Biochemical characterization of the *Mycobacterium smegmatis* rifampicin ADP-ribosylating enzyme ARR

Helen Susannah Letseka

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Master of Science.

Johannesburg, 2010

DECLARATION

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

_____

Helen Susannah Letseka

07 day of October, 2010.

DEDICATION

For my parents Alan and Louise Webb

What a beautiful piece of heartache this has all turned out to be.
Lord knows we've learned the hard way all about healthy apathy.
I use these words pretty loosely.
There is so much more to life than words
Over the Rhine

ACKNOWLEDGEMENTS

I would like to thank my family (the Webbs and the Letsekas) for emotional and financial support during my studies. Also the NRF and Wits Post Graduate Merit Award for funding.

A special thanks to Jonathan Burke who started this work with me but sadly moved on to greener pastures before it was complete, my supervisors Dr Yasien Sayed and Prof Eric Dabbs and also to Prof Heinrich Dirr for allowing me the opportunity to work in his lab.

I wish to acknowledge the following people's invaluable contribution to this work: the members of the PSFRU, Abe from the Electron Microscopy unit at Wits, Rodney Hull for help with light microscopy, Simon Thome, without whom this thesis would have been lost forever and my husband Ntutu.

Table of Contents

Abstract

ARR is an antibiotic resistance protein found in *Mycobacterium smegmatis* with several homologues in other species and genera. ARR inactivates rifampicin through the transfer of an ADP-ribose group from NAD^+ to C_{23} of rifampicin. This work aimed to characterize ARR both biochemically and functionally. ARR was heterologously expressed and purified to a homogenous level. Using far-UV circular dichroism it was found that ARR has a mixed alpha helical and beta sheet secondary structure. It was also shown that ARR retains a high proportion of this structure after heating from 10 °C to 85 °C. Fluorescence spectroscopy showed the two tryptophans (W59 and W107) in ARR to be solvent exposed. These findings are congruous with the crystal structure of ARR. Study of cell wall morphology revealed that expression of ARR led to increased cell length in *E. coli* and branching in *M. smegmatis.* We suggest other possible targets for ARR in *M. smegmatis* based on the *bld* pathway in *S. coelicolor* which involves an ADP-ribosyltransfer reaction leading to changes in cell wall morphology. These targets are extracellular binding protein, sugar binding lipoprotein and D-xylose binding periplasmic protein. These targets suggest a role for ARR in nutrient sensing and stress response.

Frequently used abbreviations

ADP Adenosine diphosphate

AIDS Acquired Immune Deficiency Syndrome

ART Mono ADP-ribosyltransferase

ATP Adenosine triphosphate

Bp Base pair

C Carbon

$CaCl_2$ Calcium chloride

CD Circular dichroism

cm Centimetre

Da Daltons

DNA Deoxyribonucleic acid

dNTPs Deoxynucleotide triphosphates

DTT Dithiothreitol

EDTA Ethylenediaminetetraacetic acid

EtBr Ethidium bromide

g Gram

HIV Human Immuno-deficiency Virus

His-tag Histidine-tagged

IPTG Isopropyl-β-D-thiogalactoside

kb Kilobase

kDa Kilo Daltons

LA Luria-Bertani Agar

LB Luria-Bertani Broth

µg Microgram

µl Microlitre

M Molar

MES 2-(*N*-morpholino)ethanesulfonic acid

mg Milligram

ml Millilitre

mm Millimetre

mM Millimolar

NaCl Sodium chloride

NAD Nicotinamide adenine dinucleotide

Ni-NTA Nickel-nitrilotriacetic acid

ng Nanogram

nm Nanometer

PAGE Polyacrylamide gel electrophoresis

PARP Poly ADP-ribose polymerase

RNA Ribonucleic acid

rpm Revolutions per minute

SDS Sodium dodecyl sulfate

TB Tuberculosis

Tris Tris(hydroxymethyl)-aminomethane

UV Ultra violet

WHO World Health Organisation

List of Figures

List of Tables

1 Introduction

1.1 Mycobacteria

Mycobacteria are acid fast Gram positive filamentous bacteria and are the only member of the family Mycobacteriaceae within the order *Actinomycetes* (NCBI). The genus *Mycobacterium* is made up of more than 30 species and numerous subspecies which are closely related to *Nocardia* and *Corynebacterium*. Like these organisms the mycobacterial genome has a high G-C content. The most distinguishable feature of mycobacteria is their ability to produce large amounts of long chain fatty acids known as mycolic acids and incorporate them into their cell walls. The genus can be divided into the fast growing mycobacteria (e.g. *M. smegmatis*) and the slow growing mycobacteria such as *M. tuberculosis* (Springer *et al.*, 1996).

M. smegmatis was first isolated from human smegma (genital secretion). Although *M. smegmatis* is considered to generally be non-pathogenic it has been implicated in several human infections especially those around wounds after trauma (Newton *et al.*, 1993). Geographically, *M. smegmatis* is widely distributed with isolates reported in the United States of America, Canada, Australia, Switzerland and Russia (Brown *et al.*, 1999; Pennekamp *et al.*, 1997; Wallace *et al.*, 1988). *M. smegmatis* is an ideal mycobacterial model as it is non-pathogenic and fast growing (Snapper *et al.*, 1990); while some pathogenic species of *Mycobacterium* take six to seven weeks to culture, *M. smegmatis* can be grown in four days.

1.1.1 Mycobacterial cell walls

Mycobacterial cell walls are made up of three types of polymers that are covalently linked. These polymers are peptidoglycan, arabinogalactan and mycolic acids (see Figure 1.1). The peptidoglycan and arabinogalactan form what is known as the AG complex which is linked by phosphodiester bonding to the mycolic acids (Petit *et al.*, 1969). Nearly 60% of the mycobacterial cell wall is made up of lipids (Wang *et al.*, 2000); this is highly unusual in bacterial cells. Mycolic acids are unique to

Mycobacteria and closely related genera including *Nocardia, Rhodococcus* and *Corynebacterium* (Frieden, 1993).

Mycolic acids are α-alkyl-β-hydroxy fatty acids which are long and complex (See Figure 1.2). Mycolic acids have a very low fluidity and make up the inner leaflet of the lipid bilayer of mycobacterial cell walls. The presence of mycolic acids may explain the extremely low permeability of mycobacterial cell walls (Wang *et al.*, 2000). This low permeability partially accounts for the low level of antibiotic susceptibility in many mycobacterial species. Mycolic acids also help in disease persistence in mycobacterial infections by shielding the bacteria from stress during infection (Barry and Mduli, 1996).

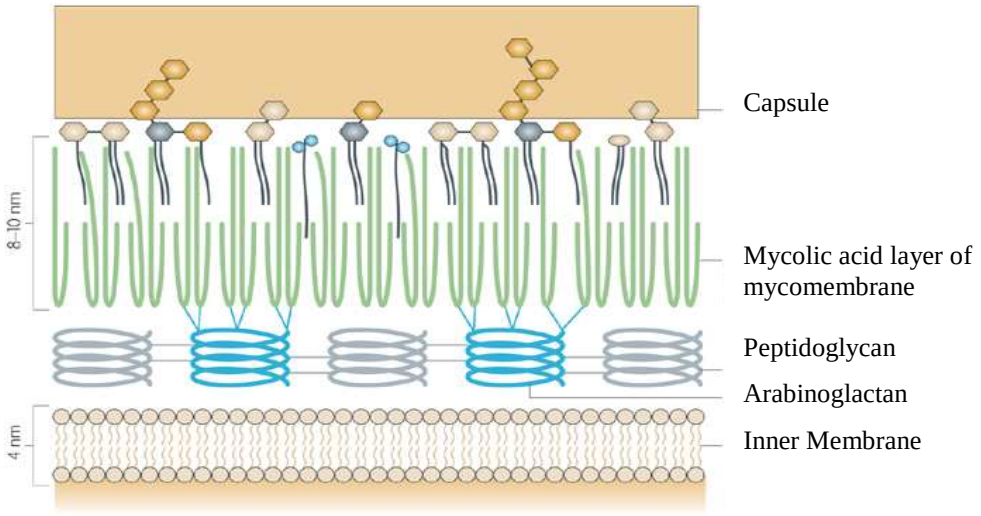


Figure 1.1 Schematic representation of the proposed structure of *M. tuberculosis* cell wall (Abdallah *et al.*, 2007). Note the mycolic acid layer which makes the membrane thick and hardy. Typically, non-pathogenic mycobacteria do not have a capsule.

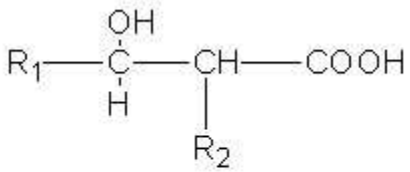


Figure 1.2 Generalized mycolic acid structure (Asselineau and Lederer, 1950).

1.1.2 Pathogenicity

By far the best known pathogenic mycobacteria are *M. tuberculosis* and *M. leprae*, the causative agent of leprosy. Both are slow growing mycobacteria. In addition to these well-known pathogens there are a number of other slow growing mycobacterial species which cause human infections (Brosch *et al.*, 2001). *M. bovis* infects a number of mammalian hosts including humans. *M. ulcerans* is the cause of an epidemic of cutaneous infections in West Africa and atypical mycobacterial infections such as the *Mycobacteria avium intracellulare* complex are increasingly found to be associated with HIV (Shafran, 1996). Of the fast growing mycobacterial species, the *M. fortuitum* group is most commonly associated with human disease.

1.1.3 *Mycobacterium tuberculosis*

Mycobacterium tuberculosis is the causative agent of tuberculosis (TB). In 2003 there were estimated to be nine million new TB infections, with the total number of infected people expected to rise to one billion by the year 2020 (WHO, 2006). TB causes more deaths worldwide than any other curable infectious disease (WHO Report, 2004). TB rarely affects young children and affects more men than women with incidence rates highest in young adults (Dye *et al.*, 2007).

In 2003, 31% of worldwide TB deaths occurred in Africa although Africa has only 11% of the world's population (See Figure 1.3).Although the rate of TB infections fell in all other regions between 1997 and 2006, the rate of infections continued to increase in Sub-Saharan Africa (Dye *et al.*, 2009). In Sub-Saharan Africa, 75% of TB deaths are HIV-related (WHO Report, 2005). In Africa, TB is the leading cause of death in HIV-infected individuals. Rates of recurring TB infection in patients with a previous history of HIV-related TB is frequent (Churchyard, 2003; Seyler *et al.*, 2005). Multidrug resistant TB is defined as a strain which is resistant to at least

rifampicin and isoniazid and is becoming increasingly common in Africa. Treatment of multidrug resistant TB is expensive, often toxic and is generally unavailable in Africa (Dean *et al.*, 2002). Multidrug resistant TB has a high fatality rate and low treatment success (Frieden, 2002).

1.2 Rifampicin

Rifampicin (known as rifampin in the United States) was first produced by Dow-Lepetit Laboratories in Milan, Italy in the mid 1960s (Sensi, 1983). Rifampicin is a semi-synthetic derivative of rifamycin B, which was isolated as a natural fermentation product of *Nocardia mediterranei*. *N. mediterranei* was initially placed in the genus *Streptomyces* but was later reclassified (Margalith and Pagani, 1961). Although five rifamycins were initially isolated from *N. mediterranei* designated rifamycin A through E, only rifamycin B was found to be stable (Sensi *et al.*, 1960).

Although rifamycin B was found to have no antibacterial activity it became 'activated' in an oxygenated aqueous solution (Sensi *et al.*, 1960). The mechanism of this 'activation' was found to be caused by the spontaneous conversion of rifamycin B into rifamycin O, which is then hydrolysed into rifamycin S by losing a single molecule of glycolic acid. Once rifamycin S is mildly reduced, rifamycin S yields rifamycin SV (Sensi *et al.*, 1960).

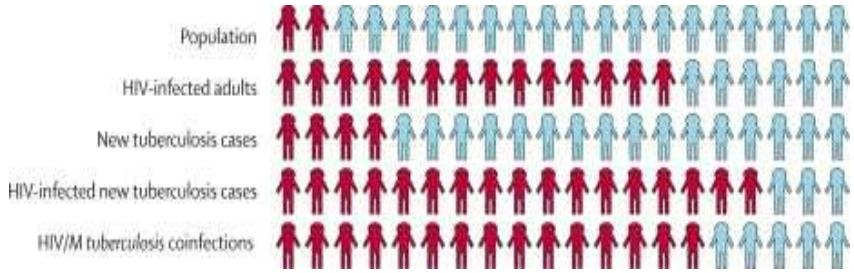


Figure 1.3 Disproportionate burden of HIV, HIV-related tuberculosis, and *M. tuberculosis* co-infections in Africa, for 2000. Every person represents 5% of the global total, with African people shown in red and the rest of the world in blue (Taken from Corbett *et al.*, 2003).

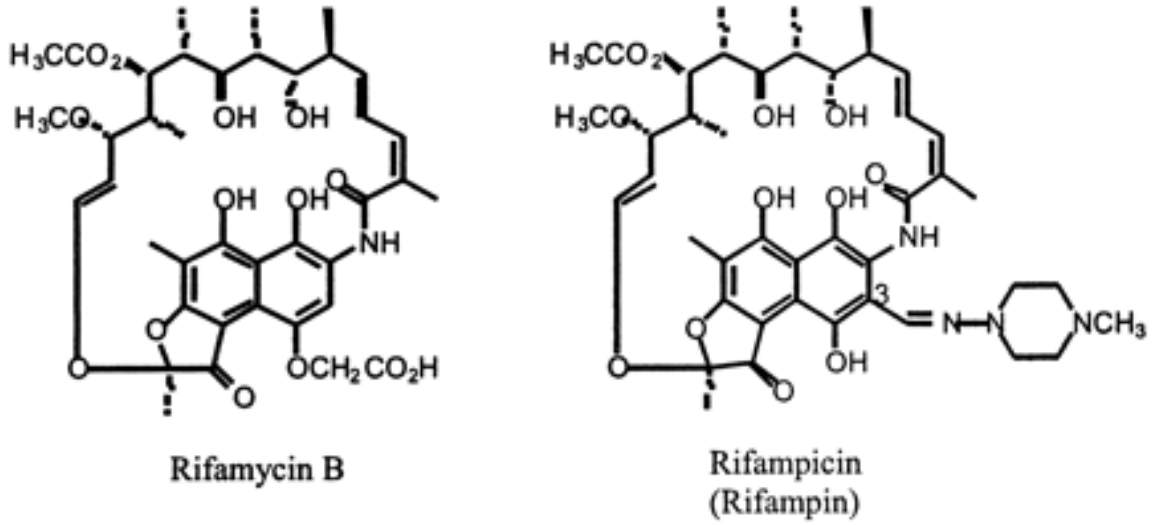


Figure 1.4 Chemical structure of rifampicin and its precursor rifamycin B (Watanabe *et al.*, 2003).

Rifamycin SV was found to be highly active against bacteria, especially Gram positives such as *Mycobacterium tuberculosis*. However it was not suitable for oral administration as it is not easily absorbed by the gastrointestinal tract and is quickly metabolised by liver enzymes leading to undetectable levels in the blood stream (Sensi, 1983). During a study which produced several hundred synthetic variations of Rifamycin SV, rifampicin was discovered.. Rifampicin is 3- formylrifamycin SV with *N*-amino-*N´*-methylpiperazine (see Figure 1.4). Rifampicin was found to be highly antimicrobial, easily absorbed and not toxic in the doses required for treatment. In 1968, rifampicin was introduced into TB therapy in Italy and was approved by the United States Food and Drug Administration (FDA) in 1971 (Sensi, 1983).

1.2.1 Mode of action

Rifampicin inhibits prokaryotic transcription through tight binding to RNA polymerase (see Figure 1.5). This was first shown by Hartman in 1967. There are five stages of transcription, namely; binding of polymerase to DNA; binding of the first nucleoside triphosphate to the DNA-polymerase complex; chain initiation through the formation of the first phosphodiester bond; chain elongation by the assembly of more nucleotides; chain termination by release of the complete polyribonucleotide. Rifampicin aborts chain elongation by preventing the formation of the second phosphodiester bond once chain initiation has taken place (Johnson and McClure, 1976).

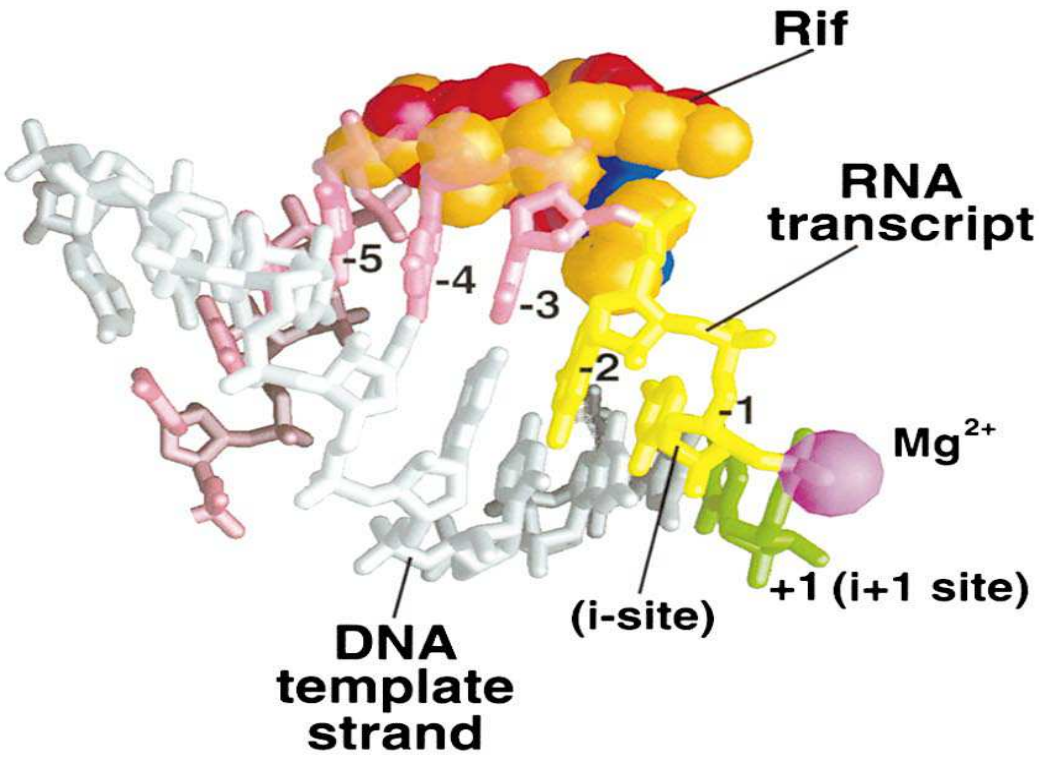


Figure 1.5 Mechanism of transcriptional inhibition by rifampicin. The RNA itself and the rest of the nucleic acids have not been depicted for reasons of clarity. Mg^{2+} is represented by the magenta sphere. The ribonucleotides are labeled with respect to the initiation site. The incoming nucleotide substrate at the -1 position is shown in green and the -1 and -2 positions that can be accommodated in the presence of rifampicin are depicted in yellow. The RNA further upstream (-3 to -8) that cannot be accommodated in the presence of rifampicin is coloured pink. The DNA template strand is grey. Rifampicin is shown positioned in its binding site on the β subunit. The carbon atoms are coloured in orange while the oxygen and nitrogen atoms are coloured in red and blue respectively (Nicholls *et al.*, 1991). The rifampicin is partially transparent, illustrating the RNA nucleotides (at -3 to -5) that sterically clash.

RNA polymerase consists of five subunits $\alpha_{(2)}\beta\beta'\sigma$. Rifampicin binds to the β subunit of prokaryotic RNA polymerase (Mindlin *et al.*, 1972). This binding is lipophilic with the lipophilic ansa chain of rifampicin extending furthest into the polymerase (Stender and Scheit, 1977). By binding in this manner, rifampicin prevents chain elongation of the di-ribonucleotide formed at the chain intiation stage.

1.2.2 Mechanisms of resistance

The most common mechanism of rifampicin resistance is the mutation of nucleotides within the *rpoB* gene which encodes the β-subunit of RNA polymerase (Jin and Gross, 1988; Telenti *et al.*, 1993; Williams *et al.*, 1994). The mutations prevent the binding of rifampicin to the RNA polymerase. *rpoB* mutations conferring rifampicin resistance are located within a 225 bp region of the gene (Jin and Gross, 1998). These mutations usually affect amino acids in regions 502-532; 560-572 and 687 (Aubry-Damon *et al.*, 1998).

Studies of *M. smegmatis*, *M. intracellular* and *M. avium* have indicated that rifampicin resistance in these organisms occurs by mechanisms other than *rpoB* mutations (Guerro *et al.*, 1994; Hetherington *et al.*, 1995). It has been established that there were four types of rifampicin resistance which inactivate the drug: glucosylation (Yazawa *et al.*, 1994); decomposition (Dabbs *et al.*, 1987); phosphorylation (Yazawa *et al.*, 1994); ribosylation (Dabbs *et al.*, 1995; Quan *et al.*, 1997).

Inactivation of rifampicin by glusosylation involves the substitution of the hydroxyl group of C_{23} in rifampicin with a β-D-glucose (Yazawa *et al.*, 1994). This form of inactivation is prominent in the *Nocardia*. Another mechanism used by *Nocardia* as well as *Rhodococcus*, is that of phosphorylation. During this process a phosphate group is added to the ansa bridge of rifampicin at C_{21} (Yazawa *et al.*, 1994).

Decomposition of rifampicin is widespread through a number of bacterial genera including *Nocardia*, *Mycobacteria* and *Rhodococcus* (Dabbs *et al.*, 1987; Andersen *et al.*, 1997). Ribosylation is a relatively newly discovered mechanism of antibiotic inactivation and appears to be limited to rifampicin inactivation (Dabbs *et al.*, 1995; Quan *et al.*, 1997, Baysarowich *et al.*, 2008).

1.2.3 Rifampicin in the treatment of infections

Rifampicin is far more effective against Gram positive organisms because there is ineffective penetration of Gram negative cell walls by rifampicin (Wehrli, 1980). Its efficacy against Gram positives resulted in its identification as a potential TB treatment early on in its development (Sensi, 1983). Today, rifampicin is recognized as the primary TB treatment along with isoniazid.

In addition to treatment of TB, rifampicin has been successfully used in the treatment of numerous infections including urinary tract infections, bone infections, endocarditis, foreign-body related infections, respiratory tract infections, bacterial meningitis, as well as skin and soft-tissue infections (Bliziotis *et al.*, 2007). Rifampicin is also being increasingly used in the treatment of drug resistant *Staphylococcus aureus* (Ellis *et al.*, 2009).

1.3 ADP-ribosyl transferases

ADP-ribosylation involves the transfer of an ADP-ribose group from NAD^+ to a substrate with the loss of nicotinamide. Collier was the first to elucidate the mechanism of an ADP-ribosylation reaction when studying the effect of diphtheria toxin on elongation factor 2 (Collier, 1967). Many well-studied ADP-ribosyl transferases are bacterial toxins. These include cholera toxin, diphtheria toxin, pertussis and *Pseudomonas aeruginosa* exotoxins A and S. All bacterial toxins which act through ADP-ribosylation act on one G-protein specific to that toxin (Shall, 1995). Some bacterial ADP-ribosyl transferase toxins are encoded by bacteriophages

which infect the bacterium (Uchida and Pappenheimer, 1971). Reversible ADP-ribosylation regulates nitrogenase in photosynthetic bacteria (Ludden, 1994). The human genome contains some ADP-ribosyl transferase genes of bacterial origin (Lander *et al.*, 2001).

In eukaryotes, ADP-ribosyl transferases are divided into mono ADP-ribosyl transferases (ARTs) and poly ADP-ribose polymerases (PARPs). The PARP superfamily consists of eighteen classes of proteins displaying a conserved catalytic domain (Amé *et al.*, 2004). The most extensively studied PARPs are nuclear-localised proteins which poly ADP-ribosylate proteins post-translationally. PARP-1 and PARP-2 are involved in the single-strand break repair and base excision repair pathways. PARP1 has also been implicated in the apoptotic pathway and has thus been identified as a possible cancer inhibitor. PARP-3 has been shown to interact with PARP-1 in the centrosome. For review articles (see Amé *et al.* (2004) and Bürkle (2005).

The ARTs are a relatively small gene family (Glowacki *et al.*, 2002). The majority of ARTs are ectoenzymes which act on extracellular substrates (Corda and Di Girolamo, 2003). Intracellular ARTs have been implicated in cell cycle regulatory pathways through binding to G-protein β-subunit and preventing modulation of effectors such as phosphoinositide 3-kinase (Lupi *et al.*, 2000). For a full list of mono ADP-ribosyltransferases see Table 1.

Table 1 Mono ADP-ribosyl transferases, their sources and substates (Corda and Di Girolamo, 2003).

Enzymes	Source	Substrate/amino acid	Effect of the reaction
Viruses			
ALT	Bacteriophage T4	RNA polymerase/Arg265	Enhances viral transcription
Mod A	Bacteriophage T4	RNA polymerase/Arg265	Enhances viral transcription
Mod B	Bacteriophage T4	RNA polymerase/Arg265, S1 ribosomal protein/Arg	Enhances viral transcription
Prokaryotes			
Toxins			
Diphtheria	*Corynebacterium diphtheriae*	EF-2/diphtamide715	Inhibits protein synthesis
Exotoxin A	*Pseudomonas aeruginosa*	EF-2/diphtamide715	Inhibits protein synthesis
Exotoxin S	*Pseudomonas aeruginosa*	Ras family/Arg41	Disrupts actin microfilaments
Cholera	*Vibrio cholerae*	$G\alpha_s$, $G\alpha_s$/Arg187	Inhibits GTPase activity
LT1, LT2	*Escherichia coli*	$G\alpha_s$, $G\alpha_s$/Arg187	Inhibits GTPase activity
Pertussis	*Bordetella pertussis*	$G\alpha_i$, $G\alpha_o$, $G\alpha_t$/Cys351	Uncouples receptor and G protein
C2, iota t	*Clostridium botulinum*	Actin	Prevents actin polymerization
C3	*Clostridium botulinum*	Rho, Rac/Asn41	Disrupts actin cytoskeleton
C3-like	*Clostridium limosum*	Rho, Rac/Asn41	Disrupts actin cytoskeleton
EDIN	*Staphylococcus aureus*	Rho/Asn41	Disrupts Golgi apparatus
VIP2	*Bacillus cereus*	Rho/Asn41	Disrupts actin cytoskeleton
SpvB	*Salmonella enterica*	Actin	Prevents actin polymerization
Mtx	*Bacillus sphaericus*	Unknown	Unknown
Intracellular			
DRAT	*Rhodospirillum rubrum*	Dinitrogenase reductase/Arg101	Inhibits dinitrogenase reductase
Eukaryotes			
Ectoenzymes			
ART1	Human, rat, mouse	Integrin, defensin/Arg	Inhibits substrate activity
ART2(A,B)	Rat, mouse	Unknown/Arg	Role in T cell proliferation
ART3	Human, rat, mouse	Unknown	Unknown
ART4	Human, rat, mouse	Unknown	Unknown
ART5	Human	Unknown/Arg	Unknown
ART6(A,B)	Chicken	p33/actin/Arg28-206	Inhibits substrate activity
ART7	Chicken	Unknown	Inhibits substrate activity
Intracellular			
Sir2p	Yeast	Histones/Sir2p/acetyl-lysine	Involved in histone deacetylation
Sirtuin2	Human	Albumin/acetyl-lysine	Involved in histone deacetylation
Pierisin1,2	Cabbage butterfly	DNA	Cytotoxic
Arginine-specific	Hamster, human	Gβ/Arg129	Inhibits substrate activity
Cysteine-specific	Human	GDH/Cys	Inhibits substrate activity

The above mentioned ADP-ribosyltransferases all act on proteins, share a similar global fold and use the same catalytic mechanism (Yates *et al.*, 2006; Hollburn *et al.*, 2006). The ADP-ribose group is added to a nitrogen atom on a susceptible amino acid, often arginine.

1.4 ADP-ribosylation in *Streptomyces coelicolor*

In the filamentous actinomycetes, a novel developmental cycle leads to the formation of an aerial mycelium at the beginning of differentiation (Kelemen and Buttner, 1998). In *S. coelicolor*, vegetative substrate hyphae are formed at spore germination. These hyphae rise to aerial hyphae which further differentiate into chains of exospores (Elliot *et al.*, 1998). The formation of the aerial hyphae coincides with the production of secondary metabolites including antibiotics. The production of aerial hyphae is known as morphological differentiation while the production of secondary metabolites is known as physiological differentiation.

Both physiological and morphological differentiation in *S. coelicolor* are under the control of the *bld* genes, known as such because colonies of *bld* mutants have a shiny "bald" appearance because they lack aerial hyphae (Kelemen and Butler, 1998). There are 10 *bld* genes: *bldJ* (formerly *bld216*), *bldA*, *bldB*, *bldC*, *bldD*, *bldF*, *bldG*, *bldH*, *bldI* and *bldK* (Merrick, 1976; Puglia and Cappelletti, 1984; Champness, 1988; Harasym *et al.*,1990; Wiley *et al.*,1993).

ADP-ribosylation is a key factor in the *bld* pathway (Shima *et al.*,1996; Sugawara *et al.*,2002). A regulatory locus designated *brgA* was found to play an essential role early on in the differentiation process by modulating the activity of proteins through ADP-ribosylation (Shima *et al.*,1996). In addition, *bldB*, *bldC* and *bldH* mutants failed to ADP-ribosylate a number of proteins which were ADP-ribosylated in parent wild type strains. *BrgA* leads to the ADP-ribosylation of at least four proteins in *S. coelicolor* (Sugawara *et al.*,2002). These proteins are BldKB, MalE and two

periplasmic proteins. It is not clear if *brgA* encodes an ART or encodes a factor which controls transcription of an ART encoding gene.

BldKB is a member of the *bldK* cluster of genes which encode the components of an oligopeptide importer-exporter which is a member of the ATP-binding cassette (ABC) membrane spanning transporter family (Nodwell *et al.*, 1996). BldK is the transporter which is responsible for importing a *bldJ* dependent extracellular signal which initiates the *bld* signaling cascade. BldKB is an extracellular subunit of BldK. MalE is involved in maltose utilization (Driessen *et al.*, 2000) and the two periplasmic proteins are part of high affinity active transport (van Wezel *et al.*, 1977).

These target proteins are all on the membrane and are involved in nutrient transport. They appear to be ADP-ribosylated on the thiol group of cysteine as shown by the susceptibility of the linkage to mercuric chloride treatment. Sugawara *et al* identified specific cysteine residues within these proteins as putative locations of ADP-ribosylation (Sugawara *et al.*, 2002). These cysteine residues are all situated in the N-terminal region of the proteins and there is a common amino acid sequence flanking them: Leu (or Ala) – Thr (or Ala) – Ala –Cys – Gly. This sequence is a known lipoprotein signal sequence and is the site at which lipoprotein modification occurs which is essential for the export of lipoproteins as well as their anchoring to the membrane (Wu *et al.*, 1996).

If these cysteine residues are the site of ADP- ribosylation, then these proteins would no longer be recognized by enzymes which form *N*-acyl-diacylglycerylcysteine and would not be exported to the membrane and would not be able to act in transport of nutrients (Sugawara *et al.*, 2002).

1.5 ARR

The only known naturally occurring ADP-ribosylation of a small non-proteinaceous target occurs in *M. smegmatis* in which rifampicin is ADP-ribosylated by the enzyme

ARR (Dabbs *et al.*, 1995; Quan *et al.*, 1997; Quan *et al.*, 1999). This ADP-ribosylation leads to the inactivation of the antibiotic. The ARR (name derived from ADP-ribosylates rifampicin) gene was identified in a genomic library of DNA of *M. smegmatis* DSM 43756 (Quan *et al.*, 1999).

The gene is 429 bp in length and begins with a GTG start codon (Quan *et al.*, 1997). Although AUG is accepted as a universal start codon, Mycobacteria frequently use GTG as a start codon. In keeping with the predicted size of the translated product of this gene the ARR protein is small ≈16kDa ,143 amino acids (see Sequence 1). The gene showed no sequence similarity to known ADP-ribosyl transferases. A putative upstream regulatory element has high sequence similarity to an upstream region of another rifampicin inactivation gene from *Rhodococcus equi iri*, suggesting a common regulatory mechanism between the two genes (Quan *et al.*, 1997).

VVARNPPKPF EVHESGAYLH GTKADLKVGD RLVPGRESNFE
AGRIMKHVYI TQTLDAAVWG AELAVGEGRG RIYIVEPEGH
IHDDPNVTDK KLPGNPTRSY RTREPVRIVG ELTDWEGHSP
EQIAAMREGL EDLRRKGLAV IYD

Sequence 1: ARR amino acid sequence.

1.5.1 Structure

In 2008, the crystal structure of ARR was determined by Baysarowich *et al*. ARR is made up of a single domain containing seven β-strands (arranged in two antiparallel β-sheets) and two α-helices (Baysarowich *et al.*, 2008). The two β-sheets (β1-4-7-2 and β3-6-5) are arranged almost perpendicularly to one another. The two α-helices form the deep cleft into which rifampicin binds. Once rifampicin has bound, an extended loop region containing β5 and β6 closes over the rifampicin forming a lid (see Figure 1.7).

It is significant that although rifampicin is quite deeply buried within ARR there are no ionic or covalent bonds formed between rifampicin and any of the amino acid side chains of ARR (Baysarowich *et al.*, 2008). Apart from two hydrogen bonds, formed with N86 and K91, the interaction is entirely hydrophobic. This allows a high degree of plasticity in this interaction suggesting that ARR may be able to act on hydrophobic substrates other than rifampicin.

In 2008 Gianniosis *et al* found that the single amino acid substitution P5T caused a significant change in charge distribution (Gianniosis *et al.*, 2008). Their work was carried out before publication of the ARR crystal structure later the same year; however, the elucidation of the structure shed no light on possible reasons for this change as P5 is not represented in the structure, as there is no electron density beyond residue six (see figure 1.6) This mutation did not affect the minimum inhibitory concentration of rifampicin in mutant cells in comparison to cells expressing the wild type protein, indicating that it does not affect activity. There was also no change to the predicted hydrodynamic volume between mutant and wild type protein, again suggesting that the mutation does not lead to global conformational changes, although it is notable that the estimated hydrodynamic volume obtained by dielectric spectroscopy is twice that of the value estimated by the crystal structure.

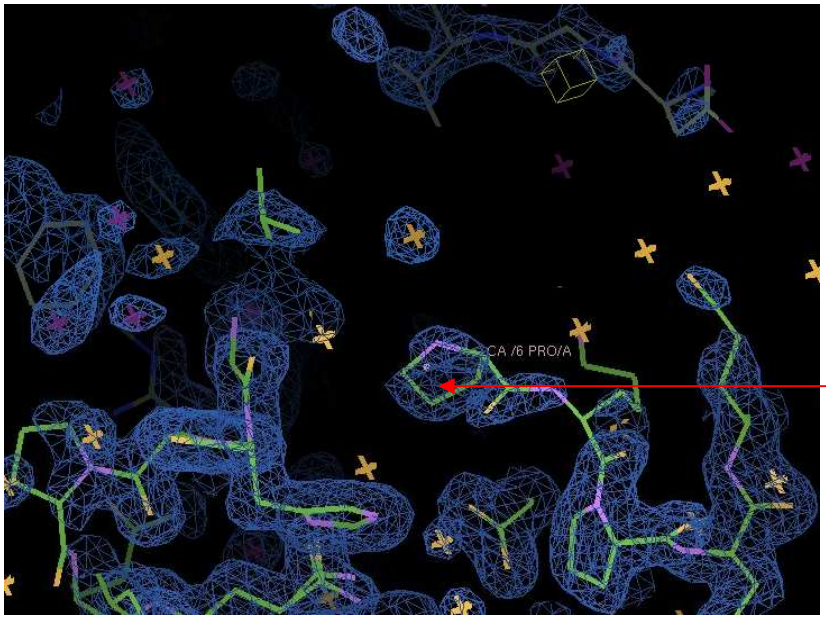


Figure 1.6: Electron density map of ARR viewed with ARR co-ordinate file viewed in Coot (Emsley and Cowtan, 2004). Electron density is shown in blue. Proline six is marked. There is no clear electron density beyond this point. PDB co-ordinate 2HW2.

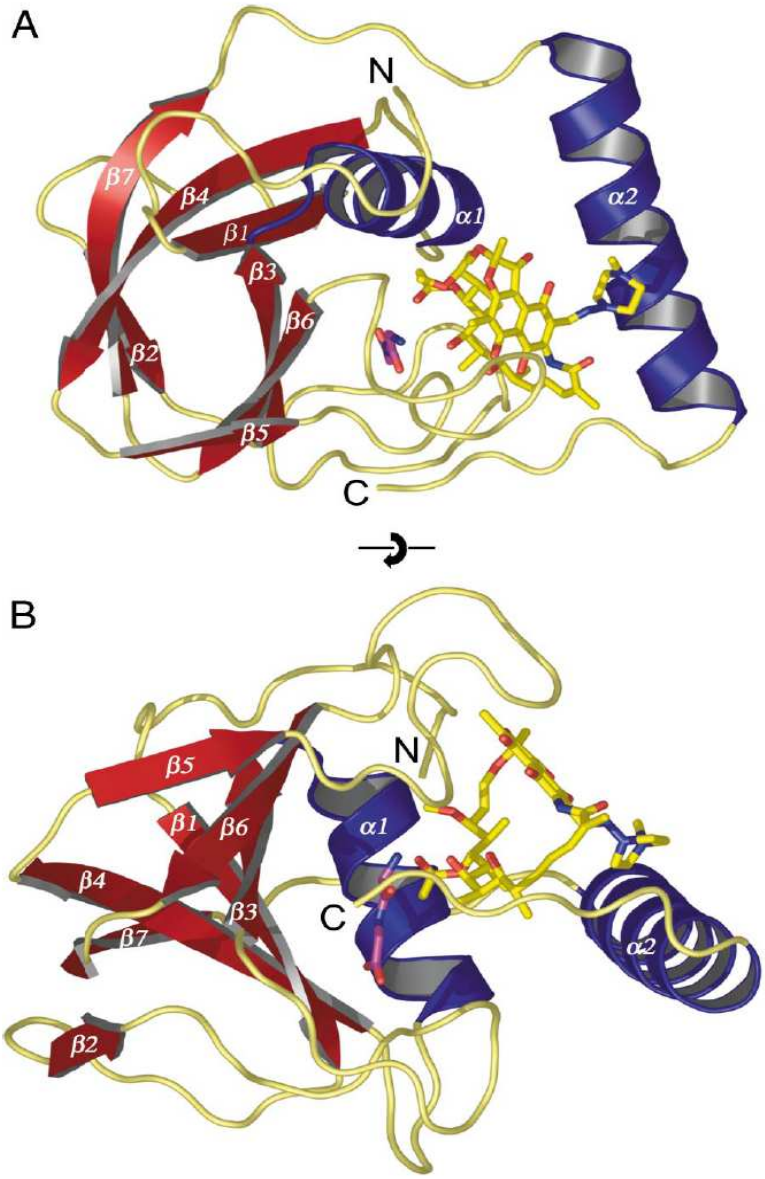


Figure 1.7 Ribbon diagram representing the crystal structure of ARR bound to rifampicin (Taken from Baysarowich *et al.*, 2008). Rifampicin is shown in bright yellow, β-strands in red and α-helices in blue. Helices and strands are numbered. The carboxy and amino termini are marked by C and N respectively. B is a 180º rotation of A about the horizontal axis.

1.5.2 Mechanism

ARR adds an ADP-ribose group onto the oxygen atom at C_{23} of rifampicin. The preferred source of ADP-ribose is NAD^+ (Baysarowich *et al.* 2008). ADP-ribosylated rifampicin has also been termed RIP-TA (Quan *et al.*, 1999) The structure of RIP-TA (MW 1.363 kDa) has been determined by ^1H and ^{13}C nuclear magnetic resonance (NMR) analyses; ^1H-^1H correlated spectroscopy, pulsed-field gradient heteronuclear multiquantum coherence, and pulsed-field gradient heteronuclear multiple-bond correlation and its chemical formula has been determined as $C_{58}H_{79}N_9O_{25}P_2$ (see Figure 1.8).

In the predicted molecular mechanism of ARR-mediated ADP-ribosylation of rifampicin (using ARR's structure and known mechanisms of other ADP-ribosyltransferases), Asp84 stabilises the development of a oxocarbenium ion allowing a transition-state which allows attack of C_{23} of rifampicin by C_1 of the ribose (see Figure 1.9). Dabbs and Quan suggested that ARR may use FAD^+ as a source of ADP-ribose, because the reaction is inhibited by light (Dabbs and Quan, 2000)

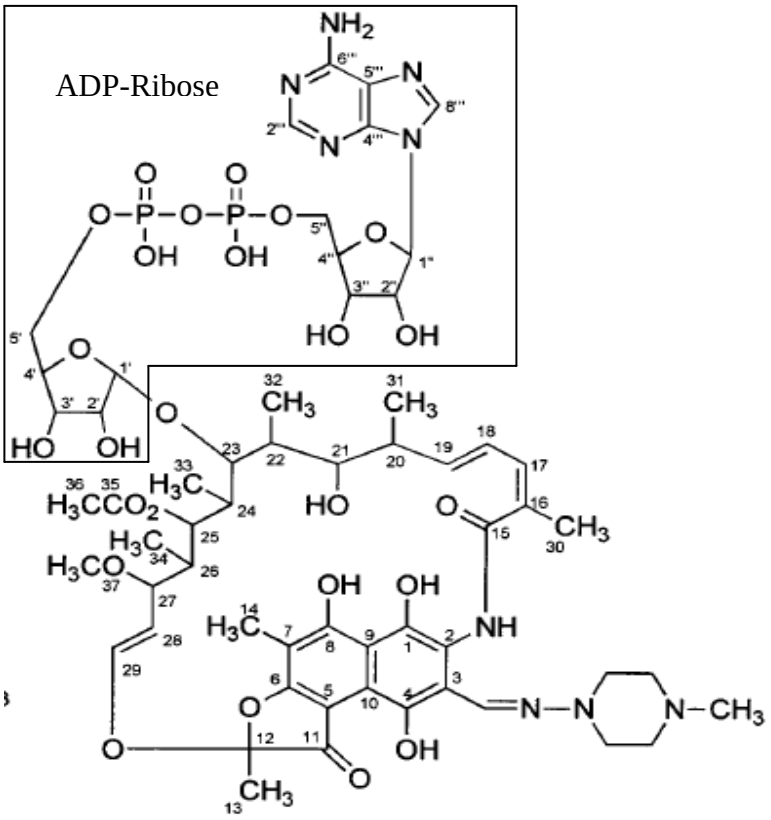


Figure 1.8 ADP-ribosylated rifampicin (taken from Quan *et al.*, 1999).

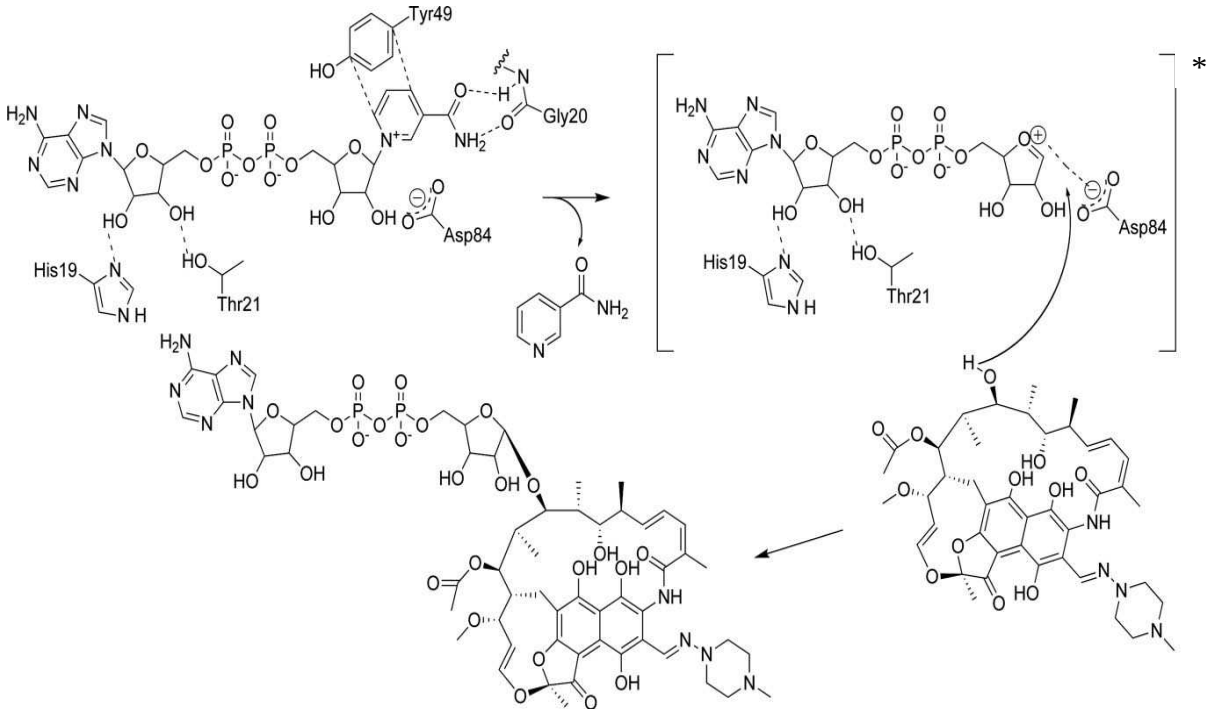


Figure 1.9 Schematic representation of ARR catalysed ADP-ribosylation of rifampicin (taken from Baysarowich *et al.*, 2008). Transition state indicated by *.

1.5.3 Enzyme activity

ARR can ADP-ribosylate a wide variety of rifamycins with diverse naphthyl substituents (Baysarowich *et al.*, 2008). These rifamycins are all of clinical importance including rifampicin (TB treatment), rifabutin (*Mycobacterium avium* complex treatment) and rifaximin (prophylactic treatment of traveller's diarrhoea).

Steady-state kinetic parameters have been determined for ARR. K_m values for binding of rifampicin and NAD^+ are in the millimolar range (Baysarowich *et al.*, 2008). k_{cat}/K_m values which indicate an enzyme's ability to catalyse a reaction at a low substrate concentration are 2.00×10^3 $s^{-1} \cdot M^{-1}$ for rifampicin and 1.54×10^4 $s^{-1} \cdot M^{-1}$ for NAD^+.

1.5.4 Homology

1.5.4.1 Sequence

ARR derived from *M. smegmatis* shows no sequence identity with other known mono or poly ADP-ribosyltransferases. However, another form of the ARR gene designated *ARR-2* shows a 55% sequence identity and is present as a mobile genetic element in numerous bacteria (Baysarowich *et al.*, 2008). *ARR-2* integrons or transposons are most notably found in several Gram-positive pathogenic bacteria including *Pseudomonas areuginosa* (Tribuddharat and Fennewald, 1999), *Escherichia coli* (Naas *et al.*,2001), *Klebsiella pneumoniae* (Arlet *et al.*, 2001) and *Acinetobacter baumannii* (Houang *et al.*, 2003). There is clear evidence for lateral gene transfer of *ARR-2*.

1.5.4.2 Structural

Although there is no significant DNA or protein sequence similarity between ARR and other ADP-ribosyltransferases there is definite structural homology. Based on their crystal structures, ADP-ribosyltransferases are divided into two classes. Class I, which includes the PARPs and diphtheria toxin, and Class II, which includes ART2

and VIP2 toxin (Baysarowich *et al.*, 2008). The two classes share a common NAD^+-binding core made up of two adjoining β-sheets made up of five β-strands. Distinguishing features between the classes included a variable number of α-helices and distinct motifs, H-Y-E in Class I and R-S-E in Class II.

Structural alignments show that ARR shows structural similarities with Class I ADP-ribosyltransferases (Baysarowich *et al.*, 2008). ARR appears to represent the minimum unit required for ADP-ribosyltransferase activity. Although sequence alignment software fails to reveal the common H-Y-E motif found in Class I ADP-ribosyltransferases, a structure-based comparison revealed two residues (H19 and Y49) likely to serve the same NAD^+-binding function in ARR. H19 is present in β1 and Y49 in β3. Mutational studies have shown that these two residues are essential for enzyme function (Baysarowich *et al.*, 2008).

There is a single glutamic acid on the front edge of β6 which is conserved in all previously known ADP-ribosyltransferases. This residue is believed to be absolutely necessary for ADP-ribosyltransferase activity; however, this residue is not present in ARR (Baysarowich *et al.*, 2008). This residue is necessary for stabilising the predicted oxocarbenium ion in the transition-state of ADP-ribosylation. In ARR it appears that the aspartic acid residue in position 84 which is in the loop connecting β5 to β6 performs this function.

1.5.5 Heat stability

Previous studies have shown that ARR functions optimally at 45 °C (Gianniosis, 2007) and that the enzyme is able to function after heating to 100 °C (Quan *et al.*, 1997). This indicates that ARR is a thermostable protein. Many species of *Mycobacterium* are able to withstand high temperatures as a consequence of their hardy cell walls. Thermophilic bacteria have been shown to have extensive ADP-ribosylation systems (Faraone-Mennella *et al.*, 2006). It is not clear whether there is a

link between ADP-ribosylation systems and the ability of a bacterium to persist at high temperatures.

1.5.6 Changes in cell wall morphology

Previous studies have shown that the expression of ARR changes cell wall morphology in both Gram-negative and Gram-positive bacteria. This has been shown by the formation of long filaments in *E. coli* expressing ARR (Puhaca, 2005; Gianniosis, 2007) and a change in colony morphology in *R. equi* (Quan, 1997).

The reason for this change in cell wall morphology is not clear. It appears to be independent of ADP-ribosylation of rifampicin and uses ARR residues, distinct to those involved in rifampicin binding, in binding this target (Puhaca, 2005). This second function of ARR would explain the extreme plasticity of binding to rifampicin as well as the large number of ARR-like sequences from environmental isolates where there has been no obvious human pressure (Baysarowich *et al.*, 2008).

1.5.7 A possible threat to TB treatment

Baysarowich found that the majority of *arr*-like sequences were present either as integrons or alongside transposable sequence. This suggests a high level of lateral gene transfer of this gene. It has been shown that mycobacteria can acquire pathogenicity factors through lateral gene transfer (Paustian *et al*., 2005). With the high number of HIV patients infected with TB and other opportunistic pathogens, it is becoming increasingly likely that *M. tuberculosis* will come into contact with another pathogen carrying *arr* as a mobile genetic element and gene transfer would be possible.

Should *arr* be transferred to *M. tuberculosis* the resultant strain would pose a huge threat to the treatment of TB with rifampicin as a frontline drug. Recently, an

orthologue of rifampicin, in which the acetate group on carbon 25 is substituted with a carbamate group, was found to be effective at disabling bacterial transcription but could not be used as a substrate for ARR (Combrink *et al.*, 2007) . Should strains of *M. tuberculosis* containing the *arr* gene arise, this compound will be invaluable in treating the resulting rifampicin resistant TB.

1.6 Aims of this project

This project aims to optimize over-expression and purification of ARR, as well as characterize ARR biochemically by use of fluorescence and circular dichroism spectroscopy, as very little biochemical information is available on ARR. It also aims to further investigate changes in cell wall morphology by microscopy and, if possible, explain these changes.

2 Materials and Methods

2.1 Bacterial strains

Bacterial strains used in this work are listed in Table 2.1.

2.2 Plasmids

pQE-30 was used for protein expression. pGEM-T Easy was used for sequencing. (see Table 2.2).

2.3 Plasmid purification

Plasmid purification was carried out using the Genejet Miniprep Kit (Fermentas). This kit is based on the alkaline lysis method of Birnboin and Doly (Birnboin and Doly, 1979). Alkaline lysis (by sodium dodecyl sulphate and sodium hydroxide) is used to lyse cells and denature cell components. When the lysate is neutralized plasmid DNA is able to renature because of its small size and also because it has been somewhat protected from denaturation due to its circular nature. Once plasmid DNA has renatured, it resolubilizes while all other denatured cell components (chromosomal DNA, proteins, lipids) remain insoluble and are pelleted out. The plasmid DNA is then purified further by binding the DNA to a silica membrane in the presence of high concentrations of chaotropic salt (in this case guanidinium chloride). After an ethanol rinse, the DNA is eluted from the membrane under low salt conditions.

Table 2.1 Bacterial strains used in this work.

Species	**Strain**	**Selection Markers**	**Source**
E. coli			
	MM2942	-	E. Dabbs
	BL21 (DE3) pLysS	-	Promega
	SG13009 (pREP)	Kanamycin resistance	Qiagen
	E. cloni	-	Lucigen
M. smegmatis			
	DSM 43657	-	E. Dabbs
	DSM 43657Km1	Kanamycin resistance	S. Quan

Table 2.2 Plasmids used in this work

Plasmid	**Insert**	**Size**	**Source**
pQE-30	-	~ 3.5 Kb	Qiagen
pQE-30	ARR	~ 4 Kb	J. Burke
pGEM-T Easy	-	~ 3 Kb	Promega
pGEM-T Easy	pQE-30 ARR fragment	~ 3.6 Kb	This work

2.3.1 Determination of DNA concentrations

DNA concentrations were determined spectrophotometrically using the NanoDrop ND-1000 spectrophotometer. The absorbance at 260 nm was used given that 50 μg/ml of double stranded DNA has an A_{260} of 1.0. This value is based on the Beer Lambert law and the extinction coefficient of double stranded DNA.

2.3.2 Agarose gel electrophoresis

Gels of 1% agarose were prepared using 1 X TAE buffer (40 mM Tris acetate; 1 mM EDTA). Two microlitres of ethidium bromide (1 mg/ml) was added to 50 ml of gel solution. Samples were loaded with 6 X Sample Loading Buffer Fermentas (10 mM Tris, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% glycerol, 60 mM EDTA). Gels were electrophoresed in 1 X TAE buffer until the bromophenol blue band of the tracking dye in the sample buffer had electrophoresed to approximately 3 cm of the bottom edge of the gel. Gels were viewed under a UV-transilluminator at 320 nm.

2.4 Polymerase chain reaction (PCR)

PCR is a commonly used molecular biological technique used to produce large amounts of DNA from a small amount of template through amplification using DNA-dependent DNA polymerase (Mullis *et al.*, 1986).

2.4.1 Oligonucleotide primer design

DNA-dependent DNA polymerase requires a short DNA primer to function. This primer is complimentary to the template and upon annealing to the template is extended in a template dependent manner by the polymerase enzyme. Primers were designed to bind about 75 bp away from the T-overhangs in pGEM-T Easy. Primers were checked for propensity to form secondary structures and primer dimers using

the GeneRunner version 3.01 software (copyright © 1996 Hastings Software Inc). These primers were named pQE-30seq forward and pQE-30seq reverse.

2.4.2 PCR of the ARR fragment from pQE-30

The best annealing temperature was determined by thermocycling eight reactions (pQE-30seq forward and reverse 2 µM each; pQE-30seq reverse; template 100 ng/µl; 1X Fermentas Mastermix) through the same cycle but with varying annealing temperatures in a BioRAD MJ mini thermocycler. The annealing temperatures used were: 60 ºC; 59.3 ºC; 58 ºC; 56.2 ºC; 54 ºC; 52.1 ºC; 50.8 ºC and 50 ºC. These temperatures were chosen by taking the average T_M of the forward and reverse primers and using a gradient that ranged from ±5 ºC of this average. These eight reactions were thermocycled using plasmid extracted from *E. coli* BL21 (DE3) pLysS and MM2942 cells to determine which was a more efficient template.

In order to obtain the highest yield of product with the lowest amount of non-specific amplification, various parameters were tested (Table 2.3). These included $MgCl_2$ concentration, template DNA concentration and extension time during cycling.

Thermocyling conditions were as follows: initial denaturation at 94 °C for 1 minute, 30 cycles of denaturation at 94 °C for 30 seconds, primer annealing at 50-60 °C for 30 seconds, and extension at 72 °C for 30 seconds to 60 seconds, and final extension at 72 °C for 5 minutes. PCR products were further purified using the Promega Wizard SV Gel and PCR Clean Up System and subsequently evaluated by agarose gel elctrophoresis on a 1% gel.

Table 2.3 Components of PCR reactions for amplification of the ARR fragment from pQE-30.

Component	***Reaction 1***	***Reaction 2***	***Reaction 3***	***Reaction 4***
dNTP Mix	0.2 μM	0.2 μM	0.2 μM	0.2 μM
$MgCl_2$	1 mM	2 mM	1 mM	2 mM
Forward Primer	1 μM	1 μM	1 μM	1 μM
Reverse Primer	1 μM	1 μM	1 μM	1 μM
Template DNA	50 ng	50 ng	150 ng	150 ng
Taq reaction buffer	To 1X	To 1X	To 1X	To 1X
Taq polymerase	5 U	5 U	5 U	5 U
Nuclease free water	To 50 μL	To 50 μL	To 50 μL	To 50 μL

2.5 Cloning of ARR fragment from pQE-30 into pGEM-T Easy

Due to difficulties in sequencing the ARR insert in pQE-30 it was necessary to sequence the insert in an alternative vector. pGEM-T Easy (Promega) is a TA cloning vector. In addition to a multiple cloning sites, the vector has 5′ T-overhangs. Undigested PCR products can be ligated to these overhangs because polymerases, which do not have 3′-5′ editing activity, add 3′ terminal adenosine bases during PCR (Clark, 1988). *Taq* is one such polymerase.

2.5.1 Ligation

Ligation reactions were set up as follows: 1 µl Promega T4 Ligase (3 Weiss units/µl); 5 µl Ligation buffer (2 X); 1 µl pGEM-T Easy (50 ng/µl); 2 µl ARR insert (50 ng/µl); 1 µl nuclease free water. The reactions were left at 20 °C for 18 hours.

2.5.2 Transformation of competent cells with plasmid DNA

Transformations were carried out using the heat-shock method. Competent cells (as listed in Table 2.1) were thawed on ice. Approximately 50 ng of chilled DNA was added to 50 µl of cells. The cells were then incubated on ice for 30 min and then heat-shocked at 42 °C for 45 sec and placed back on ice for a further 5 min. After this, 500 µl of SOC (Super Optimal broth with Catabolite) media Novagen (2 % typtone, 0.5 % yeast, 10 mM NaCl, 2.5mM KCl, 10 mM $MgCl_2$, 40 mM glucose), was added and the cells were incubated with agitation at 37 °C for one hour to allow selection genes to be expressed.

Although, until recently, the exact mechanism by which heat-shock improved transformation efficiencies was unknown it has now been shown that the heat pulse from 0°C to 42°C back to 0°C forms pores in the membrane by causing the release of lipids from the cell surface and reduces the cell wall potential allowing the DNA to pass through (Panja *et al.*, 2008).

2.5.3 Confirmation of positive clones

pGEM-T Easy containing the ARR fragment was transformed into E. cloni cells (Lucigen) and the resultant transformants were screened for plasmid presence by overnight growth on Luria-Bertani broth (LB) Agar (1% tryptone; 0.5% yeast extract; 0.5% NaCl; 1.5% agar) supplemented with 100 μg/ml ampicillin. These colonies were then screened for the insert by colony PCR. Colony PCR was performed using single colonies, 1 μM each of T7 forward primer and M13 reverse primer and Fermentas PCR Master Mix.

2.6 Sequencing of ARR fragment

Plasmid DNA was extracted from E. cloni cells transformed with pGEM-T Easy containing the ARR fragment. This DNA was sent to Inqaba Biotec (Pretoria, South Africa) for sequencing using the universal T7 forward and M13 reverse primers.

2.7 Induction studies

Heterologous expression of genes in pQE-30 is under the control of the phage T5 promoter coupled to two lac operators. This system is inducible by isopropyl β-D-1-thiogalactopyranoside (IPTG). Although the T5 promoter is recognised by *E. coli* RNA polymerase and can thus be used in any *E. coli* strain, Qiagen advise that best results are obtained using Qiagen cell lines specifically designed for use with the pQE vector range (QIA-Express Cloning Manual). One of these cell lines is *E. coli* SG13009 (pREP).

2.7.1 Preparation of competent *E. coli* SG13009 (pREP) cells

This methodology is based on that of Hanahan (1983) and uses low temperatures and divalent ions to induce competency in cells. An overnight culture of *E. coli* SG13009 (pREP) cells was grown in LB supplemented with 25 μg/ml kanamycin (to maintain the pREP plasmid) at 37 ºC with shaking at 250 rpm. 3ml of this culture was added to 50 ml fresh LB and grown until the optical density at 600 nm was 0.6. The cells were then decanted into pre-cooled centrifuge tubes and left on ice for 15 minutes

The chilled cells were pelleted by centrifugation at 2500 rpm in a chilled Sorvall SS-34 rotor for 12 minutes. The supernatant was discarded and the cells were resuspended in 15 ml transformation buffer (10 mM Mes pH 6.2, 100 mM KCl, 45 mM $MnCl_2$, 10 mM $CaCl_2$, 3 mM $HACoCl_3$) and chilled on ice for 15 minutes then centrifuged for 10 minutes at 2500 rpm in a chilled Sorvall SS-34 rotor. The supernatant was discarded and the cells were resuspended 4 ml of transformation buffer. Dimethylformamide (DMF) was added to a concentration of 3.5 %. The cells were placed on ice for 5 minutes and then 140 μl of 1 M DTT was added. After chilling on ice for a further 10 minutes, 140 μl of DMF was added and the cells were chilled for 5 minutes before aliquoting.

2.7.2 Induction conditions

Cells (*E. coli* MM2942; BL21 (DE3) pLysS and SG13009 (pREP)) were grown in LB (1% tryptone; 0.5% NaCl; 0.5% yeast extract) at 37 ºC overnight shaking at 250 rpm. A 100 X dilution of the overnight culture was prepared in LB and grown until the optical density at 600 nm reached 0.4 (early log phase), 0.6 (mid log phase) and 0.8 (late log phase). Cells were then induced with either 0.5 or 1 mM IPTG and samples taken every 2 hours (see Table 2.4). These samples were then prepared for SDS-PAGE and electrophoresed as described below.

Table 2.4 Conditions examined in induction study.

Growth stage	**Cell Lines**	**Induction times (Hours)**	**IPTG concentrations**
Early- log phase	*E. coli* BL21 (DE3) pLysS; *E. coli* MM294; *E. coli* SG13009 (pREP)	0; 2; 4; 6; 8; >10	0.5mM; 1mM
Mid-log phase	*E. coli* BL21 (DE3) pLysS; *E. coli* MM294; *E. coli* SG13009 (pREP)	0; 2; 4; 6; 8; >10	0.5mM; 1mM
Late-log phase	*E. coli* BL21 (DE3) pLysS; *E. coli* MM294; *E. coli* SG13009 (pREP)	0; 2; 4; 6; 8; >10	0.5mM; 1mM

2.7.3 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed by the discontinuous method described by Laemmli (1970). A 15 % acrylamide separating gel was prepared with 0.375 M Tris/HCl buffer (pH 8.8) and 0.1% SDS. A 4% acrylamide stacking gel was prepared with 0.0625 M Tris/HCl buffer (pH 6.8) and 0.1% SDS. Samples were prepared for electrophoresis by diluting them 1:1 with SDS sample buffer (2% SDS, 5% β-mercaptoethanol, 10% glycerol and 0.02% bromophenol blue) and heating at 95 ºC for 5 minutes. Molecular weight markers were also prepared in the same manner.

After application of samples to wells, gels were electrophoresed at 120 V until the bromophenol blue had migrated to approximately 1 cm of the bottom of the gel. The electrophoresis buffer was made up of 0.025 M Tris and 0.192 glycine buffer (pH 8.6) containing 1% SDS. After electrophoresis gels were stained (2% Coomassie blue R250, 13.3% glacial acetic acid and 18.75% ethanol) and destained (40% ethanol and 10 % glacial acetic acid).

2.8 Purification of ARR

2.8.1 Overexpression of ARR

E. coli BL21 (DE3) pLysS cells containing the pQE-30 plasmid with ARR insert were grown in LB, supplemented with 100 μg/ml ampicillin, overnight at 37 ºC with shaking at 250 rpm. Four litres of a 100 X dilution of this culture were made up in fresh ampicillin-supplemented LB and grown under the same conditions until the optical density at 600 nm reached 0.6. At this stage, IPTG was added to a final concentration of 1 mM. The cells were allowed to grow for 18 hours.

The cells which had been induced overnight were pelleted by a 20 min centrifugation at 5 000 rpm in a Sorvall SLA-3000 rotor and resuspended in 1/100 the volume

phosphate buffer (50 mM Na_2HPO_4 pH7.5). Resuspensed cells were disrupted by five cycles of 30 second pulse sonication and centrifuged at 10 000 rpm for 20 min in a chilled Sorvall SS-34 rotor in order to separate the soluble component of the cell from the insoluble component. The insoluble pellet was discarded and the soluble supernatant was used in purification.

2.8.2 Immobilised metal affinity chromatography (IMAC)

The pQE-30 vector expresses a 6 X His tag upstream of the multiple cloning site resulting in N-terminally His-tagged proteins. IMAC is the most commonly used purification method of His-tagged proteins and Ni^{2+} columns are frequently used. However, there can be problems of low protein purity and metal leaching from the column. Columns charged with cobalt claim to minimize these problems. His-Bind resin (Novagen) is a Ni-NTA (nickel-nitrilotriacetic acid) resin. His-pur Cobalt resin (Pierce) is a tetradentate, cobalt charged, 6% cross-linked agarose bead resin.

2.8.2.1 His-Bind Purification

A 10 ml column of His-Bind resin was charged according to manufacturer's instructions and equilibrated with ten column volumes of binding buffer (0.5 M NaCl; 20 mM Tris HCl pH 7.9; 5 mM imidazole). After loading the soluble cell fraction 5 column volumes of binding buffer were run through the column. The column was then washed with ten column volumes wash buffer (0.5 mM NaCl; 20 mM Tris HCl pH 7.9; 60 mM imidazole) The protein was eluted using five column volumes elution buffer (0.5 M NaCl; 20 mM Tris-HCl pH 7.9; 1 M imidazole).

2.8.2.2 His-pur Purification

A 5 ml His-pur cobalt resin column was set up according to the manufacturer's instructions. The column was then equilibrated with 10 column volumes equilibration buffer (50 mM Na_2HPO_4, 300 mM NaCl, 10 mM imidazole, pH7.4). After

equilibration, the ARR containing soluble cell fraction was loaded onto the column and allowed to flow by gravity. Once the soluble fraction had passed through the column, another 10 column volumes of equilibration buffer were passed through the column to remove any proteins non-specifically bound to the column and finally ARR was eluted from the column using four column volumes of elution buffer (50 mM Na_2HPO_4, 300 mM NaCl, 150 mM imidazole, pH 7.4).

2.8.3 Determination of ARR purity and concentration

The purity of ARR was determined by SDS-PAGE. Protein concentration was determined spectrophotometrically using the Beer-Lambert law (Equation 1)

$$A = \varepsilon_\lambda c l \qquad (1)$$

where A is absorbance, ε_λ is the molar extinction coefficient at wavelength λ, c is concentration and l is path length. In order to determine the molar extinction coefficient of ARR at 280 nm the method of Perkins (1986) was used (Equation 2). This method uses the known molar coefficients of the amino acids tryptophan (5550 $M^{-1}cm^{-1}$), tyrosine (1340 $M^{-1}cm^{-1}$) and cysteine (150 $M^{-1}cm^{-1}$) at 280 nm.

$$\varepsilon(280)(M^{-1}cm^{-1}) = 5550\ \Sigma Trp + 1340\ \Sigma Tyr + 150\ \Sigma Cys \qquad (2)$$

By this calculation, the molar extinction coefficient of ARR at 280 nm is calculated as 18450 ($M^{-1}cm^{-1}$).

2.9 Spectroscopic studies

2.9.1 Far-UV circular dichroism (CD)

CD spectroscopy reports the difference in absorbance of left- and right-handed circularly polarized light by optically active modules. In proteins, there are two

primary optically active elements; the peptide backbone (active between 250 and 170 nm) and the aromatic side chains of amino acids (active between 300 and 250 nm). Far-UV CD (that of the backbone) offers valuable information about a protein's secondary structure and gives an indication of the α-helical, β-strand and random coil content (For a review of the use of circular dichroism of proteins see Yang *et al.*, 1986 and Woody, 1995).

Far-UV CD spectra were generated between 180 and 250 nm with a cuvette of 2 mm. The spectra were averages of ten accumulations. Thermal unfolding curves were generated by tracking the level of ellipticity at 222 nm while ramping the temperature of the sample at 1 ºC/min from 5 ºC to 85 ºC. ARR was dialyzed into 50 mM phosphate buffer (pH 7.5) to remove all chloride ions before all CD measurements were recorded.

2.9.2 Fluorescence spectroscopy

Fluorescent molecules are able to absorb light of a certain wavelength (excitation wavelength) and emit some of this light at a different wavelength (emission wavelength). This phenomenon, known as fluorescence, is due to electrons entering a higher energy state as a result of energy from the absorbed light. These electrons fall back to their lower energy state, causing energy to be released in the form of light. Because the light emitted is always of a lower energy (higher wavelength) than the light absorbed, the excitation wavelength of a molecule will always be shorter than the emission wavelength (Lakowicz, 1983).

Tryptophan is a fluorophore which is very sensitive to its local environment and can thus be used as a local probe in a protein. From the crystal structure of ARR it can be deduced that the tryptophan at position 57 is solvent exposed in unbound ARR but buried when ARR is bound to rifampicin (see Figure 2.1). It may thus be possible to monitor the amount of ARR bound to rifampicin versus the unbound ARR in a

solution with rifampicin by monitoring the solution's fluorescence at 295 nm where only tryptophan is contributing to the fluorescence.

All fluorescence measurements were carried out on a Perkin-Elmer fluorimeter with slit-widths set to 5 nm. ARR was dialyzed into 50 mM phosphate buffer (pH 7.5) prior to fluorescence measurements. Intrinsic fluorescence of ARR was measured at 280 nm (tyrosine and tryptophan fluorescence) and 295 nm (tryptophan fluorescence) excitation wavelengths. For binding studies the fluorescence was measured at 295 nm excitation wavelength. Binding was assessed with 10 μM ARR and between 24 and 212 μM rifampicin.

To ascertain if rifampicin was absorbing any of the excitation or emission light (inner filter effects). An absorbance spectrum of rifampicin was constructed between 190 nm and 400 nm.

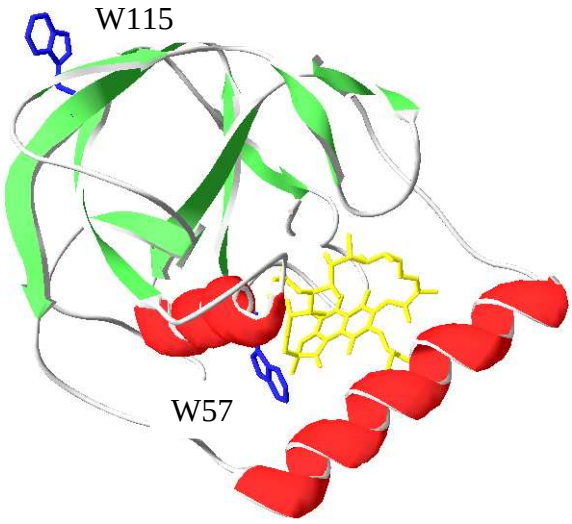


Figure 2.1 Ribbon diagram representing the crystal structure of ARR with tryptophans highlighted. Rifampicin is shown in yellow. The two tryptophans (W57 and W115) are shown in blue. Figure generated using Swiss PDB Viewer (Kaplan and Littlejohn, 2001).

2.10 Microscopic studies

Microscopic studies were undertaken to asses the possible role of ARR in cell morphology. In order to do this both *E. coli* and *M. smegmatis* strains were examined. The *E. coli* strain examined was BL21 (DE3) pLysS, one containing pQE-30 expressing ARR and one which did not. The *M. smegmatis* strains used were the wild type strain from which ARR was first isolated (DSM 43657) and a strain which carries a kanamycin resistance cassette disrupting the ARR gene (DSM 43657Kam1).

2.10.1 Light microscopy

E. coli strains were grown overnight in LB at 37 ºC with shaking at 250 rpm. *M. smegmatis* were grown for 3 days in Heart Brain Infusion broth at 37 ºC with shaking at 250 rpm. A loopful of each culture was placed onto individual glass slides and heat-fixed using an open flame. *E. coli* cells were stained using crystal violet, while *M. smegmatis* cells were stained with saffron (mycobacterial cell walls are impermeable to crystal violet). Slides were viewed with 1000 X magnification.

2.10.2 Scanning electron microscopy

Cells were cultured as described above for light microscopy. Five microlitres of each culture was dropped onto individual copper grids with carbon coating and allowed to dry. Once dry, 5 µl of a saturated uranyl acetate solution was dropped onto the grids and air-dried. The grids were viewed on a J100 scanning electron microscope (Electron Microscope Unit, University of the Witwatersrand, Johannesburg, South Africa).

2.11 Determination of heat tolerance in *E. coli*

To determine whether the changes in *E. coli* cell wall morphology caused by ARR gave the cells increased heat tolerance, *E. coli* BL21 (DE3) pLysS cells expressing ARR and *E. coli* BL21 (DE3) pLysS cells not expressing ARR were monitored for

growth at 50 ºC. This was done in LB with shaking at 250 rpm. Growth was inspected visually after 24 hours.

Chapter 3 Results

3.1 Plasmid extraction

pQE-30 (containing the ARR insert) was successfully extracted from *E. coli* BL21(DE3) pLysS and *E. coli* MM2942 cells. The plasmids did not appear to be contaminated by chromosomal DNA as shown by agarose gel electrophoresis (see Figure 3a and b). The size of the plasmid was not estimated used agarose gel electrophoresis as the plasmid was still circularized. The concentration of plasmid from the BL21 cells was spectrophotometrically determined to be 144.3 ng/µl and that of the plasmid obtained from the MM294 cells was 100.2 ng/µl.

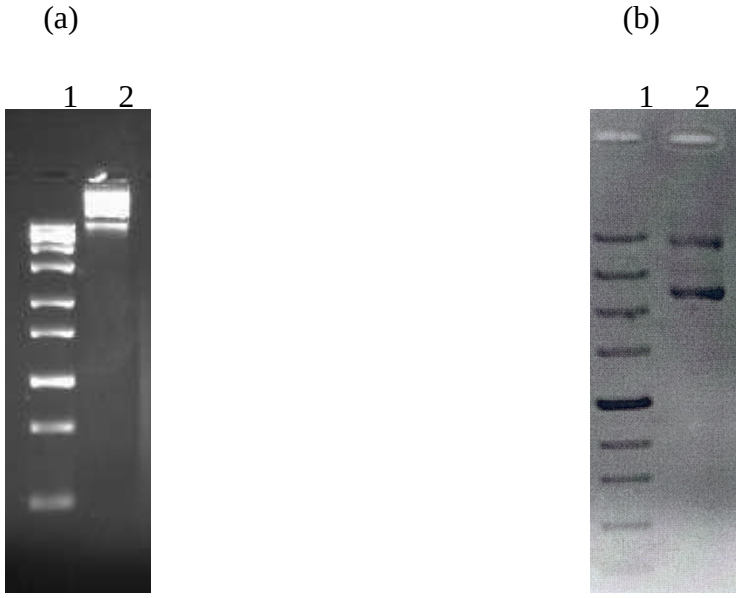


Figure 3.1: Agarose gel electrophoresis of pQE-30 containing ARR insert (a) Lane 1, Fermentas Generuler 100 bp DNA ladder; Lane 2, pQE-30 containing ARR insert purified from *E. coli* BL21 (DE3) pLysS cells (b) Lane 1, Fermentas ZipRuler 1; Lane 2, pQE-30 purified from *E. coli* MM2942 cells.

3.2 PCR

3.2.1 Oligonucleotide primer design

The forward and reverse primers (see Table 3.1) were designed to amplify the ARR fragment from pQE-30 and did not show a significant propensity for forming secondary structures or primer dimers.

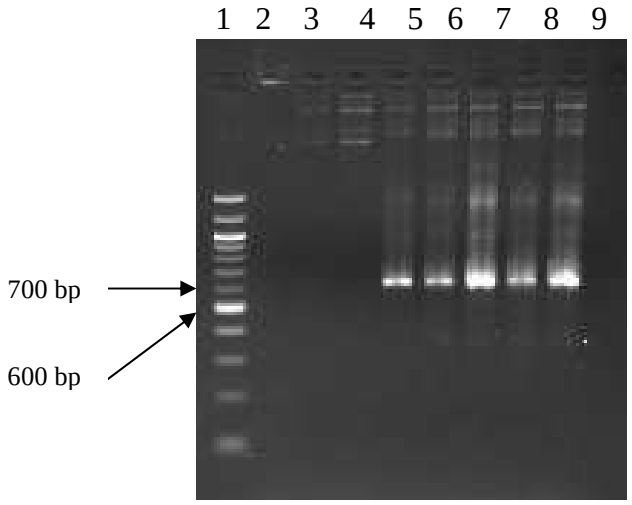
Table 3.1 Sequence of pQE-30seq primers.

Primer	*Sequence*
pQE-30seq Forward	GGAGAAATTAACTATGAGAGG
pQE-30seq Reverse	GTTCTGAGGTCATTACTGG

3.2.2 Optimisation of reaction conditions

It was found that plasmid extracted from *E. coli* BL21 (DE3) pLysS with an annealing temperature of 58 ºC gave a reasonable yield of product without excessive nonspecific product (see Figure 3.2a and b). It was still necessary to reduce non-specific binding. Reaction mix 1(Table 2.3) was found to give the purest product (see Figure 3.3). The amplicon was further purified (see Figure 3.4). The expected size of the amplicon was about 650 bp.

(a)



(b)

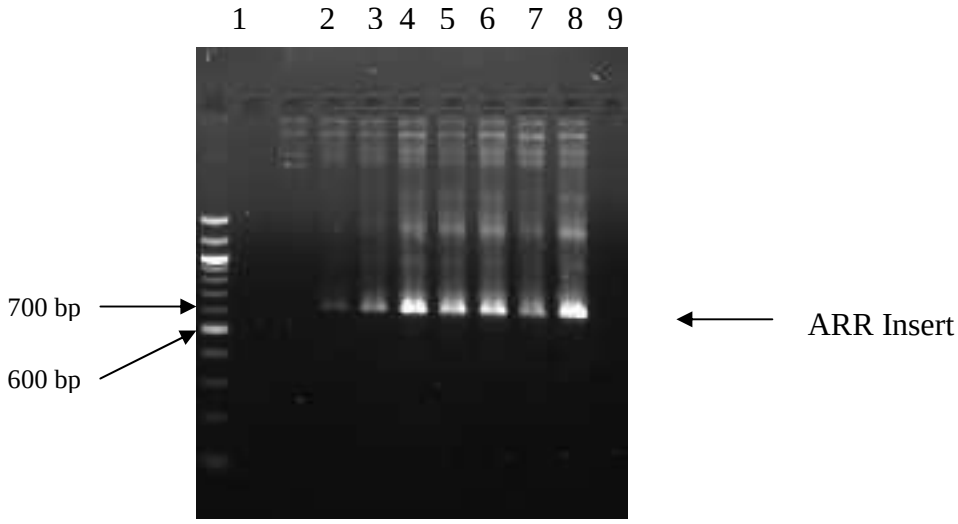


Figure 3.2 Agarose gel elecrophoresis of thermal gradient PCR (a) PCR products when template plasmid from *E. coli* MM2942 was used: Lane 1, Fermentas Generuler 100 bp DNA Ladder; Lane 2, 60 ºC annealing temperature; Lane 3, 59.3 ºC annealing temperature; Lane 4, 58 ºC annealing temperature; Lane 5, 56.2 ºC annealing temperature; Lane 6, 54 ºC annealing temperature, Lane 7, 52.1 ºC; Lane 8, 50.8 ºC; Lane 9, 50 ºC annealing temperature (b) PCR product when template plasmid from *E. coli* BL21 (DE3) pLysS was used. Lane designation same as for (a)

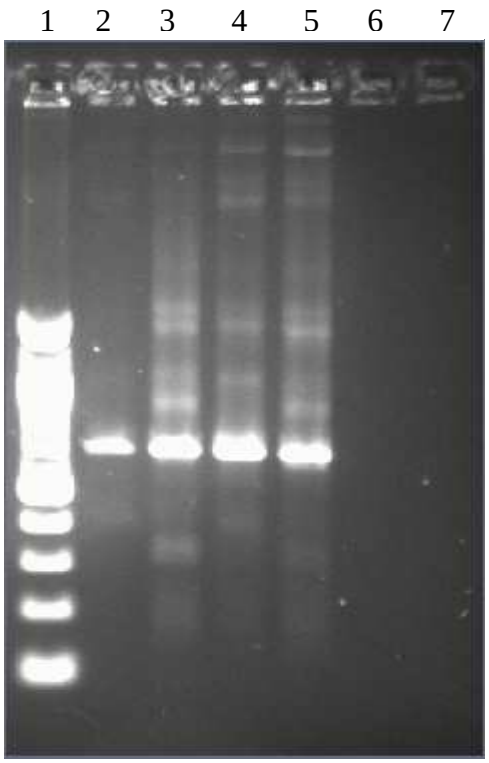


Figure 3.3 Agarose gel electrophoresis of PCR products Lane 1, Fermentas Generuler 100 bp DNA Ladder; Lane 2, reaction 1; Lane 3, reaction 2; Lane 4, reaction 3; Lane 5, reaction 4.

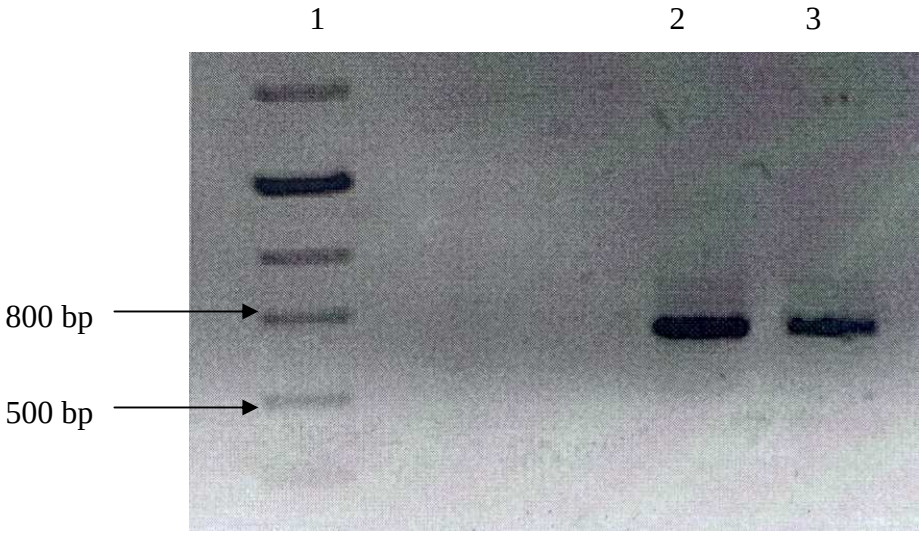


Figure 3.4 Agarose gel electrophoresis of PCR product after further purification Lane 1, Fermentas ZipRuler 1; Lanes 2 and 3 PCR product of *arr* insert after optimization and clean up.

3.3 Cloning

Eleven transformants were randomly selected after initial screening for the presence of the ARR insert. Of these eleven, three were found to contain the ARR insert by colony PCR. One of these was randomly selected for sequencing.

3.4 Sequencing

The chromatogram was checked for errors and the insert was confirmed as the known *arr* sequence with no alterations.

3.5 Over-expression

No detectible over-expression of ARR using the pQE-30 vector was found using *E. coli* MM2942 (See Figures 3.5 and 3.6), BL21 (DE3) pLysS (see Figures 3.7, 3.8 and 3.9) or SG13009 (data not shown).

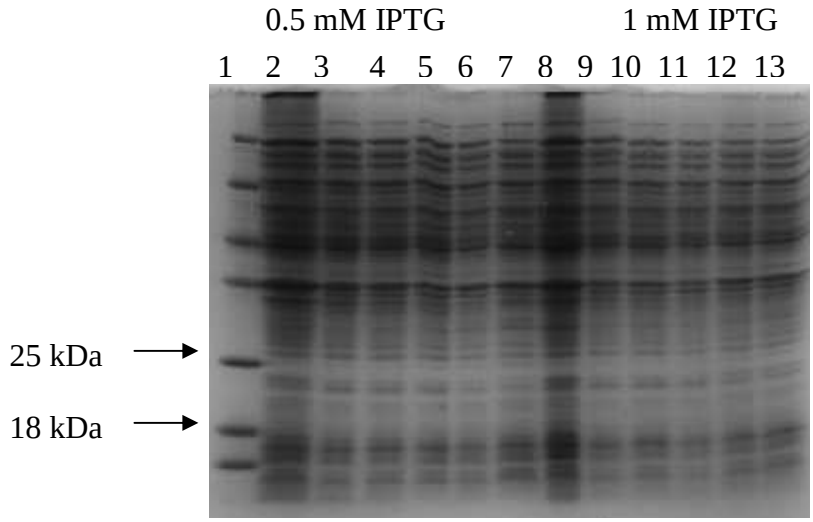


Figure 3.5 SDS-PAGE of whole cell extracts from induction study using *E. coli* MM2942 induced at mid-log phase. Lanes 1-7 0.5 mM IPTG. Lanes 8-13 1 mM IPTG. Lane 1, Fermentas Protein Marker A; Lane 2, overnight induction; Lane 3, 8 hour induction; Lane 4, 6 hour induction; Lane 5, 4 hour induction; Lane 6, 2 hour induction; Lane 7, 0 hour induction; Lane 8, >16 hours induction; Lane 9, 8 hour induction; Lane 10, 6 hour induction; Lane 11, 4 hour induction; Lane 12, 2 hour induction; Lane 13, 0 hour induction.

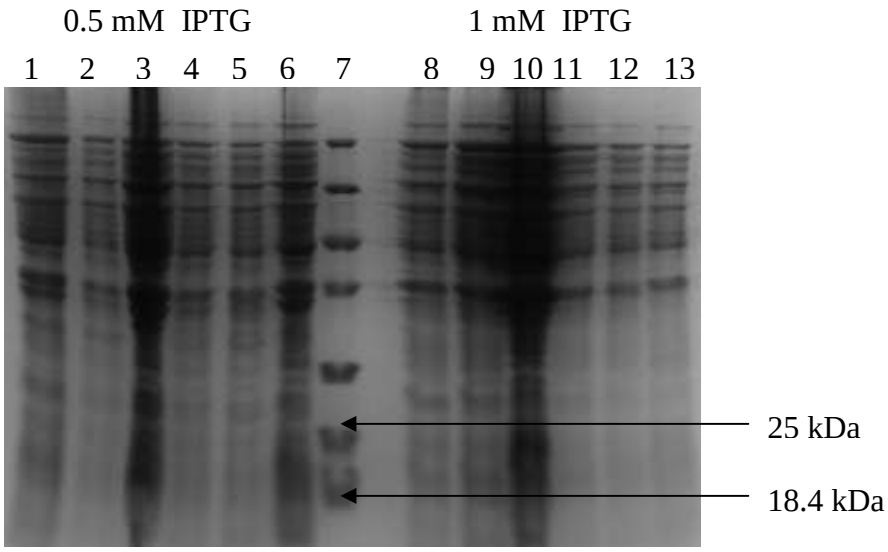


Figure 3.6 SDS-PAGE of whole cell extracts from induction study using *E. coli* MM2942 cells induced at late-log phase. Lane 1, 0 hour induction; Lane 2, 2 hour induction; Lane 3, 4 hour induction; Lane 4, 6 hour induction; Lane 5, 8 hour induction; Lane 6, >16 hours induction; Lane 7, Fermentas Protein Marker A; Lane 8, >16 hours induction; Lane 9, 8 hour induction; Lane 10, 6 hour induction; Lane 11, 4 hour induction; Lane 12, 2 hour induction; Lane 13, 0 hour induction.

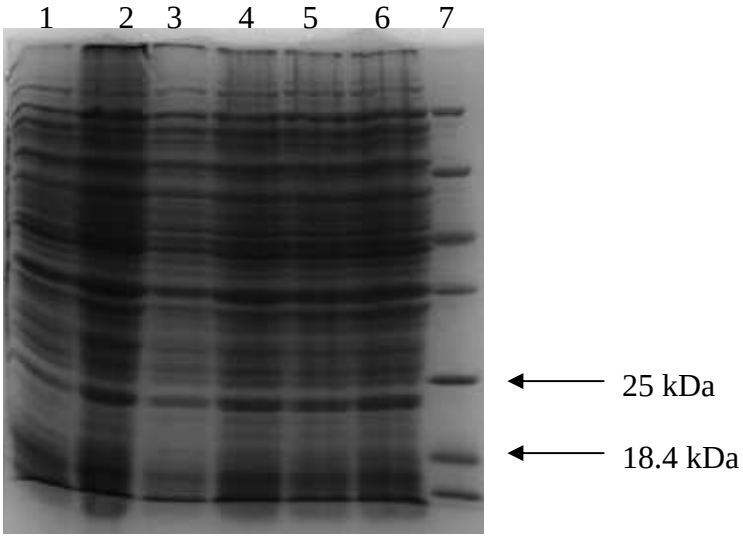


Figure 3.7 SDS-PAGE of whole cell extracts from induction study using *E. coli* BL21 (DE3) pLysS cells induced with 1mM IPTG at mid-log phase. Lane 1, 0 hours induction; Lane 2, 2 hours induction; Lane 3. 4 hours induction; Lane 4, 6 hours induction; Lane 5, 8 hours induction; Lane 6, >16 hours induction; Lane 7 Fermentas Protein Marker .

Late-log Phase Induction Mid-log Phase Induction

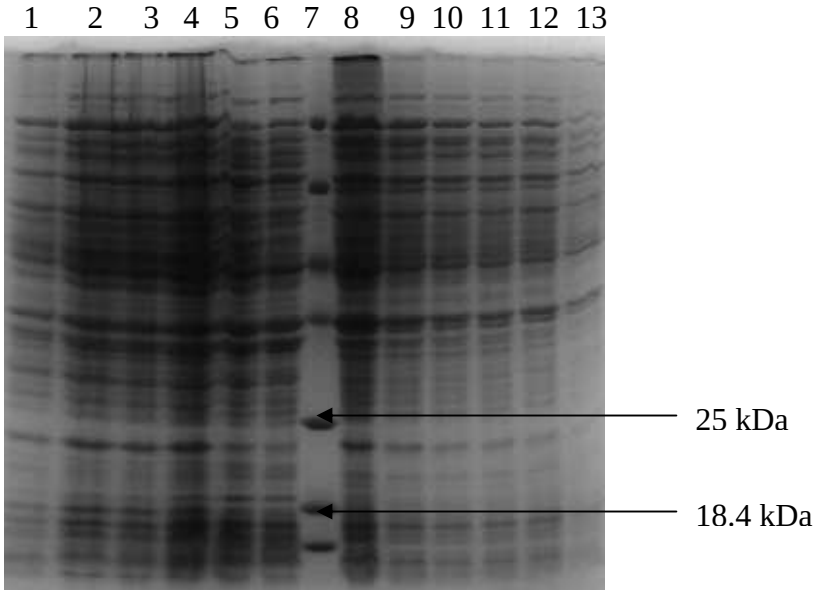


Figure 3.8 SDS-PAGE of whole cell extracts from induction study using *E. coli* BL21 (DE3) pLysS induced with 0.5 mM IPTG. Lane 1, 0 hour induction; Lane 2, 2 hour induction; Lane 3, 4 hour induction; Lane 4, 6 hour induction; Lane 5, 8 hour induction; Lane 6, overnight induction; Lane 7, Fermentas Protein Marker A; Lane 8, overnight induction; Lane 9, 8 hour induction; Lane 10, 6 hour induction; Lane 11, 4 hour induction; Lane 12, 2 hour induction; Lane 13, 0 hour induction.

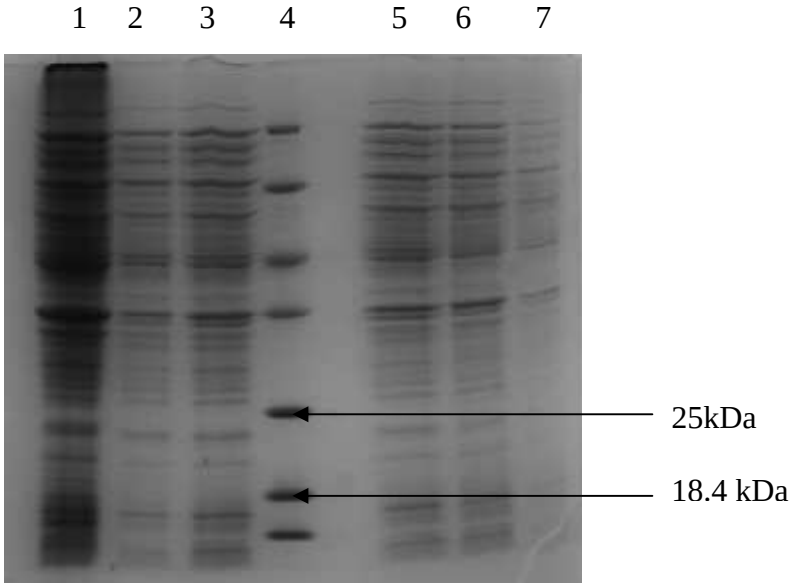


Figure 3.9 SDS-PAGE of whole cell extracts from induction study using *E. coli* BL21 (DE3) pLysS cells induced with 1 mM IPTG at late-log phase. Lane 1, Overnight induction; Lane 2, 8 hour induction; Lane 3, 6 hour induction; Lane 4, Fermentas Protein Marker A; Lane 5, 4 hour induction; Lane 6, 2 hour induction; Lane 7, 0 hour induction.

3.6 Purification

3.6.1 His-Bind purification

Purification was successful (See Figure 3.10) although ARR was heavily contaminated with other proteins (less than 85% purity by densitometry). The concentration was approximately 10 µM. Attempts to remove contaminating protein using a salt gradient through the His-Bind column and size exclusion led to significant losses of protein. His-Bind was abandoned as a purification method.

3.6.2 His-Pur purification

Purification was successful (See Figure 3.11). ARR was pure enough for further applications (greater than 95% purity by densitometry). The concentration was 27 µM but was diluted to a 10 µM working stock for further use.

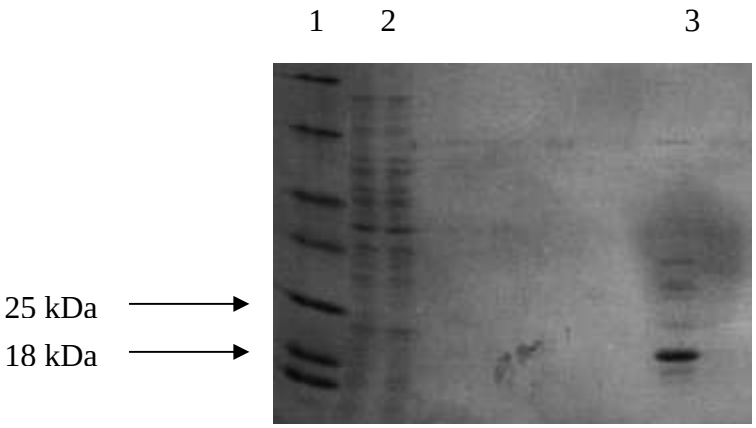


Figure 3.10 SDS-PAGE of ARR expressed in *E. coli* (DE3) pLysS purified using His-Bind resin. Lane 1, Fermentas Protein Marker A; Lane 2, Whole cell extract; Lane 3, ARR.

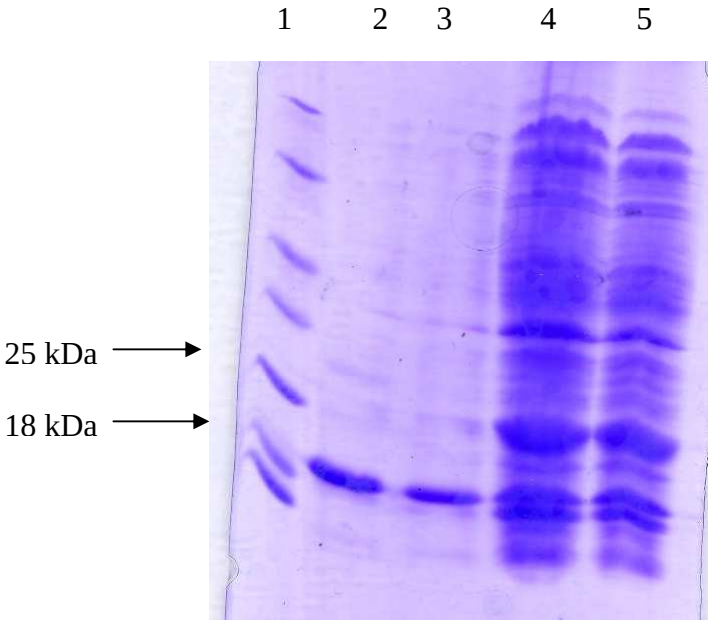


Figure 3.11 SDS-PAGE of ARR expressed in *E. coli* (DE3) pLysS purified using His-Pur resin. Lane 1, Fermentas Protein Marker A; Lane 2, and 3 ARR; Lanes 4 and 5 whole cell extract.

3.7 Spectroscopic studies

3.7.1 Far-UV Circular Dichroism

The spectrum for native ARR (See Figure 3.12) showed minima at 210 nm and 222 nm with the greater minima at 210 nm. The heat-denatured ARR spectrum showed a similar shape but the degree of ellipticity was more than half that of the native ARR (8 mdeg.cm^2.dm^{-1} compared with 21 mdeg.cm^2.dm^{-1}).

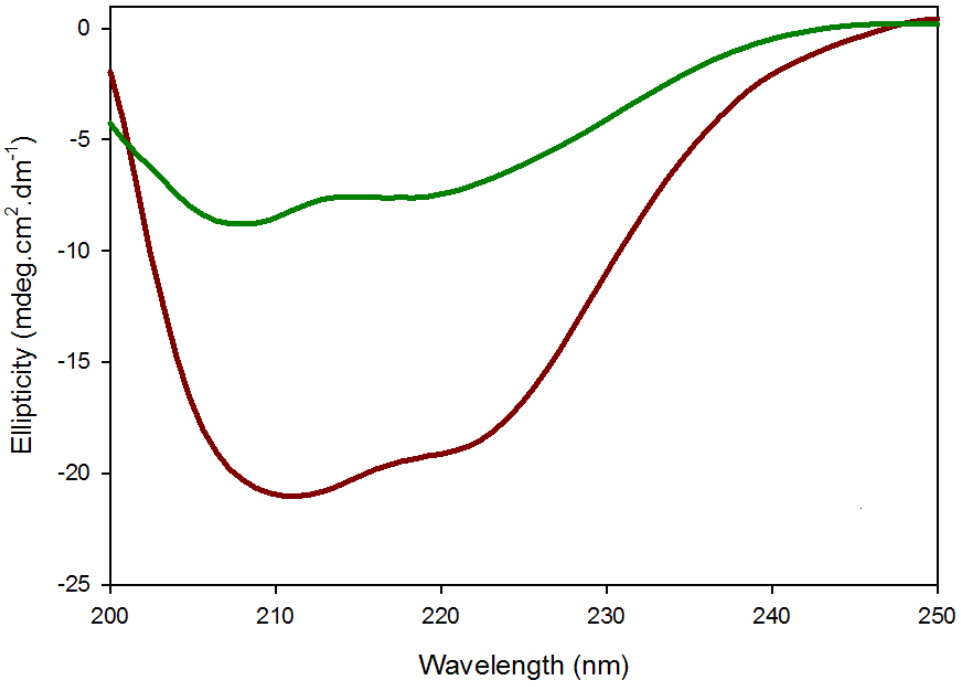


Figure 3.12 Far-UV circular dichroism spectra of 10 μM native and heat-denatured and native ARR. Red spectrum: 10 μM native ARR, green spectrum 10 μM heat denatured ARR.

3.7.2 Thermal unfolding

During thermal unfolding the degree of ellipticity decreased from 20 mdeg.cm^2.dm^{-1} to 8 mdeg.cm^2.dm^{-1}. During cooling the degree of ellipticity did not significantly change. The T_M (the temperature at which half of the protein concentration is denatured) of ARR is approximately 52 ºC (see Figure 3.13).

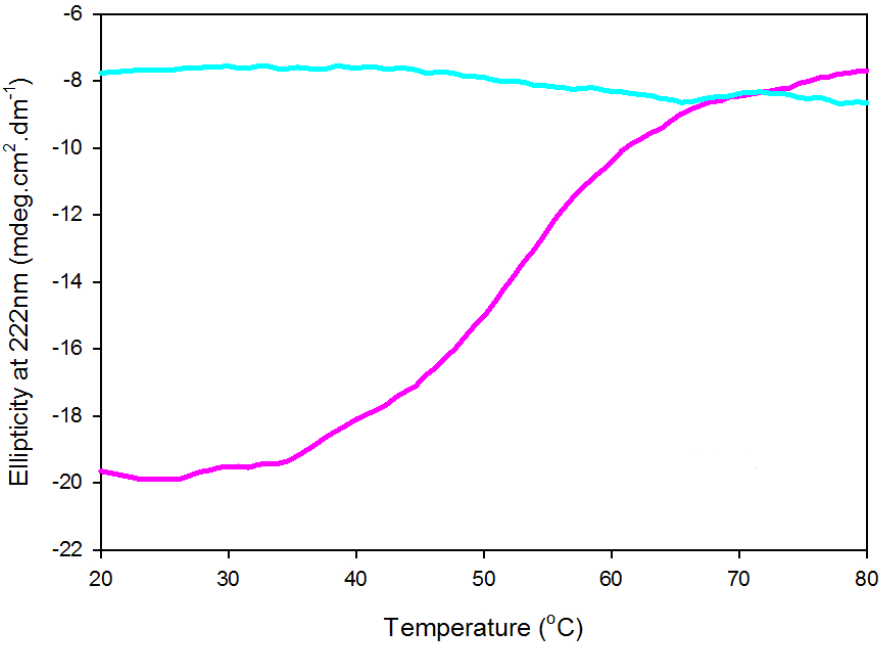


Figure 3.13 Thermal unfolding curve using far-UV CD as a probe of 10 µM ARR. Pink curve 10 µM ARR unfolding, blue curve 10 µM ARR Refolding.

3.7.3 Fluorescence

At an excitation wavelength of 280 nm (tyrosine and tryptophan fluorescence) ARR exhibited an emission maximum at approximately 350 nm and an intensity of 410 (arbitrary units). At an excitation wavelength of 295 nm (tryptophan fluorescence) the emission maximum wavelength remained unchanged at 350 nm while the intensity dropped to 300 (arbitrary units). (See Figure 3.14)

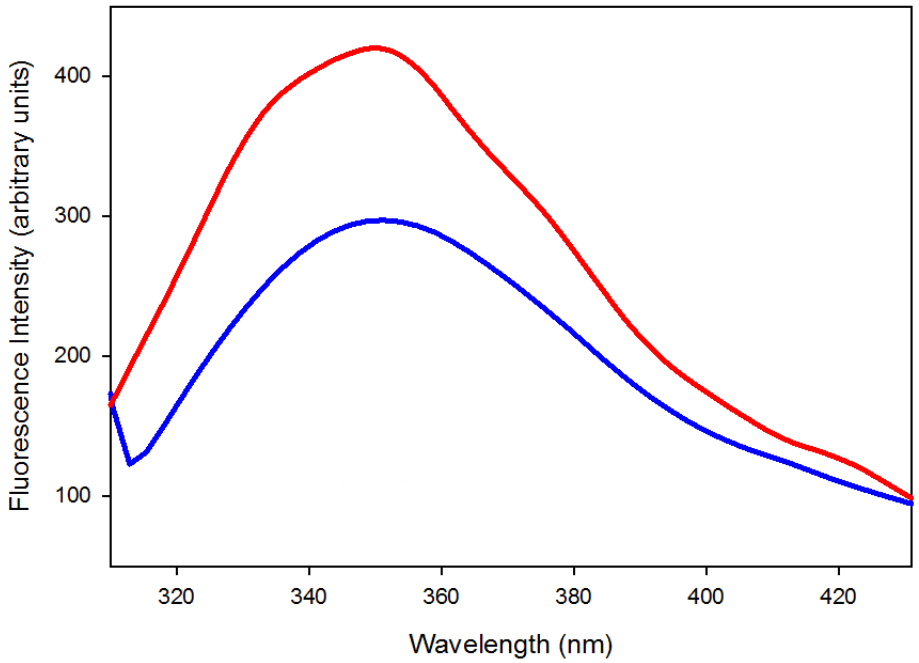


Figure 3.14 Fluorescence spectra of 10 μM ARR. Blue spectrum: 295 nm excitation wavelength, red spectrum: 280 nm excitation wavelength.

3.7.4 Binding Assay

Definite red shifting of the emission maxima was seen with increasing levels of rifampicin added to the solution (See Figure 3.15). The intensity increased as the rifampicin was increased until 58µl beyond this amount the intensity began to decrease. The absorbance of the added rifampicin at 280nm ranged from 0.14 O.D. to 1.8 O.D. (data not shown). Rifampicin showed no intrinsic fluorescence but strong absorbance at 295 nm (data not shown).

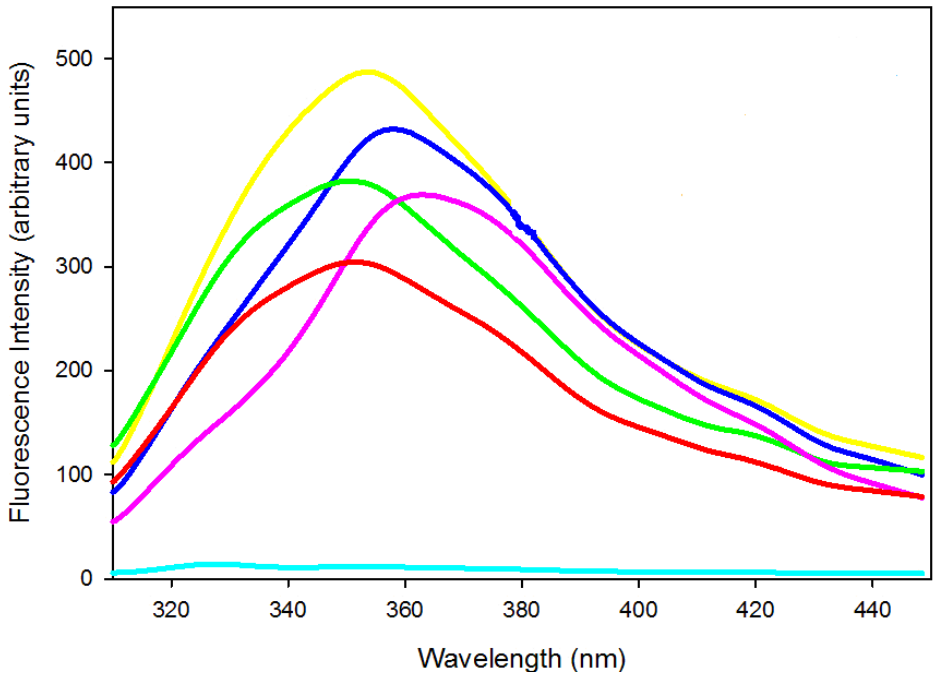


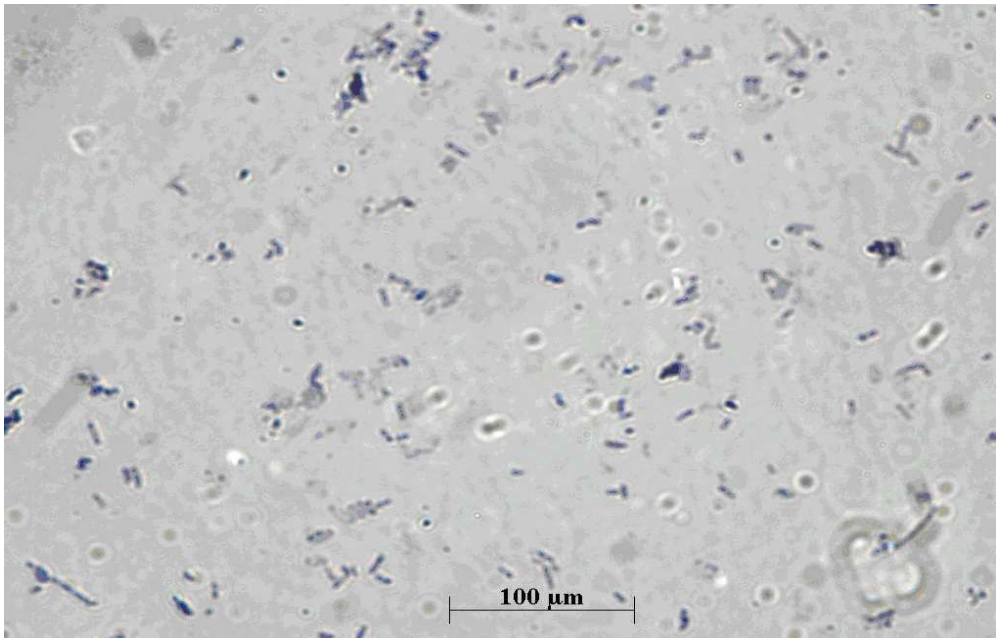
Figure 3.15 Fluorescence spectra of binding assay using 10 µM ARR and varying concentrations of rifampicin. Light blue spectrum: 24 µM rifampicin, red spectrum: 10 µM ARR, green spectrum: 10 µM ARR/ 24 µM rifampicin, yellow spectrum: 10 µM ARR/ 58 µM rifampicin, dark blue spectrum: 10 µM ARR/ 106 µM rifampicin, pink spectrum: 10 µM ARR/ 212 µM rifampicin.

3.8 Microscopic studies

3.8.1 Light microscopy

E. coli cells expressing ARR were longer than those not expressing ARR (See Figure 3.16 and 3.17). No observable change in any other cell features such as width, branching or overall shape. *M. smegmatis* cells expressing ARR showed branching compared to those not expressing ARR which were not branched (See Figure 3.18 and 3.19). There were no other visible changes to cell features.

(a)



(b)

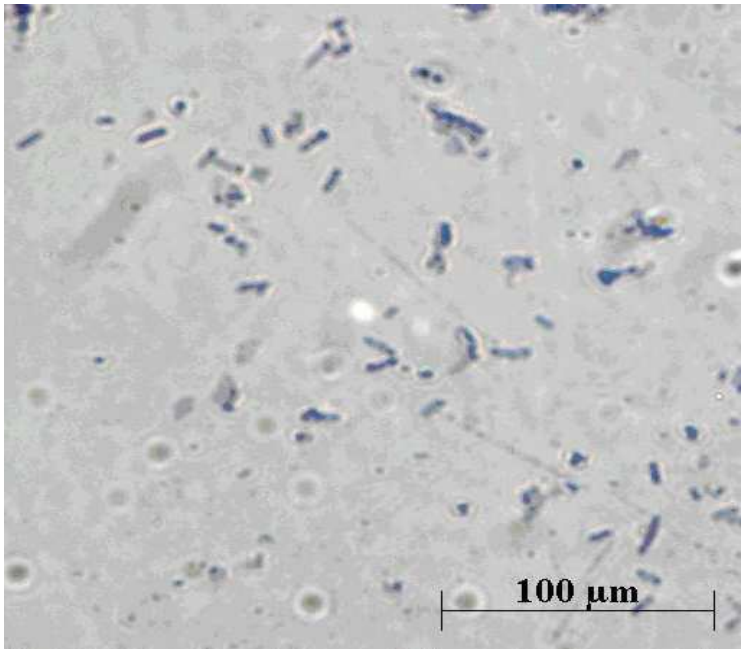
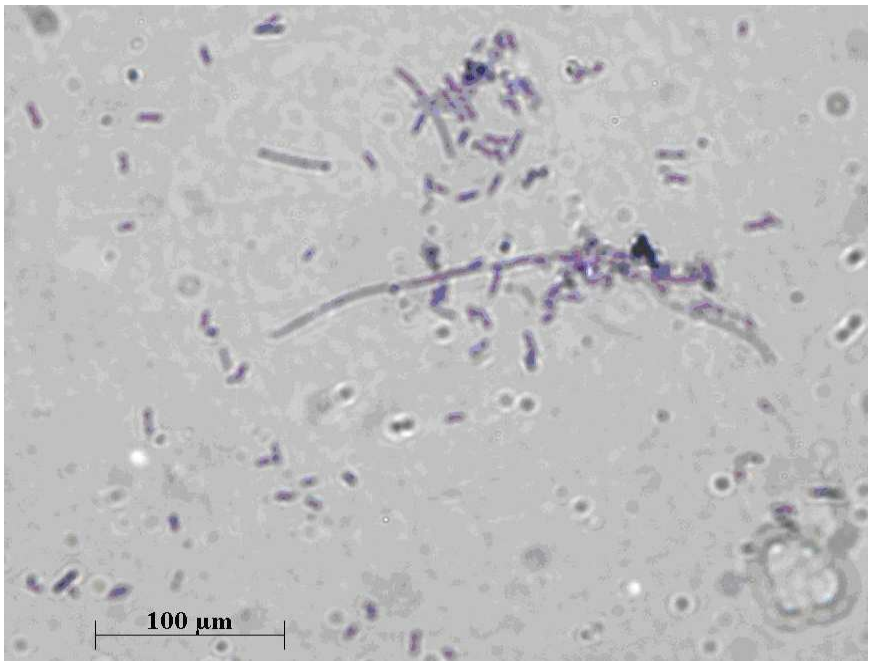


Figure 3.16 a and b Light micrographs of *E. coli* BL21 (DE3) pLysS cells not expressing ARR.

(a)



(b)

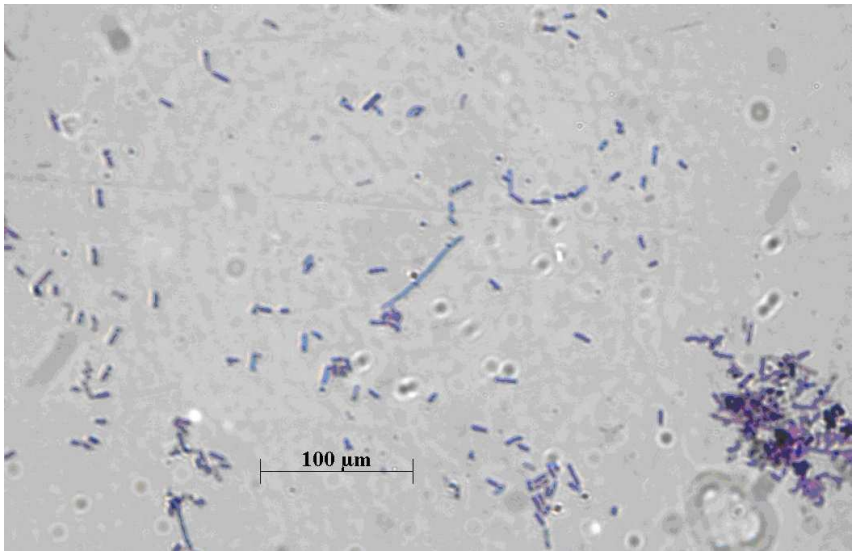


Figure 3.17 a and b Light micrographs of *E. coli* BL21 transformed with pQE-30 containing ARR insert.

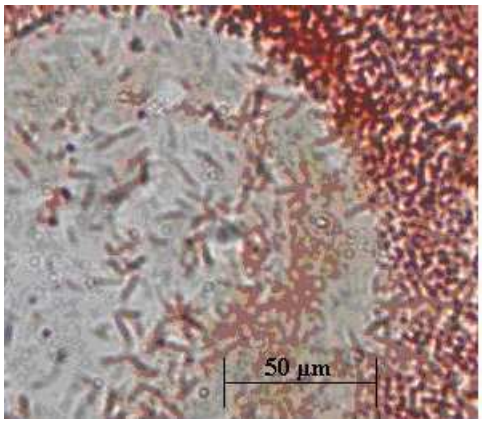


Figure 3.18 Light micrograph of *M. smegmatis* DSM 43657Km1. These cells do not express ARR.

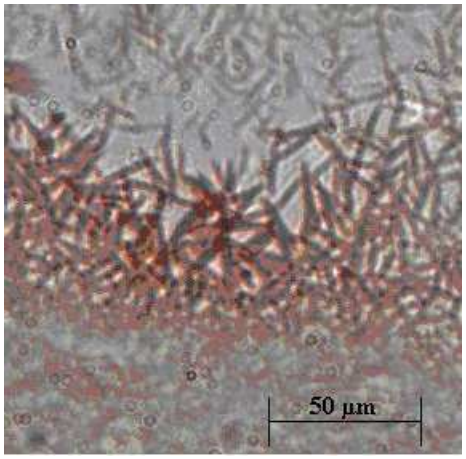
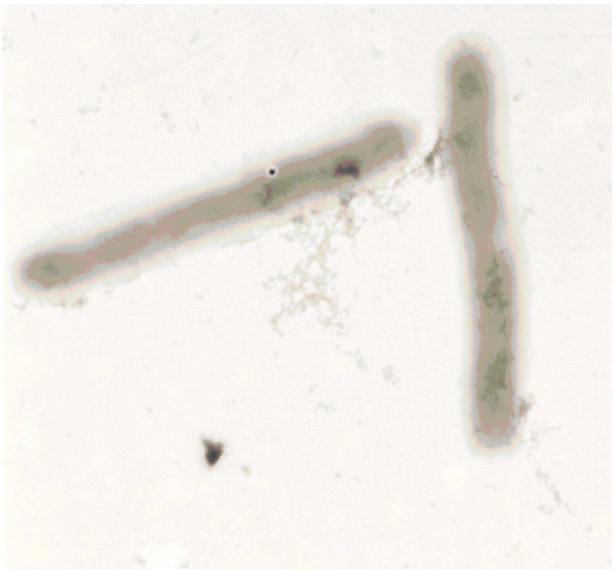


Figure 3.19 Light micrograph of *M. smegmatis* DSM 43657. These cells express ARR.

3.8.2 Electron microscopy

The results obtained from light microscopy were confirmed using electron microscopy. *E. coli* expressing ARR were longer than their non ARR-expressing counterparts. *M. smegmatis* cells which expressed ARR were branching while those which did not express ARR did not show branching (See Figures 3.20, 3.21, 3.22 and 3.23).

(a)



(b)

Figure 3.20 Transmission electron micrographs of *M. smegmatis* DSM 43657Km1 (a) 15 000 X magnification (b) 15 000 X magnification.

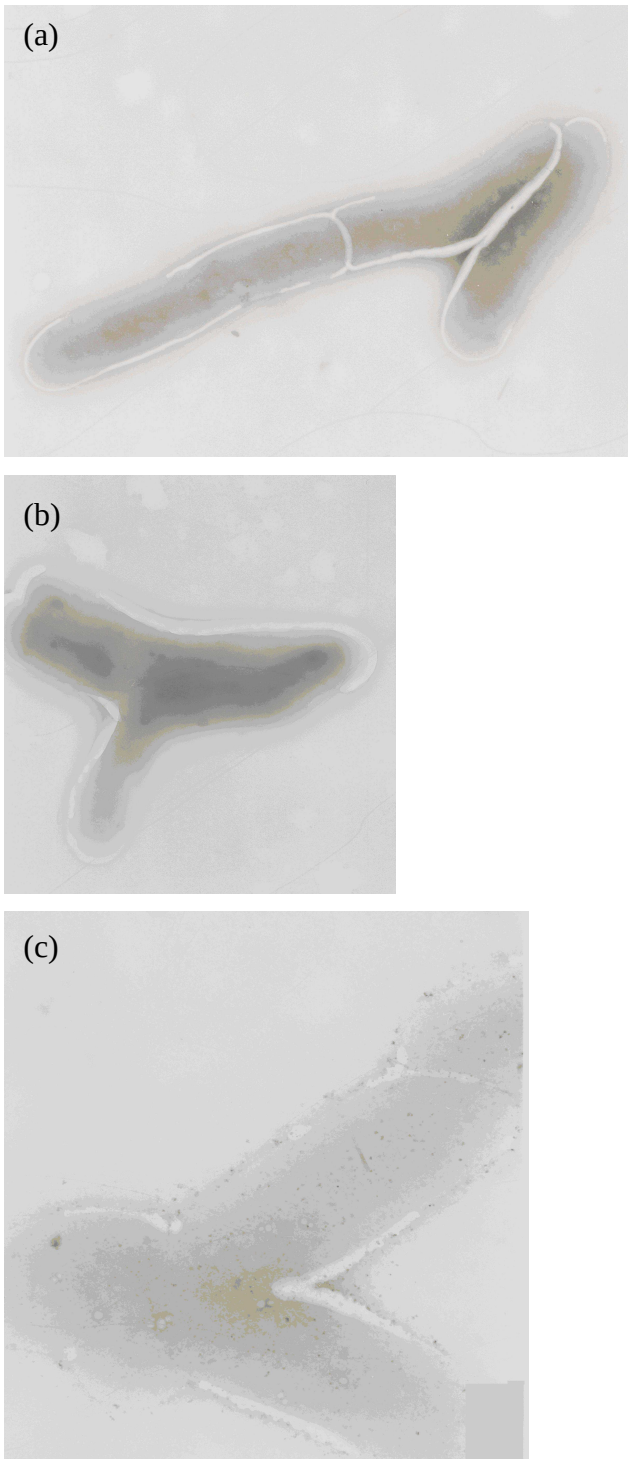


Figure 3.21 Transmission electron micrographs of *M. smegmatis* DSM 43657 (a) 15000 X magnification (b) 15 000 X magnification (c) 15 000 X magnification.

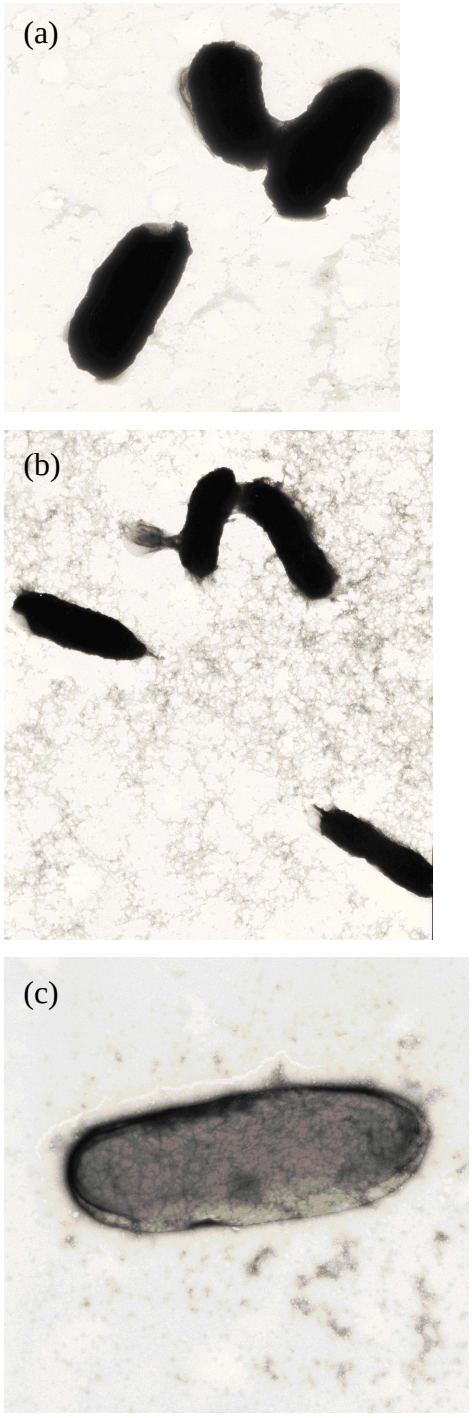


Figure 3.22 Transmission electron micrographs of *E. coli* BL21 (DE3) pLysS (a) 8000 X magnification (b) 5 000 X magnification (c) 10 000 X magnification.

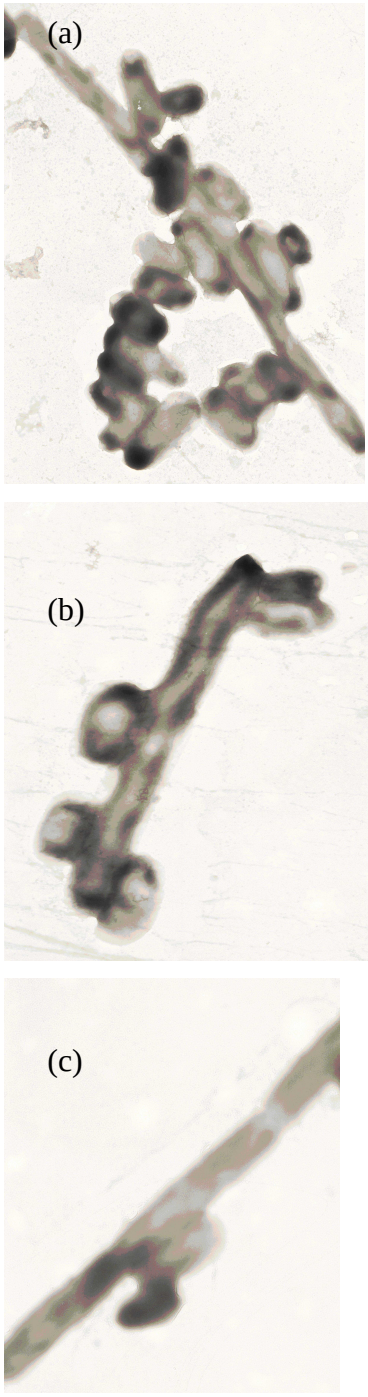


Figure 3.23 Transmission electron micrographs of *E. coli* expressing ARR (a) 5 000 X magnification (b) 8 000 X magnification (c) 8 000 X magnification.

3.9 Determination of heat tolerance in *E. coli.*

There appeared to be no significant change to heat tolerance of *E. coli* expressing ARR when compared to *E. coli* not expressing ARR. Neither culture showed any visible growth at 50 ºC.

4 Discussion

4.1 Expression of ARR

Satisfactory over-expression of ARR was not obtained using the pQE-30 plasmid. Purification of the protein was possible because of low levels of uninduced expression. Problems with expressing the protein in this vector have been reported previously (Puhaca, 2004) although in that work only *E. coli* MM2942 cells were used. Three cell lines were investigated in this work (*E. coli* MM2942; *E. coli* SG13009 (pREP) and *E. coli* (DE3) pLysS) to determine if the vector was the cause of a lack of over-expression. None of these cell lines produced favourable results, including the *E. coli* SG13009 (pREP) strain which is recommended by Qiagen for optimal expression. The choice of cell line used in over-expressing ARR from the pQE-30 vector can thus be excluded as a limiting factor.

Arr has a GUG initiation codon and in mycobacteria this is not uncommon (Quan *et al.*, 1997). However, in *E. coli* and other organisms the GUG initiation codon is not common. In eukaryotes translation is always initiated at an AUG codon however in bacteria there is slightly more flexibility and initiation can occur although rarely at GUG and even more infrequently at UUG (Ringquist *et al*., 1992)

Changing a UUG or GUG to an AUG initiation codon has been shown to drastically increase levels of expression in *E. coli* (Reddy *et al.*, 1985). In the above study it was even suggested that alternative initiation codons were used as a means to ensure low levels of expression. Not only is AUG the most efficient initiation codon in nearly all organisms but the level of transcription from an AUG start codon is also the least sensitive to the sequence of the second codon (O'Donnell and Janssen, 2000).

However in the pQE-30 vector the 6xHis tag is expressed N-terminally or at the 5′-end of the open reading frame and the vector thus uses its own start codon at the beginning of the tag sequence. This codon is an AUG initiation codon. Should the

ARR gene be cloned into another untagged or C-terminally tagged vector, the GUG codon may have to be taken into consideration.

Other factors affecting the level of expression of heterologously expressed proteins in *E. coli* involve the 5′ untranslated region (UTR) of the gene's mRNA. In other words, the vector sequence directly upstream of the initiation start site. The most important region of the UTR in terms of expression is the ribosomal binding site or Shine-Delgarno site. This site must be accessible to the 30S ribosomal subunit in order for translation to take place. It has been shown previously that expression cannot take place when mRNA forms stable stem loop structures involving this site (Wood *et al.*, 1984).

To investigate whether secondary mRNA structures play a role in the expression of ARR using the pQE-30 vector, the Mfold program (Zuker, 2003) was used to identify possible stem loop structures in the mRNA. A stem loop structure which included the Shine-Delgarno site was predicted (Figure 4.1). This structure has a predicted T_m of 65 ºC. This means that the structure would be stable at a temperature of 37 ºC (the temperature at which the cells in this study were grown during expression).

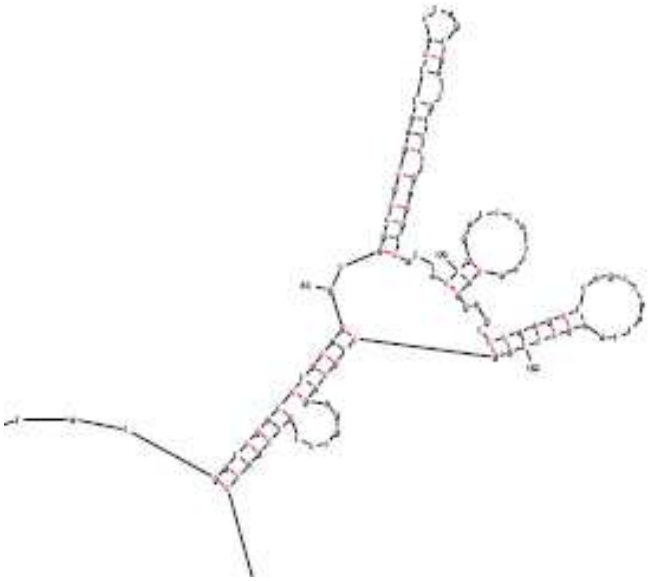


Figure 4.1 Hypothetical secondary structure of *arr* mRNA when expressed using the pQE-30 vector as predicted by the Mfold software (Zuker, 2003).

This seems the most likely cause of low expression levels when attempting to express ARR using the pQE-30 vector. To overcome this problem, another vector which does not have any complimentarity with the 5′ end of the ARR gene should be used. This is in agreement with reports that ARR has been successfully overexpressed in the pET series of vectors (Gianniosis, 2007; Baysarowich *et al.*, 2008). The hypothetical secondary structure of a pET ARR construct could be inspected for stem loop structures using the mFold program. This structure would be dependent on the choice of pET vector, as well as, which restriction sites are used within the pET vector.

Further attempts to validate the identity of the protein purified in this work could have been made using a western blot specific for his-tagged proteins, however, there is reasonable certainty that the correct protein obtained was ARR because of size and spectrographic properties.

4.2 Spectroscopic studies

4.2.1 Far-UV CD

The far-UV CD spectrum of native ARR shows two distinct minima, one at 210 nm and one at 222 nm. This spectrum is typical of proteins which contain both α-helices and β-strands The minimum, which occurs at 210 nm, is expected to be more negative in intensity than the minimum occurring at 222 nm in proteins of mixed secondary structure (Manavalan and Johnson, 1983). The far-UV CD spectrum of ARR is consistent with the crystal structure of ARR which shows ARR to be 18 % α-helical and 29% β-stranded.

Although the far-UV CD spectrum of ARR was predictable once the crystal structure of ARR had been solved (which occurred during the course of this work) and provided no novel information about the secondary structure of the protein, it may be useful as a reference point in NADP binding studies using far-UV CD as a probe similar to those done on dehydrogenase and NADH (for example Lange *et al*., 1974)

4.2.2 Thermal stability

Both the thermal unfolding curve and the far-UV CD spectrum of ARR, once it had been heated to 95 ºC, show that ARR retains about 50% structure after heating, however once denatured ARR does not refold. These findings are consistent with previous reports that ARR maintains activity after heating (Quan *et al.* ,1997, Gianoissis, 2007). This indicates that the binding sites of NAD^+ and rifampicin as well as the active site retain their structural integrity during heating. Due to its relatively high thermal stability and putative role in cell wall morphology investigations were done to determine whether ARR conferred any thermostability to the cells in which it is expressed. It was found that ARR had no influence on the cell's ability to grow at higher temperatures.

4.2.3 Intrinsic fluorescence

The intrinsic fluorescence spectra of ARR correspond well to its crystal structure. ARR contains five tyrosine residues and two tryptophan residues. The tryptophan residues are solvent exposed. The intensity of the emission maximum is much higher with an excitation wavelength of 280 nm than 295 nm because at 280 nm both tyrosine and tryptophan residues are contributing to the emission wavelength, whereas, at 295 nm only tryptophan residues are contributing.

The wavelength emission maximum of tryptophan is highly dependent on solvent accessibility. The emission maximum of free tryptophan in water is near 360 nm (Lakowicz, 1983). The more solvent exposed tryptophan residues are within a protein the closer the emission maximum will be to 360 nm. The emission maximum of 350 nm of the tryptophans in ARR is consistent with two solvent exposed tryptophan residues.

4.2.4 Binding studies

The red shift observed would be explained by an increase in the solvent accessibility of the tryptophan residues in ARR. This is contradictory to what was expected, as tryptophan 57 is buried when rifampicin binds to ARR, which would lead to blue shift. However, no conclusions can be drawn from this study because of the large inner filter effects caused by rifampicin's absorbance at 295 nm.

4.3 Changes in cell wall morphology

Although there have been previous reports of cell wall changes to *E. coli* heterologously expressing ARR, this is the first report of cell morphology changes to *M. smegmatis* in the absence of ARR production. The only other known pathway in which ADP-ribosylation plays a role in cell morphology is the *bld* pathway in *S. coelicolor*. It appears that the ADP-ribosylation which occurs via ARR in *M. smegmatis* may be very similar to the ADP-ribosylation which takes place as part of the *bld* pathway in *S. coelicolor*.

Three of the four proteins identified as ADP-ribosylation targets in the *bld* pathway (Sugawara *et al.*,2002) have homologues in *M. smegmatis*, identified using a BLAST alignment (Altschul *et al.*, 1990) see Figure 4.2- 4.4. BldKB shows significant sequence identity with the *M. smegmatis* Extracellular Binding Protein, MalE shows significant sequence identity with the *M. smegmatis* Sugar Binding Protein and the Solute Binding Protein with the *M. smegmatis* D-xylose-binding periplasmic protein. The fourth target, a periplasmic protein required in binding branched-chain amino-acids, appears to have no homologues in *M. smegmatis*.

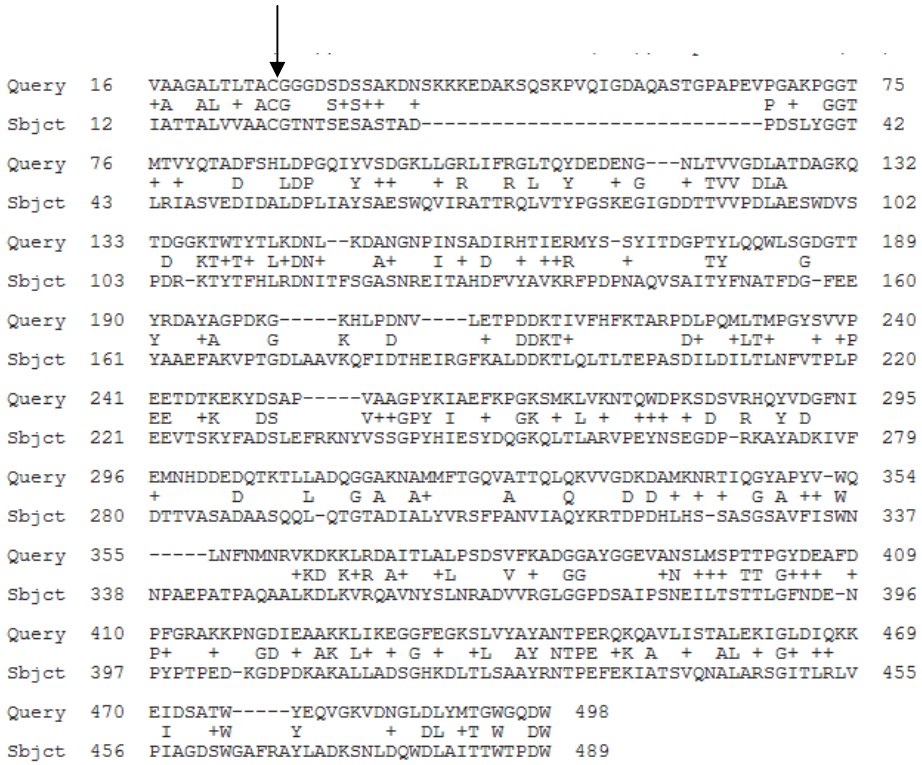


Figure 4.2 BLAST alignment of *S. coelicolor* BldKB NP_629263.1| (Query) with *M. smegmatis* extracellular binding protein YP_888816.1 (Subject). Hypothetical ADP-ribosylated cysteine is marked with an arrow.

Figure 4.3 BLAST alignment of *S. coelicolor* MalE NP_626482 (Query) with *M. smegmatis* sugar binding lipoprotein YP_884927.1 (Subject). Hypothetical ADP-ribosylated cysteine is marked with an arrow.

Figure 4.4 BLAST alignment of *S. coelicolor* solute binding protein NP_630123.1 (Query) with *M. smegmatis* D-xylose-binding periplasmic protein YP_890244.1 (Subject). The probable ADP-ribosylated cysteine is marked with an arrow.

The cysteine residue in the four ADP-ribosylated proteins in *S. coelicolor* which is thought to be ADP-ribosylated (Sugawara *et al.*,2002) is conserved in the extracellular solute binding protein, the sugar binding protein and the extracellular binding protein (in *M. smegmatis*) making the same mechanism of ADP-ribosylation possible in both species. A statistical study showed that it is highly probable that this cysteine is part of a signal sequence in the biosynthetic pathway which forms *N*-acyl-diacylglycerylcysteine a hallmark of bacterial lipoproteins (Sugawara *et al.,* 2002). If the *M. smegmatis* extracellular solute binding protein is ADP-ribosylated it would not be able to be exported to the cell surface and assist in the transport of solutes across the cell membrane.

Moreover, there may have been evolutionary pressure for *M. smegmatis* to keep the *arr* gene as it appears that the cell wall changes conferred by this gene may help bacteria endure stress. This conclusion can be drawn from the finding made by Mukherjee and Chatterji in 2005 that *arr* is down-regulated in the sigma B null mutant of *M. smegmatis*.

Bacteria often have several alternative sigma factors of core RNA polymerase which enable them to differentially regulate large sets of genes (Review by Winkler and Breaker, 2005). Sigma factors are polypeptides which associate with the DNA-dependent RNA polymerase. The sigma factors are responsible for promoter recognition. Different sigma factors recognise different sets of promoters. In this way it is possible for large sets of genes to be regulated by the regulation of corresponding sigma factors.

One of the most common needs for such large scale regulation is environmental factors, especially stress. In *M. smegmatis*, the major stress induced sigma factor is SigB (Manganelli *et al.,* 1999). Genes which are regulated by *sigB* are generally considered to be stress related. Many of the proteins found to be up-regulated by *sigB* are related to the cell wall (Mukherjee and Chatterji, 2005). It

is interesting to note that there is also an alternative sigma factor (bldN) involved in the *bld* pathway in *S. coelicolor* (Bibb and Butter, 2000).

Although the cells which showed cell wall morphology changes were not under typical stress conditions (unfavourable temperatures, high salt concentration), they were all in the stationary phase of growth which can lead to stress responses in cells due to high levels of metabolic waste in the media as well as a lack of nutrients (especially a carbohydrate source) and oxygen. A link between cell wall morphology and nutrient starvation in *M. smegmatis* has already been established (Ojha *et al.*, 2000). Cells which are nutrient-starved are shorter than cells under normal growth conditions.

There is evidence that lateral gene transfer between *M. smegmatis* and *S. coelicolor* is possible. They are both soil dwelling bacteria. *S. coelicolor* has a linear genome which has long been known to have a high level of genome instability (Chen *et al.*,2002) and *Nocardia,* evolutionarily closely related to *S. coelicolor,* and *Mycobacterium,* can conjugate with either species being the donor (Schupp *et al.*,1975). However, it does not seem likely that *arr* originated from *S. coelicolor* as at least three other proteins involved in the ADP-ribosylation which leads to changes in cell wall morphology in *S. coelicolor* are also present in *M. smegmatis*.

The locus which was determined to be responsible for ADP-ribosylation of proteins during morphological differentiation in *S. coelicolor* was named *bergA* by Ochi and colleagues and a rough location on the genome was mapped. However, no sequence was elucidated for this gene. That work was concluded before the sequencing of the *S. coelicolor* genome was undertaken. Although the genome has been largely annotated it is unclear which sequenced gene corresponds to the *bergA* locus. It is possible that the ARR homologue, ARR2, is the product of the *bergA* locus.

It appears that *M. smegmatis* has a pathway involving ADP-ribosylation which is similar to the pathway which leads to aerial spore formation in *S. coelicolor*. In *M. smegmatis*, this pathway may be involved in stress response (possibly the physiological stress the cell experiences during the stationary phase of growth). It is unclear whether ADP-ribosylation of the putative targets of ARR (identified in this work) is responsible for the changes in cell wall morphology or whether they are involved in other, as yet, undiscovered aspects of this pathway.

In order to determine whether *M. smegmatis* has a pathway similar to that of the *bld* pathway of *S. coelicolor*, it would be necessary to evaluate the ADP-ribosylation patterns of *M. smegmatis* with and without the production of ARR using the method of Sugawara *et al.*, 2002. If the three *M. smegmatis* proteins identified as putative ARR targets in this work are differentially ADP-ribosylated, the role of the cysteine identified as a possible site of ADP-ribosylation could be determined using site-directed mutagenesis. It would also be interesting to determine whether ARR is involved in a stress response to stationary phase conditions by assessing mRNA levels during different growth phases by use of quantitative real-time PCR.

If ARR is part of a stress response pathway, it is not clear how or when the protein acquired the ability to inactivate rifampicin. It does not appear that this function is present in *S. coelicolor* as there have been no reports of innate rifampicin resistance via ADP-ribosylation.

5 References

Abdallah AM, Gey van Pittius NC, Champion PA, Cox J, Luirink J, Vandenbroucke-Grauls CM, Appelmelk BJ, Bitter W. (2007) Type VII secretion--mycobacteria show the way. Nat Rev Microbiol. 5(11):883-91.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990) Basic local alignment search tool. J. Mol. Biol. 256(3):403-410.

Amé JC, Spenlehauer C, de Murcia G. (2004) The PARP superfamily. Bioessays. 26(8):882-93.

Andersen SJ, Quan S, Gowan B, Dabbs ER. (1997) Monooxygenase-like sequence of a Rhodococcus equi gene conferring increased resistance to rifampin by inactivating this antibiotic. Antimicrob Agents Chemother. 41(1):218-21.

Arlet G, Nadjar D, Herrmann JL, Donay JL, Lagrange PH, Philippon A. (2001) Plasmid-mediated rifampin resistance encoded by an arr-2-like gene cassette in Klebsiella pneumoniae producing an ACC-1 class C beta-lactamase. Antimicrob Agents Chemother. 45(10):2971-2.

Asselineau J, Lederer E. (1950) Structure of the mycolic acids of mycobacteria. Nature. 4;166(4227):782-3.

Aubry-Damon H, Soussy CJ, Courvalin P. (1998) Characterization of mutations in the rpoB gene that confer rifampin resistance in Staphylococcus aureus. Antimicrob Agents Chemother. (10):2590-4.

Balian SC, Ribeiro P, Vasconcellos SA, Pinheiro SR, Ferreira Neto JS, Guerra JL, Xavier JG, Morais ZM, Telles MA. (1997) Tuberculosis lymphadenitis in slaughtered

swine from the State of São Paulo, Brazil: microscopic histopathology and demonstration of mycobacteria. Rev Saude Publica. 31(4):391-7.

Barry CE 3rd, Mdluli K. (1996) Drug sensitivity and environmental adaptation of mycobacterial cell wall components. Trends Microbiol. 4(7):275-81.

Baysarowich J, Koteva K, Hughes DW, Ejim L, Griffiths E, Zhang K, Junop M, Wright GD. (2008) Rifamycin antibiotic resistance by ADP-ribosylation: Structure and diversity of Arr. Proc Natl Acad Sci U S A. 25;105(12):4886-91.

Bibb MJ, Buttner MJ. (2003) The *Streptomyces coelicolor* developmental transcription factor sigmaBldN is synthesized as a proprotein. J Bacteriol. 185(7):2338-45.

Birnboim HC, Doly J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7(6):1513-23.

Bliziotis IA, Ntziora F, Lawrence KR, Falagas ME. (2007) Rifampin as adjuvant treatment of Gram-positive bacterial infections: a systematic review of comparative clinical trials. Eur J Clin Microbiol Infect Dis. 26(12):849-56.

Brosch R, Pym AS, Gordon SV, Cole ST. (2001) The evolution of mycobacterial pathogenicity: clues from comparative genomics. Trends Microbiol. 9(9):452-8.

Brown BA, Springer B, Steingrube VA, Wilson RW, Pfyffer GE, Garcia MJ, Menendez MC, Rodriguez-Salgado B, Jost KC Jr, Chiu SH, Onyi GO, Böttger EC, Wallace RJ Jr. (1999) *Mycobacterium wolinskyi* sp. nov. and *Mycobacterium goodii* sp. nov., two new rapidly growing species related to *Mycobacterium smegmatis* and associated with human wound infections: a cooperative study from the International Working Group on Mycobacterial Taxonomy. Int J Syst Bacteriol. 49 Pt 4:1493-511.

Bürkle A. (2005) Poly(ADP-ribose). The most elaborate metabolite of NAD+. FEBS J. 272(18):4576-89.

Champness WC. (1988) New loci required for *Streptomyces coelicolor* morphological and physiological differentiation. J Bacteriol. 170(3):1168-74.

Chen CW, Huang CH, Lee HH, Tsai HH, Kirby R. (2002) Once the circle has been broken: dynamics and evolution of *Streptomyces* chromosomes. Trends Genet. 18(10):522-9.

Churchyard GJ, Fielding K, Charalambous S, Day JH, Corbett EL, Hayes RJ, Chaisson RE, De Cock KM, Samb B, Grant AD. (2003) Efficacy of secondary isoniazid preventive therapy among HIV-infected Southern Africans: time to change policy? AIDS. 26;17(14):2063-70.

Churchyard GJ. (2003) Stable incidence rates of tuberculosis (TB) among human immunodeficiency virus (HIV)-negative South African gold miners during a decade of epidemic HIV-associated TB. J Infect Dis. 15;188(8):1156-63.

Clark JM. (1988) Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. Nucleic Acids Res. 16(20):9677-86.

Collier RJ. (1967) Effect of diphtheria toxin on protein synthesis: inactivation of one of the transfer factors. J Mol Biol. 14;25(1):83-98.

Combrink KD, Denton DA, Harran S, Ma Z, Chapo K, Yan D, Bonventre E, Roche ED, Doyle TB, Robertson GT, Lynch AS. (2007) New C25 carbamate rifamycin derivatives are resistant to inactivation by ADP-ribosyl transferases. Bioorg Med Chem Lett1. 7(2):522-6.

Corbett EL, Charalambous S, Fielding K, Clayton T, Hayes RJ, De Cock KM,

Corda D, Di Girolamo M. (2003) Functional aspects of protein mono-ADP-ribosylation. EMBO J. 22(9):1953-8.

Dabbs ER, Yazawa K, Tanaka Y, Mikami Y, Miyaji M, Andersen SJ, Morisaki N, Iwasaki S, Shida O, Takagi H, et al. (1995) Rifampicin inactivation by *Bacillus* species. J Antibiot. 48(8):815-9.

Dabbs ER and Quan S. (2000) Light inhibits rifampicin inactivation and reduces rifampicin resistance due to a cloned mycobacterial ADP-ribosylation gene. FEMS Microbiol. Lett. 182:105-9.

Dean GL, Edwards SG, Ives NJ, Matthews G, Fox EF, Navaratne L, Fisher M, Taylor GP, Miller R, Taylor CB, de Ruiter A, Pozniak AL. (2002) Treatment of tuberculosis in HIV-infected persons in the era of highly active antiretroviral therapy. AIDS. 4;16(1):75-83.

Driessen, AJ, Rosen, BP Konings WN. (2000) Diversity of transport mechanisms: common structural principles. Trends Biochem. Sci., 25, 397-401.

Dye C, Hosseini M, Watt C. (2007) Did we reach the 2005 targets for tuberculosis control? Bull World Health Organ. 85(5):364-9.

Dye C, K Lönnroth K, Jaramillo E, Williams BG, Raviglione M. (2009) Trends in tuberculosis incidence and their determinants in 134 countries. Bull World Health Organ 87 (9)

Frieden TR. (2002) Can tuberculosis be controlled? Int J Epidemiol. 31(5):894-9.

Elliot M, Damji F, Passantino R, Chater K, Leskiw B. (1998) The *bldD* gene of *Streptomyces coelicolor* A3(2): a regulatory gene involved in morphogenesis and antibiotic production. J Bacteriol. 180(6):1549-55.

Ellis MW, Griffith ME, Jorgensen JH, Hospenthal DR, Mende K, Patterson JE. (2009) Presence and molecular epidemiology of virulence factors in methicillin-resistant *Staphylococcus aureus* strains colonizing and infecting soldiers. J Clin Microbiol. 47(4):940-5.

Emsley P, Cowtan K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D60, 2126-2132.

Faraone-Mennella MR, De Maio A, Petrella A, Romano M, Favaloro P, Gambacorta A, Lama L, Nicolaus B, Farina B. (2006) The (ADP-ribosyl)ation reaction in thermophilic bacteria. Res Microbiol. 157(6):531-7.

Gianniosis M, Berardi V, Bonincontro A, Dabbs E, Risuleo G.(2008) Charge redistribution in adenosylribosyl transferase caused by substitution of a single amino acid residue. Z Naturforsch C. 63(11-12):889-92.

Glowacki G, Braren R, Firner K, Nissen M, Kühl M, Reche P, Bazan F, Cetkovic-Cvrlje M, Leiter E, Haag F, Koch-Nolte F. (2002) The family of toxin-related ecto-ADP-ribosyltransferases in humans and the mouse. Protein Sci. 11(7):1657-70.

Gianniosis, M. (2006). Studies on ARR gene expression. [MSc Thesis] University of the Witwatersrand.

Hanahan D. (1983) Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol. 166(4):557-80.

Harasym M, Zhang LH, Chater K, Piret J. (1990)The *Streptomyces coelicolor* A3(2) *bldB* region contains at least two genes involved in morphological development. J Gen Microbiol. 136(8):1543-50.

Hartmann G, Honikel KO, Knüsel F, Nüesch J. (1967) The specific inhibition of the DNA-directed RNA synthesis by rifamycin. Biochim Biophys Acta. 145(3):843-4.

Houang ET, Chu YW, Lo WS, Chu KY, Cheng AF. (2003) Epidemiology of rifampin ADP-ribosyltransferase (arr-2) and metallo-beta-lactamase (blaIMP-4) gene cassettes in class 1 integrons in *Acinetobacter* strains isolated from blood cultures in 1997 to 2000. Antimicrob Agents Chemother. 47(4):1382-90.

Jin DJ, Gross CA. (1998) Mapping and sequencing of mutations in the *Escherichia coli rpoB* gene that lead to rifampicin resistance. J Mol Biol. 5;202(1):45-58.

Johnson D and McClure W. (1976) in RNA Polymerase (Lossick R and Chamberlain M. eds) pp 413-428. Cold Spring Harbor Press, Cold Spring Harbor NY.

Kaplan W and Littlejohn TJ. (2001) Swiss-PDB Viewer (Deep View) Briefings in Bioinformatic 2(2):195-197.

Kelemen GH, Buttner MJ. (1998) Initiation of aerial mycelium formation in *Streptomyces*. Curr Opin Microbiol. 1(6):656-62.

Laemmli UK. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 227(5259):680-5.

Lakowicz, JR. (1983) Principles of Fluorescence Spectroscopy. Plenum, New York.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al. (2001) Initial sequencing and analysis of the human genome. Nature. 409(6822):860-921.

Lange LG, Riordan JF, Vallee BL, (1974) Functional arginyl residues as NADH binding sites of alcohol dehydrogenases. Biochemistry 13(21):4361-4370.

Ludden PW. (1994) Reversible ADP-ribosylation as a mechanism of enzyme regulation in procaryotes. Mol Cell Biochem. 138(1-2):123-9.

Lupi R, Corda D, Di Girolamo M. (2000) Endogenous ADP-ribosylation of the G protein beta subunit prevents the inhibition of type 1 adenylyl cyclase. J Biol Chem. 275(13):9418-24.

Margalith P, Pagani H. (1961) Rifomycin. XIV. Production of rifomycin B. Appl Microbiol. 9:325-34.

Mavalan P, Johnson WC. (1983) Sensitivity of circular dichroism to protein tertiary structure class. Nature 305, 831 – 832.

Merrick MJ. (1976) A morphological and genetic mapping study of bald colony mutants of Streptomyces coelicolor. J Gen Microbiol. 96(2):299-315.

Mindlin SZ, Ilyina TS, Voeykova TA, Velkov VV. (1972) Genetical analysis of rifampicin resistant mutants of *E. coli* K 12. Mol Gen Genet. 115(2):115-21.

Mukherjee R, Chatterji D. (2005) Evaluation of the role of sigma B in *Mycobacterium smegmatis*. Biochem Biophys Res Commun. 338(2):964-72.

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction.. Biotechnology. 24:17-27.

Naas T, Mikami Y, Imai T, Poirel L, Nordmann P. (2001) Characterization of In53, a class 1 plasmid- and composite transposon-located integron of *Escherichia coli* which carries an unusual array of gene cassettes. J Bacteriol. 183(1):235-49.

Newton JA Jr, Weiss PJ, Bowler WA, Oldfield EC 3rd. (1993) Soft-tissue infection due to *Mycobacterium smegmatis*: report of two cases. Clin Infect Dis. 16(4):531-3.

Nodwell JR, McGovern K, Losick R. (1996) An oligopeptide permease responsible for the import of an extracellular signal governing aerial mycelium formation in *Streptomyces coelicolor*. Mol Microbiol. 22(5):881-93.

O'Donnell SM, Janssen GR. (2001) The initiation codon affects ribosome binding and translational efficiency in Escherichia coli of cI mRNA with or without the 5' untranslated leader. J Bacteriol. 183(4):1277-83.

Ojha AK, Mukherjee TK, Chatterji D. (2000) High intracellular level of guanosine tetraphosphate in *Mycobacterium smegmatis* changes the morphology of the bacterium. Infect Immun. 68(7):4084-91.

Panja S, Aich P, Jana B, Basu T. (2008) Plasmid DNA binds to the core oligosaccharide domain of LPS molecules of *E. coli* cell surface in the CaCl2-mediated transformation process. Biomacromolecules. 9(9):2501-9.

Paustian ML, Kapur V & Bannantine JP (2005) Comparative genomic hybridizations reveal genetic regions within the Mycobacterium avium complex that are divergent

from Mycobacterium avium subspecies paratuberculosis isolates. J Bacteriol 187: 2406–2415.

Pennekamp A, Pfyffer GE, Wüest J, George CA, Ruef C. (1997) *Mycobacterium smegmatis* infection in a healthy woman following a facelift: case report and review of the literature. Ann Plast Surg. 39(1):80-3.

Perkins SJ.(1986) Protein volumes and hydration effects. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. Eur J Biochem. 157(1):169-80.

Petit JF, Adam A, Wietzerbin-Falszpan J, Lederer E, Ghuysen JM. Chemical structure of the cell wall of *Mycobacterium smegmatis*. I. Isolation and partial characterization of the peptidoglycan. Biochem Biophys Res Commun. 22;35(4):478-85.

Puglia AM, Cappelletti E. (1984) A bald superfertile U.V.-resistant strain in *Streptomyces coelicolor* A3(2). Microbiologica. 7(3):263-6.

Puhaca, S. (2004) Purification and molecular characterization of the ADP-Ribosyltransferase responsible for ribosylative inactivation of the rifampicin. [MSc Thesis] University of the Witwatersrand.

Quan, S. (1997) Cloning and characterization of rifampicin inactivation genes from *Mycobacterium* and closely related organisms. [PhD Thesis] University of the Witwatersrand.

Quan S, Venter H, Dabbs ER. (1997) Ribosylative inactivation of rifampin by *Mycobacterium smegmatis* is a principal contributor to its low susceptibility to this antibiotic. Antimicrob Agents Chemother. 41(11):2456-60.

Quan S, Imai T, Mikami Y, Yazawa K, Dabbs ER, Morisaki N, Iwasaki S, Hashimoto Y, Furihata K. (1999) ADP-ribosylation as an intermediate step in inactivation of rifampin by a mycobacterial gene. Antimicrob Agents Chemother. 43:181-184

Reddy P, Peterkofsky A, McKenney K. (1985) Translational efficiency of the *Escherichia coli* adenylate cyclase gene: mutating the UUG initiation codon to GUG or AUG results in increased gene expression. Proc Natl Acad Sci U S A. 82(17):5656-60.

Ringquist, S., S. Shinedling, D. Barrick, L. Green, J. Binkley, G. D. Stormo, and L. Gold. (1992). Translation initiation in *Escherichia coli*: sequences within the ribosome-binding site. Mol. Microbiol. 6:1219-1229.

Sensi P. (1983) History of the development of rifampin. Rev Infect Dis. 5 Suppl 3:S402-6.

Sensi P, Timbal MT, Maffii G. (1960) Rifomycin IX. Two new antibiotics of rifomycin family: rifomycin S and rifomycin SV. Preliminary report. Experientia. 15;16:412.

Seyler C, Toure S, Messou E, Bonard D, Gabillard D, Anglaret X. (2005) Risk factors for active tuberculosis after antiretroviral treatment initiation in Abidjan. Am J Respir Crit Care Med. 172(1):123-7

Shafran SD, Singer J, Zarowny DP, Phillips P, Salit I, Walmsley SL, Fong IW, Gill MJ, Rachlis AR, Lalonde RG, Fanning MM, Tsoukas CM. (1996) A comparison of two regimens for the treatment of *Mycobacterium avium* complex bacteremia in AIDS: rifabutin, ethambutol, and clarithromycin versus rifampin, ethambutol, clofazimine, and ciprofloxacin. Canadian HIV Trials Network Protocol 010 Study Group. N Engl J Med. 8;335(6):377-83.

Shall S. (1995) ADP-ribosylation reactions. Biochimie. 77(5):313-8.

Shima J, Penyige A, Ochi K. (1996) Changes in patterns of ADP-ribosylated proteins during differentiation of *Streptomyces coelicolor* A3(2) and its development mutants. J Bacteriol. 178(13):3785-90.

Shupp T, Hutter R, Hopwood DA (1975) Genetic recombination in *Nocardia mediterranei*. J Bacteriol. 121(1): 128-136.

Snapper SB, Melton RE, Mustafa S, Kieser T, Jacobs WR Jr. (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. Mol Microbiol. 4(11):1911-9.

Springer B, Wu WK, Bodmer T, Haase G, Pfyffer GE, Kroppenstedt RM, Schröder KH, Emler S, Kilburn JO, Kirschner P, Telenti A, Coyle MB, Böttger EC. (1996) Isolation and characterization of a unique group of slowly growing mycobacteria: description of *Mycobacterium lentiflavum* sp. nov. J Clin Microbiol. 34(5):1100-7.

Stender W, Scheit KH. (1977) Studies of the topography of the binding site of DNA-dependent RNA polymerase from *Escherichia coli* for the antibiotic rifamycin SV. Eur J Biochem. 15;76(2):591-600.

Sugawara K, Dohmae N, Kasai K, Saido-Sakanaka H, Okamoto S, Takio K, Ochi K. (2002) Isolation and identification of novel ADP-ribosylated proteins from *Streptomyces coelicolor* A3(2). Biosci Biotechnol Biochem. 66(10):2292-6.

Telenti A, Imboden P, Marchesi F, Schmidheini T, Bodmer T. (193) Direct, automated detection of rifampin-resistant *Mycobacterium tuberculosis* by polymerase chain reaction and single-strand conformation polymorphism analysis. Antimicrob Agents Chemother. 37(10):2054-8.

Tribuddharat C, Fennewald M. (1999) Integron-mediated rifampin resistance in *Pseudomonas aeruginosa*. Antimicrob Agents Chemother. 43(4):960-2.

Uchida T, Gill DM, Pappenheimer AM Jr. (1971) Mutation in the structural gene for diphtheria toxin carried by temperate phage. Nat New Biol. 233(35):8-11.

van Wezel, G. P., White, J., Bibb, M. J., and Postma, P. W., The *malEFG* gene cluster of *Streptomyces coelicolor* A3(2): characterization, disruption and transcriptional analysis. *Mol. Gen. Genet.*, 254, 604-608 (1997)

Wade C. Winkler, Ronald R. Breaker (2005) Regulation of bacterial gene expression by riboswitches. Annual Review of Microbiology 59:487-517.

Wallace RJ Jr, Nash DR, Tsukamura M, Blacklock ZM, Silcox VA. (1988) Human disease due to *Mycobacterium smegmatis*. J Infect Dis. 158(1):52-9.

Wang L, Slayden RA, Barry CE 3rd, Liu J. (2000) Cell wall structure of a mutant of *Mycobacterium smegmatis* defective in the biosynthesis of mycolic acids. J Biol Chem. 10;275(10):7224-9.

Watanabe K, Rude MA, Walsh CT, Khosla C. (2003) Engineered biosynthesis of an ansamycin polyketide precursor in *Escherichia coli.* Proc Natl Acad Sci U S A. 19;100(17):9774-8.

Wehrli W. (1980) Rifampin: mechanisms of action and resistance. Rev Infect Dis. 5 Suppl 3:S407-11.

Williams DL, Waguespack C, Eisenach K, Crawford JT, Portaels F, Salfinger M, Nolan CM, Abe C, Sticht-Groh V, Gillis TP. (1994) Characterization of rifampin-resistance in pathogenic mycobacteria. Antimicrob Agents Chemother. 38(10):2380-6.

Wood CR, Boss MA, Patel TP, Emtage JS. (1984) The influence of messenger RNA secondary structure on expression of an immunoglobulin heavy chain in *Escherichia coli*. Nucleic Acids Res. 12(9):3937-50.

Woody RW. Circular dichroism. (1995) Methods Enzymol. 246:34-71.
Yang JT, Wu CS, Martinez HM. (1986) Calculation of protein conformation from circular dichroism. Methods Enzymol. 130:208-69.

Wu, H. C., Biosynthesis of Lipoproteins. In "*Esherichia coli* and *Salmonella*: Cellular and Molecular Biology, Second Edition, volume 1", eds. Neidhardt, F. C., Curtiss, R., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., ASM Press, Washington, D.C., pp. 1005-1014 (1996)

Yang JT, Wu CS, Martinez HM. (1986) Calculation of protein conformation from circular dichroism. Methods Enzymol. 130:208-69.

Yates SP, Jørgensen R, Andersen GR, Merrill AR. (2006) Stealth and mimicry by deadly bacterial toxins. Trends Biochem Sci. 31(2):123-33.

Yazawa K, Mikami Y, Maeda A, Morisaki N, Iwasaki S. (1994) Phosphorylative inactivation of rifampicin by *Nocardia otitidiscaviarum*. J Antimicrob Chemother. 33(6):1127-35.

Zuker M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31 (13), 3