gen. Appl. Microbiol. **49:**141-154.

**Tamaoka, J., and K. Komagata.** 1984. Determination of DNA base composition by reverse-phase high performance liquid chromatography. FEMS Microbiol. Lett. **25:**125-128.

**Tanaka, Y., H. Komaki, K. Yazawa, Y. Mikami, A. Nemoto, T. Tojyo, K. Kadowaki, H. Shigemori, and J. Kobayashi.** 1997. Brasilinode A, a new macrolide

antibiotic produced by *Nocardia brasiliensis*: producing strain, isolation and biological activity. J. Antibiot. (Tokyo) **50:**1036-1041.

**Thacker, P. D.** 2003. Set a microbe to kill a microbe. JAMA **290:**3183.

**Thomas, C. M.** 2000. Paradigms of plasmid organization. Mol. Microbiol. **37:**485-491.

**Thomson, C. J., E. Power, H. Ruebsamen-Waigmann, and H. Labischinski.** 2004. Antibacterial research and development in the 21st century – an industry perspective of the challenges. Curr. Opin. Microbiol. **7:**445-450.

**Tiemann, B., R. Depping, and W. Ruger.** 1999. Overexpression, purification, and partial characterization of ADP-ribosyltransferases modA and modB of bacteriophage T4. Gene Express. **8:**187-196.

**Valdez-Cruz, N. A., L. Caspeta, N. O. Pérez, O. T. Ramírez, M. A. Trujillo-Roldán.** 2010. Production of recombinant proteins in *E. coli* by the heat inducible expression system based on the phage lambda pL and/or pR promoters. Microb. Cell Fact. **9:**18.

**Ventura, M., C. Canchaya, A. Tauch, G. Chandra, G. F. Fitzgerald, K. F. Chater, and D. van Sinderen.** 2007. Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum. Microbiol. Mol. Biol. Rev. **71:**495-548.

**von Bargen, K., and A. Haas.** 2009. Molecular and infection biology of the horse pathogen *Rhodococcus equi*. FEMS Microbiol. Rev. **33:**870-891.

**Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay.** 1982. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1:**945-951.

**Watve M. G., R. Tickoo, M. M. Jog, B. D. Behole.** 2001. How many antibiotics are produced by the genus *Streptomyces*? Arch. Microbiol. **176:**386-390.

**Wei, P., and C. R. Stewart.** 1993. A cytotoxic early gene of *Bacillus subtilis* bacteriophage SPO1. J. Bacteriol. **175:**7887-7900.

**Werbowy, K., H. Cieśliński, and J. Kur.** 2009. Characterization of a cryptic plasmid pSFKW33 from *Shewanella* sp. 33B. Plasmid **62:**44-49.

**World Health Organization.** Guidelines for the programmatic management of drug-resistant tuberculosis. WHO/HTM/TB/2006.361. Geneva, Switzerland: WHO, 2006.

**Wu, G., L. Nie, and W. Zhang.** 2006. Predicted highly expressed genes in *Nocardia farcinica* and the implication for its primary metabolism and nocardial virulence. Antonie van Leeuwenhoek **89:**135-146.

**Wu, L. J., P. J. Lewis, R. Allmansberger, P. M. Hauser, and J. Errington.** 1995. A conjugation-like mechanism for prespore chromosome partitioning during sporulation in *Bacillus subtilis*. Genes Dev. **9:**1316-1326.

**Wyatt, G. R., and S. S. Cohen.** 1953. The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. Biochem. J. **55:**774-782.

**Xia, H.-Y., Y.-Q. Tian, R. Zhang, K.-C. Lin, and Z.-J. Qin.** 2006. Characterization of *Nocardia* plasmid pXT107. Acta Biochim. Biophys. Sin. **38:**620-624.

**Yamamoto, K. R., B. M. Alberts, R. Benzinger, L. Lawhorne, and G. Treiber.** 1970. Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification. Virology **40:**734-744.

**Yao, W., Y. Yang, and J. Chiao.** 1994. Cloning vector system for *Nocardia* spp. Current Microbiology **29:**223-227.

**Yazawa, K., Y. Mikami, A. Maeda, M. Akao, N. Morisaki and S. Iwasaki.** 1993. Inactivation of rifampin by *Nocardia brasiliensis*. Antimicrob. Agents Chemother. **37:**1313-1317.

**Yu, Y. -T. N., and L. Snyder.** 1994. Translational elongation factor Tu cleaved by a phage-exclusion system. Proc. Natl. Acad. Sci. USA **91:**802-806.

**Zabeau, M., and K. K. Stanley.** 1982. Enhanced expression of the cro-β-galactosidase fusion under the control of the PR promoter of the bacteriophage lambda. EMBO J. **1:**1217-1224.

**Zilling, W., H. Fujiki, W. Blum, D. Janekovic, M. Schweiger, H. J. Rahmsdorf, H. Ponta, and M. Hirsch-Kauffmann.** 1975. *In vivo* and *in vitro* phosphorylation of DNA-dependent RNA polymerase of *Escherichia coli* by bacteriophage-T7-induced protein kinase. Proc. Natl. Acad. Sci. USA **7:**2506-2510.

**Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. **31:**3406-3415.